EMBEDDING SPANNING BIPARTITE GRAPHS OF SMALL BANDWIDTH

FIACHRA KNOX AND ANDREW TREGLOWN

Abstract. Böttcher, Schacht and Taraz [6] gave a condition on the minimum degree of a graph $G$ on $n$ vertices that ensures $G$ contains every $r$-chromatic graph $H$ on $n$ vertices of bounded degree and of bandwidth $o(n)$, thereby proving a conjecture of Bollobás and Komlós [15]. We strengthen this result in the case when $H$ is bipartite. Indeed, we give an essentially best-possible condition on the degree sequence of a graph $G$ on $n$ vertices that forces $G$ to contain every bipartite graph $H$ on $n$ vertices of bounded degree and of bandwidth $o(n)$. This also implies an Ore-type result. In fact, we prove a much stronger result where the condition on $G$ is relaxed to a certain robust expansion property. Our result also confirms the bipartite case of a conjecture of Balogh, Kostochka and Treglown [1] concerning the degree sequence of a graph which forces a perfect $H$-packing.

1. Introduction

A central problem in graph theory is to establish conditions on a graph $G$ which ensure that $G$ contains another graph $H$ as a spanning subgraph. Perhaps the best-known example of such a problem is when $H$ is a Hamilton cycle. Dirac’s theorem [10] states that any graph $G$ on $n$ vertices with minimum degree $\delta(G) \geq n/2$ contains a Hamilton cycle. The Pósa-Seymour conjecture (see [11] and [24]) states that any graph $G$ on $n$ vertices with $\delta(G) \geq rn/(r + 1)$ contains the $r$th power of a Hamilton cycle. (The $r$th power of a Hamilton cycle $C$ is obtained from $C$ by adding an edge between every pair of vertices of distance at most $r$ on $C$.) Komlós, Sárközy and Szemerédi [17] proved this conjecture for sufficiently large graphs.

There has also been significant attention on establishing minimum degree conditions which ensure a graph $G$ contains a perfect $H$-packing: Given a graph $H$, a perfect $H$-packing in a graph $G$ is a collection of vertex-disjoint copies of $H$ which covers all the vertices in $G$. (Perfect $H$-packings are also referred to as $H$-factors or perfect $H$-tilings.) A seminal result in the area is the Hajnal-Szemerédi theorem [12] which states that every graph $G$ whose order $n$ is divisible by $r$ contains a perfect $K_r$-packing provided that $\delta(G) \geq (r - 1)n/r$. (Corrádi and Hajnal [9] had earlier proved this result in the case when $r = 3$.) Notice that in the case when $r + 1$ divides $|G|$, a necessary condition for a graph $G$ to contain the $r$th power of a Hamilton cycle is that $G$ contains a perfect $K_{r+1}$-packing. Thus, the Pósa-Seymour conjecture implies the Hajnal-Szemerédi theorem. Kühn and Osthus [19, 20] characterised, up to an additive constant, the minimum degree which ensures a graph $G$ contains a perfect $H$-packing for an arbitrary graph $H$.

It is desirable to find conditions that ensure a graph $G$ contains $H$ as a spanning subgraph where $H$ is any graph from a large collection of graphs. That is, rather than finding individual results for specific graphs $H$, one seeks more general, wide-reaching results. A graph $H$ on $n$ vertices is said to have bandwidth at most $b$, if there exists a labelling of the vertices of $H$ by the numbers $1, \ldots, n$ such that for every edge $ij \in E(H)$ we have $|i - j| \leq b$. Clearly every graph $H$ has bandwidth at most $|H| - 1$. Thus, a perfect $H$-packing has bandwidth at most $|H| - 1$. Further, a Hamilton cycle has bandwidth 2, and in general the $r$th power of a Hamilton cycle has bandwidth at most $2r$. Böttcher et al. [4] proved that every planar graph $H$ on $n$ vertices with bounded maximum degree has bandwidth at most $O(n/\log n)$.

Date: January 5, 2012.
The following result of Böttcher, Schacht and Taraz [6] gives a condition on the minimum degree of a graph $G$ on $n$ vertices that ensures $G$ contains every $r$-chromatic graph on $n$ vertices of bounded degree and of bandwidth $o(n)$, thereby proving a conjecture of Bollobás and Komlós [15].

**Theorem 1** (Böttcher, Schacht and Taraz [6]). Given any $r, \Delta \in \mathbb{N}$ and any $\gamma > 0$, there exist constants $\beta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds. Suppose that $H$ is an $r$-chromatic graph on $n \ge n_0$ vertices with $\Delta(H) \le \Delta$ and bandwidth at most $\beta n$. If $G$ is a graph on $n$ vertices with

$$\delta(G) \ge \left(\frac{r-1}{r} + \gamma\right)n,$$

then $G$ contains a copy of $H$.

Prior to the proof of Theorem 1, Hán [13] proved the case when $H$ is bipartite and Böttcher, Schacht and Taraz [5] proved the case when $\chi(H) = 3$. In this paper our focus is on strengthening Theorem 1 in the case when $H$ is bipartite.

1.1. **Degree sequence conditions.** Dirac’s theorem and the Hajnal-Szemerédi theorem are best possible in the sense that the minimum degree conditions in both these results cannot be lowered. However, this does not mean that one cannot strengthen these results. Indeed, Chvátal [8] gave a condition on the degree sequence of a graph which ensures Hamiltonicity: Suppose that the degrees of the graph $G$ are $d_1 \le \cdots \le d_n$. If $n \ge 3$ and $d_i \ge i+1$ or $d_{n-i} \ge n-i$ for all $i < n/2$ then $G$ is Hamiltonian. Notice that Chvátal’s theorem is much stronger than Dirac’s theorem since it allows for almost half of the vertices of $G$ to have degree less than $n/2$.

Balogh, Kostochka and Treglown [1] proposed the following two conjectures concerning the degree sequence of a graph which forces a perfect $H$-packing.

**Conjecture 2** (Balogh, Kostochka and Treglown [1]). Let $n, r \in \mathbb{N}$ such that $r$ divides $n$. Suppose that $G$ is a graph on $n$ vertices with degree sequence $d_1 \le \cdots \le d_n$ such that:

- $d_i \ge (r-2)n/r + i$ for all $i < n/r$;
- $d_{n/r+1} \ge (r-1)n/r$.

Then $G$ contains a perfect $K_r$-packing.

Note that Conjecture 2, if true, is much stronger than the Hajnal-Szemerédi theorem since the degree condition allows for $n/r$ vertices to have degree less than $(r-1)n/r$.

**Conjecture 3** (Balogh, Kostochka and Treglown [1]). Suppose $\gamma > 0$ and $H$ is a graph with $\chi(H) = r$. Then there exists an integer $n_0 = n_0(\gamma, H)$ such that the following holds. If $G$ is a graph whose order $n \ge n_0$ is divisible by $|H|$, and whose degree sequence $d_1 \le \cdots \le d_n$ satisfies

- $d_i \ge (r-2)n/r + i + \gamma n$ for all $i < n/r$,

then $G$ contains a perfect $H$-packing.

In this paper we prove the following result which gives a condition on the degree sequence of a graph $G$ on $n$ vertices that ensures $G$ contains every bipartite graph on $n$ vertices of bounded degree and of bandwidth $o(n)$.

**Theorem 4.** Given any $\Delta \in \mathbb{N}$ and any $\gamma > 0$, there exists constants $\beta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds. Suppose that $H$ is a bipartite graph on $n \ge n_0$ vertices with $\Delta(H) \le \Delta$ and bandwidth at most $\beta n$. Let $G$ be a graph on $n$ vertices with degree sequence $d_1 \le \cdots \le d_n$. If

- $d_i \ge i + \gamma n$ or $d_{n-i+1} \ge n-i$ for all $i < n/2$,

then $G$ contains a copy of $H$. 
The degree sequence condition in Theorem 4 is similar to that in Chvátal’s theorem, except that now we have two error terms in the condition. Notice that Theorem 4 is much stronger than the bipartite case of Theorem 1. Furthermore, in the case when \( r = 2 \), Conjecture 3 is implied by Theorem 4.

Theorem 4 is, up to the error terms, best-possible for many graphs \( H \). Indeed, suppose that \( H \) is a bipartite graph on an even number \( n \) of vertices that contains a perfect matching. Suppose that \( m \in \mathbb{N} \) such that \( m < n/2 \). Let \( G \) be a graph on \( n \) vertices with vertex classes \( V_1, V_2, V_3 \) of sizes \( m, m - 1 \) and \( n - 2m + 1 \) respectively and whose edge set contains all possible edges except for those in \( V_1 \) and between \( V_1 \) and \( V_3 \). Let \( d_1 \leq \cdots \leq d_n \) denote the degree sequence of \( G \). Then

- \( d_i \geq i - 1 \) and \( d_{n-i+2} \geq n-i \) for all \( i < n/2 \),

but since \( |V_1| > |V_2| \), \( G \) does not contain a perfect matching and therefore \( H \).

1.2. Ore-type degree conditions. Ore-type degree conditions consider the sum of the degrees of non-adjacent vertices of a graph. The name comes from Ore’s theorem [23], which states that a graph \( G \) of order \( n \geq 3 \) contains a Hamilton cycle if \( d(x) + d(y) \geq n \) for all non-adjacent \( x \neq y \in V(G) \). Recently, Châu [7] proved an Ore-type analogue of the Pósa-Seymour conjecture in the case of a Hamilton cycle (i.e. when \( r = 2 \)).

The following Ore-type result of Kierstead and Kostochka [14] implies the Hajnal-Szemerédi theorem: Let \( n, r \in \mathbb{N} \) such that \( r \) divides \( n \). Suppose that \( G \) is a graph on \( n \) vertices such that for all non-adjacent \( x \neq y \in V(G) \), \( d(x) + d(y) \geq 2(r-1)n/r-1 \). Then \( G \) contains a perfect \( K_r \)-packing. Kühn, Osthus and Treglown [21] characterised, asymptotically, the Ore-type degree condition which ensures a graph \( G \) contains a perfect \( H \)-packing for an arbitrary graph \( H \).

It is natural to seek an Ore-type analogue of Theorem 1. The following result provides such an analogue in the case when \( H \) is bipartite.

**Theorem 5.** Given any \( \Delta \in \mathbb{N} \) and any \( \gamma > 0 \), there exists constants \( \beta > 0 \) and \( n_0 \in \mathbb{N} \) such that the following holds. Suppose that \( H \) is a bipartite graph on \( n \geq n_0 \) vertices with \( \Delta(H) \leq \Delta \) and bandwidth at most \( \beta n \). Let \( G \) be a graph on \( n \) vertices such that, for all non-adjacent \( x \neq y \in V(G) \),

\[
d(x) + d(y) \geq (1 + \gamma)n.
\]

Then \( G \) contains a copy of \( H \).

In Section 2.2 we show that Theorem 5 is a direct consequence of Theorem 4. Note that Theorem 5 is best-possible up to the error term for bipartite graphs \( H \) on \( n \) vertices which do not contain an isolated vertex. Indeed, let \( G \) consist of a copy of \( K_{n-1} \) and an isolated vertex. Then \( G \) does not contain \( H \) but \( d(x) + d(y) = n - 2 \) for all non-adjacent \( x \neq y \in V(G) \).

In light of Theorem 5, we propose the following Ore-type analogue of Theorem 1.

**Conjecture 6.** Given any \( r, \Delta \in \mathbb{N} \) and any \( \gamma > 0 \), there exists constants \( \beta > 0 \) and \( n_0 \in \mathbb{N} \) such that the following holds. Suppose that \( H \) is an \( r \)-chromatic graph on \( n \geq n_0 \) vertices with \( \Delta(H) \leq \Delta \) and bandwidth at most \( \beta n \). Let \( G \) be a graph on \( n \) vertices such that, for all non-adjacent \( x \neq y \in V(G) \),

\[
d(x) + d(y) \geq 2 \left( \frac{r-1}{r} + \gamma \right) n.
\]

Then \( G \) contains a copy of \( H \).

If true, Conjecture 6 is stronger than Theorem 1. Böttcher and Müller (see [3]) have announced a proof of the conjecture in the case when \( r = 3 \).
1.3. **Robustly expanding graphs.** An important and well-studied notion in graph theory is graph expansion. We will consider the following stronger notion of ‘robust expansion’. Roughly speaking, a graph $G$ on $n$ vertices is a robust expander if, for every ‘reasonably sized’ set $S \subseteq V(G)$, $G$ contains at least $|S| + o(n)$ vertices that are adjacent to ‘many’ vertices in $S$. More formally, let $0 < \nu \leq \tau < 1$. Suppose that $G$ is a graph on $n$ vertices and $S \subseteq V(G)$. Then the $\nu$-robust neighbourhood $RN_{\nu,G}(S)$ of $S$ is the set of vertices $v \in V(G)$ such that $|N(v) \cap S| \geq \nu n$. We say that $G$ is a robust $(\nu, \tau)$-expander if every $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$ satisfies $|RN_{\nu,G}(S)| \geq |S| + \nu n$.

The notion of robustly expanding (di)graphs was first introduced by Kühn, Osthus and Treglown in [22]. The following result is an immediate consequence of Theorem 16 from [22].

**Theorem 7** (Kühn, Osthus and Treglown [22]). Given positive constants $\nu \leq \tau \ll \eta < 1$ there exists a positive integer $n_0$ such that the following holds. Let $G$ be a graph on $n \geq n_0$ vertices with $\delta(G) \geq \eta n$ which is a robust $(\nu, \tau)$-expander. Then $G$ contains a Hamilton cycle.

(Here, and throughout the paper, we write $0 < \nu \ll \tau \ll \eta$ to mean that we can choose the constants $\nu, \tau, \eta$ from right to left. More precisely, there are increasing functions $f$ and $g$ such that, given $\eta$, whenever we choose some $\tau \leq f(\eta)$ and $\nu \leq g(\tau)$, all calculations needed in our proof are valid. Hierarchies with more constants are defined in the obvious way.)

We will use Theorem 7 to prove the following result concerning embedding bipartite graphs of small bandwidth.

**Theorem 8.** Given $\Delta \in \mathbb{N}$ and positive constants $\nu \leq \tau \ll \eta < 1$ there exists constants $\beta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds. Suppose that $H$ is a bipartite graph on $n \geq n_0$ vertices with $\Delta(H) \leq \Delta$ and bandwidth at most $\beta n$. Let $G$ be a graph on $n$ vertices with $\delta(G) \geq \eta n$ which is a robust $(\nu, \tau)$-expander. Then $G$ contains a copy of $H$.

In Section 2.2 we show that Theorem 8 implies Theorem 4 and that Theorem 4 implies Theorem 5. Thus, we only prove Theorem 8 directly.

Note that Theorem 8 is very general in the sense that it allows for the graph $G$ to have small minimum degree (although $\delta(G)$ must be linear). Furthermore, there are examples of graphs $G$ that satisfy the hypothesis of Theorem 8 and whose maximum degree is also small. Indeed, let $0 < \nu \ll \tau \ll \eta < 1$ such that $1/\eta$ is an odd integer. Further choose $n \in \mathbb{N}$ such that $\eta n \in \mathbb{N}$. Define $G$ to be the blow-up of a cycle on $1/\eta$ vertices, such that each vertex class of $G$ contains $\eta n$ vertices. Thus, $|G| = n$ and $\delta(G) = \Delta(G) = 2\eta n$. It is easy to check that $G$ is a robust $(\nu, \tau)$-expander. Given constants $0 < \nu \ll \tau \ll p < 1$, with high probability $G(n, p)$ is a robust $(\nu, \tau)$-expander with minimum degree at least $np/2$ and maximum degree at most $2np$.

2. **Notation and preliminaries**

2.1. **Notation.** Throughout this paper we omit floors and ceilings whenever this does not affect the argument. We write $|G|$ for the order of a graph $G$, $\delta(G)$ and $\Delta(G)$ for its minimum and maximum degrees respectively and $\chi(G)$ for its chromatic number. The degree of a vertex $x \in V(G)$ is denoted by $d(x)$ and its neighbourhood by $N(x)$. Given $S \subseteq V(G)$ we define $N(S) := \bigcup_{v \in S} N(v)$.

Given disjoint $A, B \subseteq V(G)$ the number of edges with one endpoint in $A$ and one endpoint in $B$ is denoted by $e_{G}(A, B)$. We write $(A, B)_{G}$ for the bipartite subgraph of $G$ with vertex classes $A$ and $B$ whose edges are precisely those edges in $G$ with one endpoint in $A$ and the other in $B$. Often we will write $(A, B)$, for example, if this is unambiguous.

2.2. **Degree sequence and Ore-type conditions forcing robust expansion.** The following result is an immediate consequence of Lemma 13 from [22].
Lemma 9 ([22]). Given positive constants \( \tau \ll \eta < 1 \) there exists an integer \( n_0 \) such that whenever \( G \) is a graph on \( n \geq n_0 \) vertices with
\[
d_i \geq i + \eta n \text{ or } d_{n-i-\eta n} \geq n - i \text{ for all } i < n/2,
\]
then \( \delta(G) \geq \eta n \) and \( G \) is a robust \((\tau^2, \tau)\)-outerexpander.

Notice that Lemma 9 together with Theorem 8 implies Theorem 4. We now show that Theorem 4 implies Theorem 5.

Lemma 10. Let \( \gamma > 0 \). Suppose \( G \) is a graph on \( n \) vertices such that, for all non-adjacent \( x \neq y \in V(G) \),
\[
d(x) + d(y) \geq (1 + 2\gamma)n.
\]
Let \( d_1 \leq \cdots \leq d_n \) denote the degree sequence of \( G \). Then
\[
d_i \geq i + \eta n \text{ or } d_{n-i-\eta n} \geq n - i \text{ for all } i < n/2.
\]

Proof. Firstly note that for \( (1 - \gamma)n/2 \leq i < n/2 \) we wish to show that either \( d_{n-i'-\gamma n} \geq n - i' \) or \( d_{i'} \geq i' + \gamma n \), where \( i' := n - i - \gamma n \). Notice that \( n/2 - \gamma n < i' \leq n/2 - \gamma n/2 \). Thus, it suffices to only consider \( i \) such that \( 1 \leq i < (1 - \gamma)n/2 \).

Suppose there is some \( 1 \leq i \leq (1 - \gamma)n/2 \) such that the statement does not hold. Then there is a set \( A \) of \( i \) vertices, each of degree less than \( i + \gamma n \leq n/2 + \gamma n/2 \). So for any \( x, y \in A \), \( d(x) + d(y) < (1 + 2\gamma)n \) and hence \( G[A] \) is a clique. Set \( B := V(G) \setminus A \). Note that \( c_G(A, B) < (\gamma n + 1)i \). Hence, there is a vertex \( x \in B \) that receives less than \( \min\{\gamma n + 1, i\} \) edges from \( A \). Therefore, there is a vertex \( y \in A \) such that \( xy \notin E(G) \). Thus, \( d(x) + d(y) < (n - i - 1 + \gamma n + 1) + (i + \gamma n) \leq (1 + 2\gamma)n \), contradicting our assumption.

3. OUTLINE OF THE PROOF OF THEOREM 8

3.1. Proof overview. The overall strategy is similar to that of the proof of Theorem 1 in [6]. Indeed, as in [6] the proof is split into two main lemmas; the Lemma for \( G \) and the Lemma for \( H \). However, many of the methods used in [6] break down in our setting so our argument proceeds somewhat differently.

The role of the Lemma for \( G \) (Lemma 23) is to obtain some special structure within \( G \) so that it will be suitable for embedding \( H \) into: We show that \( G \) contains a spanning subgraph \( G' \) which ‘looks’ like the blow-up of a cycle \( C = V_1V_2\ldots V_{2k}V_1 \). More precisely, there is a partition \( V_1, \ldots, V_{2k} \) of \( V(G) \) such that:

(i) \( (V_{2i-1}, V_{2i})_{G'} \) is a ‘super-regular’ pair of density at least \( d > 0 \) for each \( 1 \leq i \leq k \);

(ii) \( (V_{2i}, V_{2i+1})_{G'} \) is an ‘\( \varepsilon \)-regular’ pair of density at least \( d \) for each \( 1 \leq i \leq k \).

Furthermore, there are even integers \( 1 \leq i_1 \neq j_1 \leq 2k \) such that:

(iii) \( (V_{i_1}, V_{j_1})_{G'} \) is ‘\( \varepsilon \)-regular’ with density at least \( d \).

(And \( V_{i_1}V_{j_1} \) can be though of as a chord of \( C \).) Crucially, this partition is ‘robust’ in the sense that one can modify the sizes of each partition class \( V_i \) somewhat without destroying the properties (i)–(iii). (This is made precise by the Mobility lemma given in Section 6.)

Set \( c := V_{i_1}V_{j_1} \). The role of the Lemma for \( H \) (Lemma 25) is to construct a graph homomorphism \( f \) from \( H \) to \( C \cup \{c\} \) in such a way that ‘most’ of the edges of \( H \) are mapped to edges of the form \( V_{2i-1}V_{2i} \) for some \( i \). (Recall that these are the edges which correspond to super-regular pairs in \( G' \).) The homomorphism \( f \) is such that every \( V_i \in C \) receives roughly \( |V_i| \) vertices of \( H \). So \( f \) can be viewed as a ‘guide’ as to which vertex class \( V_i \subseteq V(G) \) each vertex from \( H \) is embedded into.

In particular, since the partition \( V_1, \ldots, V_{2k} \) is ‘robust’, we can alter the sizes of the classes \( V_i \) such that (i)–(iii) still hold and so that now \( |f^{-1}(V_i)| = |V_i| \) for all \( i \). Properties (i)–(iii) then allow us
to apply the Blow-up lemma [16] to embed $H$ into $G'$ and thus $G$. (Actually we apply a result from [2] which is a consequence of the Blow-up lemma.)

3.2. Techniques for the Lemma for $G$. In order to obtain the partition $V_1, \ldots, V_{2k}$ of $V(G)$ we modify a partition $V'_0, V'_1, \ldots, V'_{2k}$ obtained by applying Szemerédi’s Regularity lemma [26] to $G$. Roughly speaking, $V'_1, \ldots, V'_{2k}$ will satisfy (i)–(iii). Thus, we need to redistribute the vertices from $V'_0$ into the other vertex classes whilst retaining these properties. We also require our partition $V_1, \ldots, V_{2k}$ to satisfy $|V_{2i-1}| \approx |V_{2i}|$ for each $1 \leq i \leq k$. So we need to redistribute vertices in a ‘balanced’ way. In the Lemma for $G$ in [6], the minimum degree condition of Theorem 1 is heavily relied on to achieve this. However, our graph $G$ may have very small minimum degree. So instead we introduced the notion of a ‘shifted $M$-walk’ to help us redistribute vertices: Given a perfect matching $M$ in a graph $G$ a shifted $M$-walk is a walk whose edges alternate between edges of $M$ and edges of $R \setminus M$ (see Section 5.1 for the precise definition). Since $G$ is a robust expander, we can find short shifted $M$-walks in a reduced graph $R$ of $G$. These walks act as a ‘guide’ as to how we redistribute vertices amongst the vertex classes.

3.3. Techniques for the Lemma for $H$. In [6] the techniques used are actually strong enough to prove a more general result than Theorem 1 (and so Theorem 1 is not proved directly). For example, in the case when $r = 2$, their result concerns not only bipartite $H$ but also a special class of 3-colourable graphs $H$ where the third colour class is very small (see Theorem 2 in [6] for precise details). One example of such a graph $H$ is a Hamilton cycle $C'$ with a chord between two vertices of distance 2 on $C'$. $H$ is 3-colourable and has bounded bandwidth. However, $H$ cannot be embedded into every graph $G$ satisfying the hypothesis of Theorem 8. Indeed, consider the graph $G$ defined at the end of Section 1.3.

In particular, this means we have to approach the proof of the Lemma for $H$ differently: Since $H$ has bandwidth $o(n)$ we can chop $V(H)$ into small linear sized segments $A_1, B_1, \ldots, A_m, B_m$ where all the edges of $H$ lie in pairs of the form $(A_i, B_i)_H$ and $(B_i, A_{i+1})_H$ and such that $A := \cup_{i=1}^m A_i$ and $B := \cup_{i=1}^m B_i$ are the colour classes of $H$. Ideally we would want to construct $f$ to map the vertices of $A_1$ into $V_1$, the vertices of $B_1$ into $V_2$ and so on, continuing around $C$ many times until all the vertices have been assigned. However, since $|A|$ and $|B|$ may vary widely, this would map vertices in an unbalanced way. That is, the total number of vertices mapped to ‘odd’ classes $V_{2i-1}$ would differ widely from the total number of vertices mapped to ‘even’ classes $V_{2i}$. We get around this problem by using the chord $c = V_{1i}V_{1j}$ to ‘flip’ halfway in the process. So after this, vertices from the $B_i$ are mapped to ‘odd’ classes $V_{2i-1}$ and vertices from the $A_i$ are mapped to the ‘even’ classes $V_{2i}$. We also ‘randomise’ part of the mapping procedure to ensure that the number of vertices of $H$ assigned to each $V_i$ is approximately $|V_i|$.

4. The Regularity lemma

In the proof of the Lemma for $G$ (Lemma 23) we will use Szemerédi’s Regularity lemma [26]. In this section we will introduce all the information we require about this result. To do this, we firstly introduce some more notation. The density of a bipartite graph $G$ with vertex classes $A$ and $B$ is defined to be

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

Given any $\epsilon, d > 0$, we say that $G$ is $(\epsilon, d)$-regular if $d(A, B) \geq d$ and, for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$, we have $|d(A, B) - d(X, Y)| < \epsilon$. We say that $G$ is $(\epsilon, d)$-super-regular if additionally every vertex $a \in A$ has at least $d|B|$ neighbours in $B$ and every vertex $b \in B$ has at least $d|A|$ neighbours in $A$. We also say that $(A, B)$ is an $(\epsilon, d)$-(super-)regular pair. We will frequently use the following simple fact.

6

Fact 11. Let $\varepsilon, d > 0$. Suppose that $G = (A, B)$ is an $(\varepsilon, d)$-regular pair. Let $B' \subseteq B$ be such that $|B'| \geq \varepsilon |B|$. Then there are at most $\varepsilon |A|$ vertices in $A$ with fewer than $(d - \varepsilon) |B'|$ neighbours in $B'$.

We will also require the next simple proposition which allows us to modify a (super-)regular pair without destroying its (super-)regularity (see e.g., [5, Proposition 8]).

Proposition 12. Let $(A, B)$ be an $(\varepsilon, d)$-regular pair, and let $A'$ and $B'$ be vertex sets with $|A' \Delta A| \leq \alpha |A|$ and $|B' \Delta B| \leq \beta |B|$. Then $(A', B')$ is an $(\varepsilon', d')$-regular pair where $$\varepsilon' := \varepsilon + 3(\sqrt{\alpha} + \sqrt{\beta}) \quad \text{and} \quad d' := d - 2(\alpha + \beta).$$

If, moreover, $(A, B)$ is $(\varepsilon, d)$-super-regular and each vertex in $A'$ has at least $d |B'|$ neighbours in $B'$ and each vertex in $B'$ has at least $d |A'|$ neighbours in $A'$, then $(A', B')$ is $(\varepsilon', d')$-super-regular.

We will use the following degree form of Szemerédi’s Regularity lemma [26] which can be easily derived from the classical version.

Lemma 13 (Regularity lemma). For every $\varepsilon > 0$ and $k_0 \in \mathbb{N}$ there exists $K_0 = K_0(\varepsilon, k_0)$ such that for every $d \in [0, 1]$ and for every graph $G$ on $n \geq K_0$ vertices there exists a partition $V_0, V_1, \ldots, V_k$ of $V(G)$ and a spanning subgraph $G'$ of $G$, such that the following conditions hold:

(i) $k_0 \leq k \leq K_0$,
(ii) $d_{C'}(x) \geq d_C(x) - (d + \varepsilon)n$ for every $x \in V(G)$,
(iii) the subgraph $G'[V_i]$ is empty for all $1 \leq i \leq k$,
(iv) $|V_0| \leq \varepsilon n$,
(v) $|V_1| = |V_2| = \ldots = |V_k|$,
(vi) for all $1 \leq i < j \leq k$ each vertex $(V_i, V_j)_{C'}$ is an $(\varepsilon, d)$-regular pair or $G'[V_i, V_j]$ is empty.

We call $V_1, \ldots, V_k$ clusters, $V_0$ the exceptional set and the vertices in $V_0$ exceptional vertices. We refer to $G'$ as the pure graph. The reduced graph $R$ of $G$ with parameters $\varepsilon, d$ and $k_0$ is the graph whose vertices are $V_1, \ldots, V_k$ and in which $V_iV_j$ is an edge precisely when $(V_i, V_j)_{C'}$ is $(\varepsilon, d)$-regular.

The following result implies that the property of a graph $G$ being a robust expander is ‘inherited’ by the reduced graph $R$ of $G$. It is an immediate consequence of Lemma 14 from [22].

Lemma 14 ([22]). Let $k_0, n_0$ be positive integers and let $\varepsilon, d, \eta, \nu, \tau$ be positive constants such that $1/n_0 \ll \varepsilon \ll d \ll \nu, \tau, \eta < 1$ and such that $k_0 \ll n_0$. Let $G$ be a graph on $n \geq n_0$ vertices with $\delta(G) \geq \eta n$ and such that $G$ is a robust $(\nu, \tau)$-expander. Let $R$ be the reduced graph of $G$ with parameters $\varepsilon, d$ and $k_0$. Then $\delta(R) \geq \eta |R|/2$ and $R$ is a robust $(\nu/2, 2\tau)$-expander.

5. Useful results

5.1. Shifted walks and robust expanders. Let $G$ be a graph containing a perfect matching $M$. A shifted $M$-walk in $G$ with endpoints $a = v_1$ and $b = v_{2\ell}$ is a walk $v_1v_2 \ldots v_{2\ell}$ in $G$ such that $v_{2i}v_{2i+1} \in M$ for every $1 \leq i \leq \ell - 1$ and $v_{2\ell-1}v_{2\ell} \notin M$ for any $1 \leq i \leq \ell$. A shifted $M$-walk is simple if it contains each edge of $M$ at most twice. Note that a path containing a single edge $v_1v_2 \notin M$ is a (simple) shifted $M$-walk for any perfect matching $M$.

Lemma 15. Let $G$ be a graph containing a perfect matching $M$, and let $W$ be a shifted $M$-walk in $G$ with endpoints $a$ and $b$. Then $W$ contains a simple shifted $M$-walk $W'$ with endpoints $a$ and $b$.

Proof. We proceed by induction. Let $W = v_1 \ldots v_{2\ell}$. If $W$ is already simple then we set $W' := W$; otherwise, there exists an edge $xy \in M$ which appears at least three times in $W$. Let the first three appearances of $xy$ be $v_i, v_{i+1}, v_{2i}v_{2i+1}$ and $v_i, v_{i+1}$ (in order). Now each of $v_i$, $v_{i+1}$ and $v_{i+1}$ is either $x$ or $y$, and so without loss of generality we can assume that $v_i = v_{i+1} = x$. Now we have a shorter shifted $M$-walk $v_1 \ldots v_i v_{i+1} v_{2i+1} \ldots v_{2\ell}$.
Lemma 16. Let $M$ be a perfect matching in a graph $G$ and let $A \subseteq V(M)$ be a set containing at most one vertex from each edge of $M$. Suppose that $W$ is a shifted $M$-walk both of whose endpoints lie in $A$. Then $W$ contains a shifted $M$-walk $W'$ such that the endpoints of $W'$ both lie in $A$ and no other vertices of $W'$ lie in $A$.

Proof. Let $W = v_1 \ldots v_{2\ell}$. Let $1 \leq i_1 \leq \ell$ be minimal such that $v_{2i_1} \in A$, and let $1 \leq i_2 \leq i_1$ be maximal such that $v_{2i_2-1} \in A$. Now $v_{2i_2-1}v_{2i_2} \ldots v_{2i_1-1}v_{2i_1}$ is the desired shifted $M$-walk. □

Let $G$ be a graph containing a perfect matching $M$, and let $A \subseteq V(G)$. For each $v \in V(G)$, let $v' \in V(G)$ be the unique vertex such that $vv' \in M$. The shifted $M$-neighbourhood of $A$ is the set $SN_M(A) = \{ v' \mid v \in N(A) \}$. $SN_M^r(A)$ is defined recursively by $SN_M^1(A) := SN_M(A)$ and $SN_M^r(A) := SN_M(SN_M^{r-1}(A))$ for $r \geq 2$.

Lemma 17. Let $0 < \nu < \tau < \eta \ll 1$ be constants. Suppose $G$ is a graph on $n$ vertices with $\delta(G) \geq \eta n$ which is a robust $(\nu, \tau)$-expander, and let $G$ be a perfect matching in $G$. Then for any $a \in V(G), G$ contains a shifted $M$-walk of length at most $3/\nu$ which both starts and finishes at $a$.

Proof. The minimum degree condition implies that $|SN_M(a)| = |N(a)| \geq \eta n \geq \tau n$. Since $G$ is a robust $(\nu, \tau)$-expander,

$$|SN_M^{r-1}(a)| = |N(SN_M^{r-1}(a))| \geq \min \{ |SN_M^{r-1}(a)| + \nu n, (1-\tau+\nu)n \},$$

for all $r \geq 2$. Hence $|SN_M^{1/2\nu}(a)| \geq (\tau + 1/2 - \nu)n$ and so

$$|N(SN_M^{1/2\nu}(a))| \geq (1/2 + \tau)n.$$

Thus, there exists some edge $vv' \in M$ such that both $v$ and $v'$ lie in $N(SN_M^{1/2\nu}(a))$. This implies that there exists a shifted $M$-walk $P$ with endpoints $a$ and $v$ and a shifted $M$-walk $P'$ with endpoints $a$ and $v'$, each of length at most $1/\nu + 1$. Now $P \cup vv' \cup P'$ forms a shifted $M$-walk of length at most $2/\nu + 3 \leq 3/\nu$ which starts and finishes at $a$. □

Lemma 18. Let $0 < \alpha < \nu \leq \tau \ll 1$ be constants. Suppose that $G$ is a graph on $n$ vertices which is a robust $(\nu, \tau)$-expander and let $S \subseteq V(G)$ be a set of size $\alpha n$. Then $G - S$ is a robust $(\nu - \alpha, \tau + \alpha)$-expander.

Proof. Let $G' := G - S$ and $n' := |G'|$. Consider any $A \subseteq V(G')$ such that $(\tau + \alpha)n' \leq |A| \leq (1-\tau-\alpha)n'$. Set $A' := A \cup S$. Then

$$\tau n \leq (\tau + \alpha)n' + \alpha n \leq |A'| \leq (1-\tau-\alpha)n' + \alpha n \leq (1-\tau)n.$$

So $|RN_{\nu,G}(A')| \geq |A'| + \nu n$. Now every vertex of $RN_{\nu,G}(A')$ has at least $\nu n$ neighbours in $A'$ and since $|S| = \alpha n$, at least $(\nu - \alpha)n \geq (\nu - \alpha)n'$ of these must lie in $A$. Hence every vertex of $RN_{\nu,G}(A') \setminus S$ lies in $RN_{\nu - \alpha,G'}(A)$, and so $|RN_{\nu - \alpha,G'}(A)| \geq |A'| + \nu n - |S| \geq |A| + (\nu - \alpha)n'$, as desired. □
5.2. Probabilistic bounds. The following two probabilistic bounds will be used in the proof of the Lemma for \( H \) (Lemma 25).

**Lemma 19** ([18], Lemma 2.1). Suppose that \( 1/k \ll p, (1 - p), \varepsilon \), that \( n \geq k^3 / 6 \), and that \( X \sim \text{Bin}(n, p) \). Then for any \( 0 \leq r \leq k - 1 \),

\[
\frac{1 - \varepsilon}{k} \leq \mathbb{P}(X \equiv r \mod k) \leq \frac{1 + \varepsilon}{k}.
\]

**Lemma 20** ([25], Proposition 1.1). Let \( X_1, \ldots, X_n \) be random variables taking values in \([0, 1]\), such that for each \( 1 \leq k \leq n \),

\[
\mathbb{E}[X_k | X_{k-1}, \ldots, X_1] \leq a_k.
\]

Let \( \mu := \sum_{k=1}^{n} a_k \). Then for any \( 0 < \delta \leq 1 \),

\[
\mathbb{P}\left[ \sum_{k=1}^{n} X_k > (1 + \delta)\mu \right] \leq e^{-\frac{\delta^2 \mu}{4}}.
\]

We also require the following expectation bound.

**Lemma 21.** Suppose that \( X \) and \( Y \) are integer-valued random variables and that \( B \) is an event, such that for each \( x, y \in \mathbb{Z} \), \( \mathbb{P}[X = x | B \cap (Y = y)] = \mathbb{P}[X = x | Y = y] \). Then

\[
\mathbb{E}[X | B] \leq \max_{y \in \mathbb{Z}} \mathbb{E}[X | Y = y].
\]

**Proof.** Note that for each \( x \in \mathbb{Z} \),

\[
\mathbb{P}[X = x | B] = \sum_{y \in \mathbb{Z}} \mathbb{P}[(X = x) \cap (Y = y) | B].
\]

Further,

\[
\mathbb{P}[(X = x) \cap (Y = y) | B] = \frac{\mathbb{P}[(X = x) \cap (Y = y) \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[X = x] \mathbb{P}(Y = y) \cap B]}{\mathbb{P}[Y = y] \cap B]} \cdot \frac{\mathbb{P}[Y = y] \cap B]}{\mathbb{P}[B]}
\]

Hence,

\[
\mathbb{E}[X | B] = \sum_{x \in \mathbb{Z}} x \mathbb{P}[X = x | B] = \sum_{x \in \mathbb{Z}} x \sum_{y \in \mathbb{Z}} \mathbb{P}[X = x | Y = y] \cdot \mathbb{P}[Y = y | B]
\]

\[
= \sum_{y \in \mathbb{Z}} \mathbb{P}[Y = y | B] \sum_{x \in \mathbb{Z}} x \mathbb{P}[X = x | Y = y] = \sum_{y \in \mathbb{Z}} \mathbb{P}[Y = y | B] \cdot \mathbb{E}[X | Y = y]
\]

\[
\leq \max_{y \in \mathbb{Z}} \mathbb{E}[X | Y = y]. \quad \square
\]

6. The Mobility lemma

In order to state our next result we first introduce a slight variant of the notion of a reduced graph. Let \( \varepsilon, \varepsilon', d, d' > 0 \). Suppose that \( G \) is a graph and \( V_1, \ldots, V_k \) is a partition of \( V(G) \). We say that a graph \( R \) is an \((\varepsilon, d, d')\)-reduced graph of \( G \) on \( V_1, \ldots, V_k \) if the following holds:

- \( V(R) = \{V_1, \ldots, V_k\} \);
- If \( V_i V_j \in E(R) \) then \((V_i, V_j)_G \) is an \((\varepsilon, d)\)-regular pair (for all \( 1 \leq i \neq j \leq k \)).
Suppose $V'_1, \ldots, V'_k$ is another partition of $V(G)$. We also say that $R$ is an $(\varepsilon', d')$-reduced graph of $G$ on $V'_1, \ldots, V'_k$ if the following holds:

- If $V_i V_j \in E(R)$ then $(V'_i, V'_j)_G$ is an $(\varepsilon', d')$-regular pair for all $1 \leq i \neq j \leq k$.

Suppose that $V_1, \ldots, V_k$ and $V'_1, \ldots, V'_k$ are both partitions of the vertex set of a graph $G$. Given a cluster $V = V_i$ for some $1 \leq i \leq k$, we will often denote by $V'$ the cluster $V'_i$.

We will apply the next result in the Lemma for $G$ (Lemma 23) so that we can alter a particular partition of a graph $G$ somewhat without destroying the structure of our reduced graph $R$.

**Lemma 22 (Mobility lemma).** Let $k \in \mathbb{N}$, and let $\xi, \varepsilon, \varepsilon', d, d'$ be positive constants such that

$$0 < \xi \ll 1/k \ll \varepsilon \ll \varepsilon' \ll d' \ll d \ll 1.$$  

Suppose $G$ is a graph on $n$ vertices, $A_1, B_1, A_2, B_2, \ldots, A_k, B_k$ is a partition of $V(G)$ such that $|A_i|, |B_i| \geq n/3k$ for all $1 \leq i \leq k$ and $R$ is an $(\varepsilon, d)$-reduced graph on $A_1, B_1, \ldots, A_k, B_k$. Let $(a_i)_{i=1}^k$ and $(b_i)_{i=1}^k$ be integers. Suppose that the following conditions hold:

(i) $R$ contains the Hamilton cycle $C = A_1 B_1 A_2 B_2 \ldots A_k B_k A_1$;

(ii) $R$ contains an edge $A_i A_j$ for some $i \neq j$;

(iii) $R$ contains an edge $B_i B_j$ for some $i \neq j$;

(iv) The pair $(A_i, B_i)_G$ is $(\varepsilon, d)$-super-regular for all $1 \leq i \leq k$;

(v) $|a_i|, |b_i| < \xi n$ for each $1 \leq i \leq k$;

(vi) $\sum_{i=1}^k a_i + \sum_{i=1}^k b_i = 0$;

(vii) $|\sum_{i=1}^k a_i| = |\sum_{i=1}^k b_i| \leq \xi n$.

Then there exists a partition $A'_1, B'_1, A'_2, B'_2, \ldots, A'_k, B'_k$ of $V(G)$ such that $|A'_i| = |A_i| + a_i$ and $|B'_i| = |B_i| + b_i$ for each $1 \leq i \leq k$, $R$ is an $(\varepsilon', d')$-reduced graph of $G$ on $A'_1, B'_1, A'_2, B'_2, \ldots, A'_k, B'_k$, and $(A'_i, B'_i)_G$ is $(\varepsilon', d')$-super-regular for each $1 \leq i \leq k$.

**Proof.** Without loss of generality we may assume that $\sum_{i=1}^k a_i \geq 0$. (As a consequence of this assumption we will in fact only need the edge $B_1 B_2$, and not the edge $A_1 A_{j_1}$.) Note that by (iii) and Fact 11 there are at least $(1-\varepsilon)|B_{j_2}| \gg \xi n$ vertices in $B_{j_2}$ with at least $(d-\varepsilon)|B_{j_2}|$ neighbours in $B_{j_2}$. Pick $\sum_{i=1}^k a_i \leq \xi n$ of these vertices and move them from $B_{i_1}$ into $A_{j_2}$. Call the resulting sets $B^*_i$ and $A^*_i$, respectively.

We now perform an iterative procedure which will reassign vertices among the vertex classes $(A_i)_{i=1}^k$ and, separately, $(B_i)_{i=1}^k$. Initially we define the classes $A^*_i = A_i$ for each $i \neq j_2$ and $B_i^* = B_i$ for each $i \neq j_2$. Roughly speaking, $A_i^*$ (or $B_i^*$) will be the current version of $A_i$ (or $B_i$).

The choice of how we defined $B^*_i$ and $A^*_i$ is such that, initially,

$$\sum_{i=1}^k |A^*_i| = \sum_{i=1}^k |A_i| + \sum_{i=1}^k a_i \quad \text{and} \quad \sum_{i=1}^k |B^*_i| = \sum_{i=1}^k |B_i| + \sum_{i=1}^k b_i.$$  

Throughout the procedure we will ensure that (1) holds. Furthermore, throughout we will ensure that

$$|A^*_i \Delta A_i|, |B^*_i \Delta B_i| \leq 5k\xi n \leq \varepsilon |A_i|, \varepsilon |B_i|$$  

for every $1 \leq i \leq k$. We will also ensure that whenever a vertex $v$ is moved to a cluster $A^*_i$, $v$ has at least $(d-\varepsilon)|B_i|$ neighbours in $B_i$, and vice versa. We will terminate the procedure when $|A^*_i| = |A_i| + a_i$ and $|B^*_i| = |B_i| + b_i$, and then set $A'_i := A^*_i$ and $B'_i := B^*_i$ for each $1 \leq i \leq k$.

Each iteration proceeds as follows: Let $1 \leq i \leq k$ be such that $|A^*_i| < |A_i| + a_i$ and let $j \neq i$ be such that $|A^*_j| > |A_j| + a_j$. (Such $i$ and $j$ exist by (1).) Suppose that $i < j$. Note that (i) implies that $(B_{j-1}, A_j)_G$ is an $(\varepsilon, d)$-regular pair. So by (2) and Fact 11 there is a vertex $v$ in $A^*_j$ which has at least $(d-\varepsilon)|B_{j-1}|$ neighbours in $B_{j-1}$. We move $v$ from $A^*_j$ to $A^*_{j-1}$. Similarly we move one
that the following properties are satisfied. 

Suppose \( A_i \) is a graph on \( n \) vertices with \( \delta(G) \geq 2n \) which is a robust \((\nu, \tau)\)-expander. Then there exists an integer \( k \) such that \( \xi \ll 1/k \ll \varepsilon \ll d \ll \nu \ll \tau \ll \eta \ll 1 \).

Suppose \( G \) is a graph on \( n \geq n_0 \) vertices with \( \delta(G) \geq \eta n \) which is a robust \((\nu, \tau)\)-expander. Then there exists an integer \( k \) such that \( \xi \ll 1/k \ll \varepsilon \ll d \ll \nu \ll \tau \ll \eta \ll 1 \).

7. The Lemma for \( G \)

**Lemma 23** (Lemma for \( G \)). Let \( n_0 \in \mathbb{N} \) and let \( \lambda, \xi, \varepsilon, d, \nu, \tau, \eta \) be positive constants such that \( 0 < 1/n_0 \ll \lambda \ll \xi \ll \varepsilon \ll d \ll \nu \ll \tau \ll \eta \ll 1 \).

Suppose \( G \) is a graph on \( n \geq n_0 \) vertices with \( \delta(G) \geq \eta n \) which is a robust \((\nu, \tau)\)-expander. Then there exists an integer \( k \) such that \( \xi \ll 1/k \ll \varepsilon \ll d \ll \nu \ll \tau \ll \eta \ll 1 \).
Proof. Choose additional constants $\varepsilon'$ and $d'$ such that 
\[ \xi \ll \varepsilon' \ll \varepsilon \ll d \ll d' \ll \nu. \]

Apply the Regularity Lemma (Lemma 13) with parameters $\varepsilon'$, $d'$ and $k_0 := 1/\varepsilon'$ to obtain clusters $V_1, \ldots, V_{k'}$ of size $m$ (where $(1 - \varepsilon')n/k' \leq m \leq n/k'$), an exceptional set $V_0$, a pure graph $G' \subseteq G$ and the reduced graph $R$ of $G$ with parameters $\varepsilon'$, $d'$ and $k_0$. Since $\xi \ll \varepsilon'$ we may assume that 
\[ \xi \ll 1/k' \leq \varepsilon'. \]

If $k'$ is odd then we delete $V_{k'}$ from $R$ and add all of the vertices of $V_{k'}$ to $V_0$. So $|V_0| \leq \varepsilon'n + m \leq 2\varepsilon'n$. We now refer to this modified reduced graph as $R$ and redefine $k' = |R|$. By Lemma 14, $R$ originally had minimum degree at least $\eta k'/2$ and was a robust $(\nu/2, 2\tau)$-expander. So $R$ still has minimum degree at least $\eta k'/3$ and by Lemma 18, $R$ is still a robust $(\nu/3, 3\tau)$-expander.

Set $k := k'/2$. Since $1/k' \ll \nu \leq \tau < \eta < 1$, Theorem 7 implies that $R$ contains a Hamilton cycle $C = A_1B_1 \ldots A_kB_kA_1$. Since $|C| = 2k$ is even, $C$ contains a perfect matching $M = \{A_1B_1, \ldots, A_kB_k\}$. Notice that $R$ contains an edge $A_iA_j$ for some $1 \leq i_1 \neq j_1 \leq k$ and an edge $B_iB_j$ for some $1 \leq i_2 \neq j_2 \leq k$. Indeed, let $A := \{A_i\}_{i=1}^k$ and note that since $R$ is a robust $(\nu/3, 3\tau)$-expander we have \(|RN_vR(A)| \geq k + \nu k'\). This implies that $A \cap RN_vR(A) \neq \emptyset$ and hence that $R$ contains some edge $A_iA_j$. Similarly $R$ contains an edge $B_iB_j$.

Fact 11 implies that we can replace each cluster in $V(R)$ with a subcluster of size $m' := (1 - \varepsilon')m$ such that for every edge $A_iB_j \in M$ the chosen subclusters of $A_i$ and $B_j$ form a $(2\varepsilon', d'/2)$-super-regular pair in $G'$. We add all of the vertices not in these subclusters to $V_0$, and from now on we refer to the subclusters as the clusters of $R$. So $(V, W)_{G'}$ is still a $(2\varepsilon', d'/2)$-regular pair for all $VW \in E(R)$. Note that $|V_0| \leq 2\varepsilon'n + \varepsilon'n = 3\varepsilon'n$.

Our next task is to incorporate the vertices of $V_0$ into the clusters $V_1, \ldots, V_{k'}$ such that the pairs $(A_j, B_j)_{G'}$ remain ‘super-regular’ and such that the pairs $(V_i, V_j)_{G'}$ remain ‘regular’ for all $V_iV_j \in E(R)$. Let $V_0 = \{x_1, \ldots, x_t\}$ where $t \leq 3\varepsilon'n$. We will assign the vertices of $V_0$ in such a way that:

(a) At most $8\varepsilon'm'/\eta$ vertices are assigned to each cluster $V \in V(R);
(b) Whenever a vertex $x_i \in V_0$ is assigned to a cluster $A_j$, $x_i$ has at least $\eta m'/4$ neighbours in $B_j$. Similarly any vertex from $V_0$ assigned to $B_j$ has at least $\eta m'/4$ neighbours in $A_j$.

Suppose we have assigned $x_1, \ldots, x_{t-1}$ to clusters in $V(R)$ such that (a) and (b) are satisfied. Call a cluster $V \in V(R)$ full if it has already been assigned $8\varepsilon'm'/\eta$ vertices of $V_0$. Let $F$ be the set of full clusters. Since $|V_0| \leq 3\varepsilon'n$ we have $|F| \leq (3\varepsilon'n)/\eta \leq \eta k$. Thus, as $\delta(G) \geq \eta n$,
\[ |N_G(x_i) \backslash (V_0 \cup \bigcup_{V \in F} V)| \geq \eta m - 3\varepsilon'n - (\eta k)m' \geq \eta n/3. \]

Hence, by the pigeonhole principle there exists some $V \in V(R) \backslash F$ such that $|N_G(x_i) \cap V| \geq \eta m/3k' \geq \eta m'/4$. Now if $V = A_j$ for some $1 \leq j \leq k$ then we add $x_i$ to $B_j$; otherwise $V = B_j$ for some $1 \leq j \leq k$ and we add $x_i$ to $A_j$. Repeating this process for each $x_i$ we indeed assign all of the vertices of $V_0$ to the clusters of $R$ in such a way that (a) and (b) are satisfied. We now incorporate all of the assigned vertices into their respective clusters. Further, we add all those edges from $G$ with endpoints in $V_0$ to $G'$. Note that

(c) $m' \leq |V| \leq m' + 8\varepsilon'm'/\eta \leq (1 + \sqrt{\varepsilon'})m'$ for all $V \in V(R),$
(d) $(V, W)_{G'}$ is a $(\sqrt{\varepsilon'}, d'/4)$-regular pair for every edge $VW \in E(R)$ and
(e) $(V, W)_{G'}$ is a $(\sqrt{\varepsilon'}, d'/4)$-super-regular pair for every edge $VW \in E(M)$.

Next we will perform an algorithm which redistributes vertices among the clusters in $R$ in such a way that $|A_i| - |B_i| \leq \lambda n$ for each $1 \leq i \leq k$. We define $V(R)^* = \{A_i^*, B_i^*\}_{i=1}^k$, $R^*$ and $M^*$ as
follows: Initially we set $A^*_i := A_i$ and $B^*_i := B_i$ for all $1 \leq i \leq k$, $R^* := R$ and $M^* := M$. At each step we will redefine each $A^*_i$ and $B^*_i$, $R^*$ and $M^*$ and reassign vertices so that the quantity
\[ \Sigma^* = \sum_{1 \leq i \leq k, |A^*_i| - |B^*_i| > \lambda n} ||A^*_i| - |B^*_i|| \]
decreases by at least $\lambda n$. We terminate the algorithm when $\Sigma^* = 0$, i.e., when $||A^*_i| - |B^*_i|| \leq \lambda n$ for all $1 \leq i \leq k$. (We would also terminate the algorithm if all of the clusters were removed from $R^*$, but as we will prove this does not occur before $\Sigma^* = 0$.) Initially $\Sigma^* \leq 8\epsilon m^2/k/\eta \leq 4\epsilon^2 n/\eta$ by (c), and hence we need at most $4\epsilon^2/\eta \lambda$ steps to complete the process. $R^*$ will always be an induced subgraph of $R$ and at each step we set $M^*$ to be the submatching of $M$ induced by $V(R^*)$.

We will ensure that the inequality
\[ |R^*| \geq (1 - \nu/12) k' \]
holds throughout, and that $M^*$ is a perfect matching in $R^*$. Further we will ensure that
\[ |A^*_i \setminus A_i| \leq (\epsilon')^{1/3} m' \quad \text{and} \quad |B^*_i \setminus B_i| \leq (\epsilon')^{1/3} m'; \]
\[ |A_i \setminus A^*_i| \leq (\epsilon')^{1/3} m' \quad \text{and} \quad |B_i \setminus B^*_i| \leq (\epsilon')^{1/3} m' \]
for all $1 \leq i \leq k$.

Each step proceeds as follows: Call a vertex $v$ well-connected to a cluster $V \in V(R)$ if $v$ has at least $d' m'/8$ neighbours in $V$. Recall that if $VW \in E(R)$ then $(V,W)_{G'}$ is a $(\sqrt{\eta}, d'/4)$-regular pair and so $V$ contains at least $m'/2$ vertices $v$ which are well-connected to $W$. In what follows we will ensure that every vertex we redistribute to a cluster $A^*_i$ is well-connected to $B_i$ and vice versa. Since (5) holds throughout the process, $A^*_i$ will always contain at least $m'/2 - (\epsilon')^{1/3} m' \geq m'/3 \gg \lambda$ vertices that are well-connected to $B_i$. The analogous statement also holds for the $B^*_i$. Thus, at any point during the algorithm we may choose a set of $\lambda$ well-connected vertices from $A^*_i$ or $B^*_i$. (When it is clear from the context, we will not explicitly specify which cluster a vertex $v$ is well-connected to.)

Let $S$ be the set of clusters $V^* \in V(R)^*$ such that either $V^* = A^*_i$ and $|A^*_i| > |B^*_i| + \lambda n$ or $V^* = B^*_i$ and $|B^*_i| > |A^*_i| + \lambda n$. If $S$ is empty then we have $\Sigma^* = 0$ and the algorithm terminates; otherwise, choose $V^* \in S$ arbitrarily. Since $R$ is a robust $(\nu/3, 3\tau)$-expander and $\delta(R) \geq \eta k/4$ and Lemma 18 implies that $\delta^*(R^*) \geq \eta |R^*|/4$ and Lemma 17 implies that $R^*$ is a robust $(\nu/4, 4\tau)$-expander. Hence Lemma 17 implies that $R^*$ contains a shifted $M^*$-walk $P'$ of length at most $12/\nu$ which starts and finishes at $V^*$. By Lemma 15, $P'$ contains a simple shifted $M^*$-walk $P''$ which also starts and finishes at $V^*$. Now apply Lemma 16 to $P''$ to obtain a simple shifted $M^*$-walk $P$ of length at most $12/\nu$, such that the endpoints of $P$ both lie in $S$ and no other vertices of $P$ lie in $S$. We call $P$ the active walk of this step of the algorithm.

Let $P = W_1 W_2 W_3 \ldots W_{\ell-1} W_{\ell} W_\ell$ such that $W_i W_i' \in E(M^*)$ for each $2 \leq i \leq \ell - 1$. Let $W_1$ and $W_\ell$ denote the clusters such that $W_1 W_1' \in E(M^*)$. Move $\lambda n/2$ well-connected vertices from $W'_1$ into $W'_2$, $\lambda n/2$ well-connected vertices from $W'_2$ into $W'_3$, and so on until we have moved $\lambda n/2$ well-connected vertices from $W'_{\ell-1}$ to $W'_\ell$. Then move $\lambda n/2$ well-connected vertices from $W_2$ into $W_1$, $\lambda n/2$ well-connected vertices from $W_3$ into $W_2$, and so on until we have moved $\lambda n/2$ well-connected vertices from $W_\ell$ to $W_{\ell-1}$. Note that since $P$ is simple, each cluster loses at most $\lambda n$ vertices and gains at most $\lambda n$ vertices. Further for each $1 < i < \ell$ the quantity $||W_i| - |W_i'||$ remains unchanged (in fact, $|W_i|$ and $|W'_i|$ remain unchanged). For $i = 1, \ell$, $||W_i| - |W'_i||$ decreases by precisely $\lambda n$ (or $2\lambda n$ if $W'_i = W_i$).

In order to ensure that (4) holds, we remove from $R^*$ every pair $\{A_i, B_i\}$ of clusters such that $|A'_i \setminus A_i| \geq (\epsilon')^{1/3} m' - \lambda n$ or $|B'_i \setminus B_i| \geq (\epsilon')^{1/3} m' - \lambda n$. Since each cluster gains at most $\lambda n$ vertices in each step, any clusters which are not removed will still satisfy (4) at the end of the next step.
Similarly, to ensure that (5) holds, we remove from $R^*$ every pair $\{A_i, B_i\}$ of clusters such that $|A_i \backslash A_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$ or $|B_i \backslash B_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$.

**Claim 24.** For each pair $\{A_i, B_i\}$ of clusters which is removed from $R^*$ we have that $||A_i^*| - |B_i^*|| \leq \lambda n$.

To prove the claim, suppose for a contradiction that some pair $\{A_i, B_i\}$ of clusters is removed from $R^*$ and that $||A_i^*| - |B_i^*|| > \lambda n$. In order for $\{A_i, B_i\}$ to be removed we must have that $|A_i \backslash A_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$, $|B_i \backslash B_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$, $|A_i \backslash A_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$ or $|B_i \backslash B_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$. Without loss of generality assume that $|A_i \backslash A_i^*| \geq (\varepsilon')^{1/3}m' - \lambda n$. Since in each step we add at most $\lambda n$ vertices to $A_i^*$, there must have been at least $(\varepsilon')^{1/3}m'/2\lambda n$ steps in the algorithm so far, such that $A_i^*$ is contained in the active walk $P$ of each step. By the definition of $P$, either $A_i^*$ or $B_i^*$ must be an endpoint of $P$. So $||A_i^*| - |B_i^*||$ is reduced by at least $\lambda n$ during each such step, and hence by at least $(\varepsilon')^{1/3}m'/2$ during the algorithm so far. But this is a contradiction since initially $||A_i^*| - |B_i^*|| \leq \sqrt{\varepsilon}m' \leq (\varepsilon')^{1/3}m'/2$.

It remains to show that (3) holds throughout the process. Suppose for a contradiction that at some point more than $\nu k'/12$ clusters have been removed from $R^*$. Then at least

$$(\varepsilon')^{1/3}m' - \lambda n \left( \frac{\nu k'}{48} \right) > \sqrt{\varepsilon}n$$

vertices of $G$ must have been redistributed during the process so far. But at most $12\lambda n/\nu$ vertices were redistributed during each step, and at most $4\varepsilon'/\eta \lambda$ steps were performed during the process. Hence the number of redistributed vertices is at most

$$\frac{12\lambda n}{\nu} \cdot \frac{4\varepsilon'}{\eta \lambda} < \sqrt{\varepsilon}n,$$

which contradicts (6). This proves that (3) holds throughout. Now by Claim 24, we have that $||A_i^*| - |B_i^*|| \leq \lambda n$ for all $1 \leq i \leq k$ at the point where the process terminates.

We now set $n_{2i-1} := |A_i^*|$ and $n_{2i} := |B_i^*|$ for each $1 \leq i \leq k$. Notice that $n_j \geq (1 - \varepsilon')^{1/3}m' > n/3k$ for each $1 \leq j \leq 2k$. We now relabel the clusters of $R$ in the natural way so that $V(R) = V(R)^*$. Note that by (4) and (5) we have

$$|A_i \Delta A_i^*| \leq 2(\varepsilon')^{1/3}m' \quad \text{and} \quad |B_i \Delta B_i^*| \leq 2(\varepsilon')^{1/3}m'$$

for each $1 \leq i \leq k$. Hence by Proposition 12 the pair $(V, W)_{G'}$ is $(\varepsilon')^{1/10}, d'/10)$-regular for every edge $VW \in E(R)$. Further, the pair $(A_i^*, B_i^*)_{G'}$ is $(\varepsilon')^{1/10}, d'/10)$-super-regular for every $1 \leq i \leq k$. Indeed, we ensured that every vertex $v$ which was redistributed to $A_i^*$ had at least $d'm'/8$ neighbours in $B_i^*$. Since $|B_i^* \Delta B_i^*| \leq 2(\varepsilon')^{1/3}m'$, $v$ has at least $d'|B_i^*|/10$ neighbours in $B_i^*$. Similarly every vertex $v \in B_i^*$ has at least $d'|A_i^*|/10$ neighbours in $A_i^*$.

Now suppose we are given a partition $(n'_i)_{i=1}^{2k}$ of $n$ such that $n'_i \leq n_i + \varepsilon n$ for each $1 \leq i \leq 2k$. Set $a_i := n_{2i-1}' - n_{2i-1}$ and $b_i := n_{2i}' - n_{2i}$ for all $1 \leq i \leq k$. Notice that $|a_i|, |b_i| \leq 2k \varepsilon n$ for each $1 \leq i \leq k$ and $|\sum_{i=1}^k a_i| = |\sum_{i=1}^k b_i| \leq 2k \varepsilon n$. Recall that $R$ contains the edges $A_{j_1}^* A_{j_1}^*$ and $B_{j_2}^* B_{j_2}^*$. Thus, we can apply the Mobility lemma (Lemma 22) with parameters $k, 2k \varepsilon, (\varepsilon')^{1/10}, d$ and $d'/10$ to obtain a partition $A'_1, B'_1, \ldots, A'_k, B'_k$ of $V(G)$ which satisfies conditions $(\alpha_1)$–$(\alpha_5)$.
Further, we may assume that Lemma 25 which will act as an intermediate stage between 1

Proof. Let $f$ be a bipartite graph on $n \geq n_0$ vertices with bandwidth at most $\beta n$ such that $\Delta(H) \leq \Delta$. Let $n_1, n_2, \ldots, n_{2k}$ be an integer partition of $n$ such that $n_i > n/(3k)$ for all $1 \leq i \leq 2k$ and $|n_{2i-1} - n_{2i}| \leq \xi n$ for all $1 \leq i \leq k$. Suppose $C'$ is the cycle $12\ldots(2k)1$ on $[2k]$, and let $c = (2i_1)(2i_2)$ be a chord of $C$ (for some distinct $1 \leq i_1, i_2 \leq k$). Then there exists a set $S \subseteq V(H)$ and a graph homomorphism $f : H \to C \cup \{c\}$, such that

$$(\beta_1) \ |S| \leq \xi n;$$

$$(\beta_2) \ |f^{-1}(i)| \leq n_i + \xi n \text{ for all } 1 \leq i \leq 2k;$$

$$(\beta_3) \text{ Every edge which is not in } H[S] \text{ is mapped to an edge } (2i - 1)2i, \text{ for some } 1 \leq i \leq k.$$  

Choose $\beta > 0$ and integers $n_0, m_1, m_2$ and $k_1$ such that $1/n_0 \ll \beta \ll 1/m_1 \ll 1/m_2 \ll 1/k_1 \ll 1/\Delta, 1/k, \xi$.

Further, we may assume that $m_2$ divides $m_1$. We begin by defining a new cycle $C'$ with chord $c'$ which will act as an intermediate stage between $C$ and $H$, i.e., we will construct homomorphisms $f_1 : C' \cup \{c\} \to C \cup \{c\}$ and $f_2 : H \to C' \cup \{c\}$ such that $f = f_1 \circ f_2$ is our desired homomorphism. Let $C'$ be the cycle $12\ldots(2k)1$ on $[2k']$, where $k' := \sum_{i=1}^{k} (n_{2i-1} + n_{2i})k_1/n$. Note that $k_1 \leq k' \leq k_1 + k$. Define $f_1$ as follows: For each $1 \leq j \leq k'$, let $j'$ be such that $1 \leq j' \leq j < \sum_{i=1}^{j' - 1} (n_{2i-1} + n_{2i})k_1/n < j \leq \sum_{i=1}^{j'} (n_{2i-1} + n_{2i})k_1/n$. Then set $f_1(2j - 1) = 2j' - 1$ and $f_1(2j) = 2j'$ for each $1 \leq j \leq k'$. Recall that $c = (2i_1)(2i_2)$ is a chord of $C$. Suppose that $i_1', i_2'$ are such that $f_1(2i_1') = 2i_1$ and $f_1(2i_2') = 2i_2$. Notice that as $i_1 \neq i_2$, we have that $i_1' \neq i_2'$. Thus, set $c' := (2i_1')(2i_2')$ to be the chord of $C'$.

By construction $f_1(c') = c$. Given any edge $c_1 = (2j - 1)(2j)$ on $C'$, we have that $f_1(c_1) = (2j' - 1)(2j')$. Further, consider any edge $c_2 = (2j)(2j + 1) = (2j)(2j + 1) - 1) on $C'$. Then $f_1(2j) = 2j'$ and $f_1(2j + 1) = 2j + 1 - 1$. By definition of $f_1$, either $(j + 1)' = j'$ or $(j + 1)' = j' + 1$. But both $(2j')(2j' - 1)$ and $(2j')(2j' + 1)$ are edges of $C$. So in either case $f_1$ maps $c_2$ to an edge of $C$. Therefore, indeed $f_1$ is a graph homomorphism.

Roughly speaking, we will construct $f_2$ as follows: Initially we split $H$ up into small segments $A_1, B_1, \ldots, A_{m_1}, B_{m_1}$ in such a way that almost all of the edges of $H$ lie in the pairs $(A_i, B_i)_{i=1}^{m_1}$ and the remainder lie in the pairs $(B_i, A_{i+1})_{i=1}^{m_1}$. Our ideal strategy would be to map all of the vertices of $A_1$ onto vertex 1 of $C'$, the vertices of $B_1$ onto vertex 2, the vertices of $A_2$ onto vertex 3, etc. This ensures that $f_2$ is a homomorphism and that almost all of the edges of $H$ are mapped onto the edge of the form $(2i - 1)2i$ for some $i$. However the number of vertices mapped onto each vertex of $C'$ may vary widely. To solve this problem we introduce ‘drunken’ segments of the path in which the assignment of the vertices is random, and use a probabilistic argument to show that with positive probability each vertex of $C'$ receives approximately the same number of vertices of $H$. We also use the chord $c'$ to ‘turn around’ at some point during the process, in order to eliminate the possible inequality between the number of vertices of $H$ assigned to odd and even vertices of $C'$.

**Chopping $H$ up into segments.** Since $H$ has bandwidth at most $\beta n$, there exists an ordering $x_1, x_2, \ldots, x_n$ of $V(H)$ such that for every edge $x_i x_j$ of $H$, $|i - j| \leq \beta n$. Let $(A, B)$ be a bipartition of $V(H)$. We define $\{A_i, B_i\}_{i=1}^{m_1}$ as follows: for each vertex $x \in A$ there exists $1 \leq i \leq m_1$ such that $(i-1)n/m_1 - \beta n < s \leq in/m_1 - \beta n$ (unless $s > n - \beta n$). We assign $x_1$ to $A_1$ (or to $A_{m_1}$ if $s > n - \beta n$). Similarly for each vertex $x_i \in B$ there exists $1 \leq j \leq m_1$ such that $(j - 1)n/m_1 < t \leq jn/m_1$, and we assign $x_i$ to $B_j$. Let $S$ be the set of vertices $x_1$ such that $in/m_1 - 2\beta n < s \leq in/m_1 + \beta n$ for some $1 \leq i \leq m_1$. Note that the following properties hold:

(a) $n/m_1 - \beta n \leq |A_i| + |B_i| \leq n/m_1 + \beta n$ for each $1 \leq i \leq m_1;$

(b) $\sum_{i=1}^{m_1} |A_i| \leq n/m_1 + \beta n$ and $\sum_{i=1}^{m_1} |B_i| \leq n/m_1 - \beta n$.

(c) $|S| \leq n/m_1 - \beta n$.

(d) For each $i$, the set $A_i$ lies in a segment of $H$.

(e) For each $i$, the set $B_i$ lies in a segment of $H$.

(f) For each $i$, the set $S_i$ lies in a segment of $H$.

(g) For each $i$, the set $S_i$ lies in a segment of $H$.

(h) For each $i$, the set $S_i$ lies in a segment of $H$.

(i) For each $i$, the set $S_i$ lies in a segment of $H$.
(b) \( |S| \leq 3m_1 \beta n \ll \xi n; \)
(c) Every edge of \( H \) which is not in \( H[S] \) lies in one of the pairs \((A_i, B_i)\) for some \( 1 \leq i \leq m_1; \)
(d) Every edge of \( H[S] \) lies in one of the pairs \((A_i, B_i)\) or one of the pairs \((B_i, A_{i+1})\) for some \( 1 \leq i \leq m_1 \) (where \( A_{m_1+1} := A_1 \)).

Properties (c) and (d) follow from the fact that \( H \) has bandwidth at most \( \beta n \) and that \( n/m_1 \gg \beta n. \)

We now modify the small segments so that properties (a)–(d) are still satisfied and so that every small segment has size at least \( n/(4\Delta m_1). \) Suppose a small segment \( A_i \) has size smaller than \( n/(4\Delta m_1). \) Note that \( |N_H(A_i)| \leq n/(4m_1) \) and so

\[
|B_i \setminus (S \cup N_H(A_i))| \overset{(a)}{\geq} (n/m_1 - \beta n - n/(4\Delta m_1)) - 3\beta n - n/(4m_1) \geq n/(4m_1).
\]

But (c) implies that any vertex in \( B_i \setminus (S \cup N_H(A_i)) \) must be isolated in \( H \) and so may be reassigned to \( A_i \) without affecting properties (a)–(d). Hence we may reassign sufficiently many vertices so that \( |A_i|, |B_i| \geq n/(4\Delta m_1). \) For any segment \( B_i \) which has size smaller than \( n/(4\Delta m_1) \) we proceed in an identical way. From now on we denote by \( A \) the union of small segments \( \bigcup_{i=1}^{m_1} A_i \) and by \( B \) the union \( \bigcup_{i=1}^{m_1} B_i. \)

We now group the small segments together to form large segments \( \{L_j\}_{j=1}^{m_2}, \) which are defined as

\[
L_j := \bigcup_{\frac{(j-1)m_1}{m_2} < t \leq \frac{jm_1}{m_2}} (A_t \cup B_t).
\]

Note that since \( \beta \ll 1/m_1, \) (a) implies that

\[
\frac{n}{m_2} - \sqrt{\beta n} \leq |L_j| \leq \frac{n}{m_2} + \sqrt{\beta n}
\]

for each \( 1 \leq j \leq m_2. \) In order to eliminate any inequality between the number of vertices of \( H \) assigned to odd and even vertices of \( C' \) we need to partition \( \{L_j\}_{j=1}^{m_2} \) into two parts. We will assign the vertices in each part separately. For each \( 1 \leq j \leq m_2, \) set \( s_j := |L_j \cap A| - |L_j \cap B|. \) Note that \( |s_j| \leq n/m_2 + \sqrt{\beta n} - n/(4\Delta m_2) \leq n/m_2 \) for each \( 1 \leq j \leq m_2, \) and that \( \sum_{j=1}^{m_2} s_j = |A| - |B|. \)

Suppose without loss of generality that \( |A| - |B| \geq 0. \) Then since \( \beta, 1/m_2 \ll \xi \) there exists an integer \( m_3 \) so that \( \xi m_2/20 \leq m_3 \leq (1 - \xi/20) m_2 \) and

\[
\frac{|A| - |B|}{2} - \frac{\xi n}{20} \leq \sum_{i=1}^{m_3} s_j \leq \frac{|A| - |B|}{2} + \frac{\xi n}{20}.
\]

We will embed separately the large segments \( \{L_j\}_{j=1}^{m_3} \) and the segments \( \{L_j\}_{j=m_3+1}^{m_2}. \)

For each \( 1 \leq j \leq m_3, \) let the drunken segment \( D_j \) be the union of the last \( k_2 := \xi m_1/6k' m_2 \) pairs of small segments in \( L_j \) and let the sober segment \( S_j \) be the union of the rest of the small segments in \( L_j. \)

**Defining our algorithms.** We now define three different algorithms for assigning the vertices of a segment to vertices of \( C', \) given an initial vertex \( 2i_0' - 1 \in C'. \) In each case we work mod \( 2k' \) when dealing with vertices of \( C'. \) The sober algorithm is a deterministic process which proceeds as follows: Let \( S_j \) be a sober segment whose first small segments are \( A_{i_0} \) and \( B_{i_0}. \) For every \( i \) such that \( A_i \) and \( B_i \) are small segments of \( S_j, \) assign every vertex of \( A_i \) to the vertex \( 2i' - 1 \in C' \) and every vertex of \( B_i \) to the vertex \( 2i' \in C' \) where \( i' \equiv i_0' + i - i_0 \mod k'. \) So whenever \( B_i \) is assigned to \( 2i', \) \( A_{i+1} \) is assigned to \( 2i' + 1 \mod (2k'). \) We call the vertex \( 2i' \) of \( C' \) to which the vertices of the last small segment of \( S_j \) are assigned the final vertex of the algorithm and define this term in a similar way for the remaining two algorithms.

The drunken algorithm is a randomised algorithm which proceeds as follows: Given a drunken segment \( D_j \) whose first small segment is \( A_{i_0}, \) assign every vertex of \( A_{i_0} \) to the vertex \( 2i_0' - 1 \in C' \)
and every vertex of $B_{i_0}$ to the vertex $2i_0'$ of $C'$. Then for every pair $A_{i+1}, B_{i+1}$ of small segments in $D_j$, let $2i'$ be the vertex to which the vertices of $B_i$ were assigned and let

$$i'' = \begin{cases} i' & \text{with probability } \frac{1}{2}; \\ i' + 1 & \text{with probability } \frac{1}{2}. \end{cases}$$

(All random choices are made independently.) Assign every vertex of $A_{i+1}$ to $2i'' - 1$ and every vertex of $B_{i+1}$ to $2i''$.

**Claim 26.** Suppose that the vertices of $D_j$ are assigned using the drunken algorithm with initial vertex $2i_0' - 1$, and let $i_1' \in [k']$ be arbitrary. Let the random variable $I$ be the final vertex of the drunken algorithm. Then

$$\mathbb{P}[I = 2i_1' \mid i_0'] \leq \frac{1 + \xi/20}{k'}. $$

To prove the claim, note that $I \sim 2(i_0' + \text{Bin}(k_2, 1/2))$, that $k_2 \gg (k')^3/6$ and that $1/k' \ll \xi/20$. So Lemma 19 with $\varepsilon := \xi/20$ implies that $\mathbb{P}[I - 2i_0' = 2i_1' - 2i_0'] \leq (1 + \xi/20)/k'$, and Claim 26 follows immediately.

The 2$i_1'$-seeking algorithm is a deterministic algorithm which proceeds as follows: Given a drunken segment $D_j$ whose first small segment is $A_{i_0}$, assign every vertex of $A_{i_0}$ to the vertex $2i_0' - 1$ of $C'$ and every vertex of $B_{i_0}$ to the vertex $2i_0'$ of $C'$. Then for every pair $A_{i+1}, B_{i+1}$ of small segments in $D_j$, let $2i'$ be the vertex of $C'$ to which the vertices of $B_i$ were assigned and let

$$i'' = \begin{cases} i' & \text{if } i' = i_1'; \\ i' + 1 & \text{otherwise}. \end{cases}$$

Assign every vertex of $A_{i+1}$ to $2i'' - 1$ and every vertex of $B_{i+1}$ to $2i''$. Note that the final vertex of this algorithm is always $2i_1'$, since $k' \leq \xi m_1/(6k' m_2)$.

**Applying the algorithms.** We use these algorithms to assign small segments to vertices of $C'$ as follows: Choose $i_0' \in [k']$ randomly and let $2i_0' - 1$ be the initial vertex for $S_1$. For each $1 \leq j \leq m_3 - 1$, use the sober algorithm to assign the segments of $S_j$ and then use the drunken algorithm to assign the vertices of $D_j$, where in each case the initial vertex of each segment is the successor of the final vertex of the previous segment. (So for example, if the final vertex of the drunken algorithm, when applied to $D_j$, is $2i'$, then the initial vertex of the sober algorithm, when applied to $S_{j+1}$, is $2i' + 1$.) Then assign the vertices of $S_{m_3}$ using the sober algorithm and assign the vertices of $D_{m_3}$ using the 2$i_1'$-seeking algorithm. (Recall that 2$i_1'$ was a vertex of the chord $c'$.) We explain how we assign the small segments from $\bigcup_{j=m_3+1}^{m_2} L_j$ later.

**Claim 27.** For each $1 \leq i \leq k'$, let $X_i$ be the number of vertices of $\bigcup_{j=1}^{m_3} S_j$ assigned to the vertex $2i - 1$ of $C'$. Then

$$\mathbb{P} \left[ X_i > \frac{1}{2k'} \left( \frac{m_3 n}{m_2} + \frac{|A| - |B|}{2} \right) + \frac{\xi n}{6k'} \right] \leq \frac{1}{3k'}.$$ 

For each $1 \leq j \leq m_3$, let $Y_j = |S_j \cap A|$ and let $X_{i,j}$ be the number of vertices of $S_j$ which are assigned to $2i - 1$. To prove the claim, we first use Claim 26 to bound $\mathbb{E}[X_{i,j} \mid X_{i,j-1}, \ldots, X_{i,1}]$. Let $r_j$ be the initial vertex of $S_j$ for each $j$. Let $B$ be the event $X_{i,j-1} = x_{i,j-1}, \ldots, X_{i,1} = r_{i,1}$ for some $x_{i,1}, \ldots, x_{i,j-1}$. Now for $1 \leq i' \leq k'$ and any integer $x$ we have $\mathbb{P}[X_{i,j} = x \mid B \cap (r_{j-1} = 2i' - 1)] = \mathbb{P}[X_{i,j} = x \mid r_{j-1} = 2i' - 1]$. Hence Lemma 21 implies that $\mathbb{E}[X_{i,j} \mid X_{i,j-1}, \ldots, X_{i,1}] \leq \frac{1}{2k'} |A| - |B| \frac{m_3 n}{m_2} + \frac{\xi n}{6k'}$.
max_{i' = 1}^{k'} \mathbb{E}[X_{i,j} \mid r_{j-1} = 2i' - 1]. But Claim 26 implies that
\[
\mathbb{E}[X_{i,j} \mid r_{j-1} = 2i' - 1] = \sum_{i''=1}^{k'} \mathbb{E}[X_{i,j} \mid r_{j} = 2i'' - 1] \mathbb{P}[r_{j} = 2i'' - 1 \mid r_{j-1} = 2i' - 1] \leq \frac{1 + \xi/20}{k'} \sum_{i''=1}^{k'} \mathbb{E}[X_{i,j} \mid r_{j} = 2i'' - 1] = \frac{(1 + \xi/20)Y_{j}}{k'}.
\]

Hence \( \mathbb{E}[X_{i,j} \mid X_{i,j-1}, \ldots, X_{i,1}] \leq (1 + \xi/20)Y_{j}/k' \). Set \( X_{i,j}' := X_{i,j}m_{2}/n \). Since \( |S_{j}| \leq \left( \frac{n}{m_{2}} + \sqrt{\beta n} \right) - \left( \frac{\xi m_{1}}{6k'm_{2}} \right) - \left( \frac{n}{4\Delta m_{1}} \right) \leq \frac{n}{m_{2}} \), we have that \( X_{i,j}' \in [0, 1] \), for each \( 1 \leq j \leq m_{3} \).

Let
\[
\mu = \sum_{j=1}^{m_{3}} \frac{(1 + \xi/20)Y_{j}m_{2}}{k'n},
\]
and note that
\[
\sum_{j=1}^{m_{3}} Y_{j} \leq \sum_{j=1}^{m_{3}} |A \cap L_{j}| = \frac{1}{2} \left( \sum_{j=1}^{m_{3}} |L_{j}| + s_{j} \right) \leq \left( \frac{m_{3}n}{2m_{2}} + \frac{m_{3}\sqrt{\beta n}}{2} \right) + \left( \frac{|A| - |B|}{4} + \frac{\xi n}{40} \right) \leq \frac{m_{3}n}{2m_{2}} + \frac{|A| - |B|}{4} + \frac{\xi n}{20}.
\]

Note also that \( Y_{j} \geq (n/(4\Delta m_{1})) \times (m_{1}/2m_{2}) = n/(8\Delta m_{2}) \) for each \( 1 \leq j \leq m_{1} \). Thus we have \( \mu \geq m_{3}/(8\Delta k') \gg (\log k')/\xi^{2} \). We now apply Lemma 20 with \( \delta := \xi/20 \) to obtain
\[
\mathbb{P} \left[ \sum_{j=1}^{m_{3}} X_{i,j}' > (1 + \xi/20)\mu \right] = e^{-\frac{\xi^{2}n}{16m_{2}} \leq \frac{1}{3k'}.}
\]

It follows that with probability at least \( 1 - 1/3k' \),
\[
X_{i} \leq \frac{(1 + \xi/20)^{2}}{k'} \sum_{i=1}^{m_{3}} Y_{j} \leq \frac{(1 + \xi/20)^{2}}{k'} \left( \frac{m_{3}n}{2m_{2}} + \frac{|A| - |B|}{4} + \frac{\xi n}{20} \right) \leq \frac{1}{2k'} \left( \frac{m_{3}n}{m_{2}} + \frac{|A| - |B|}{2} \right) + \frac{\xi n}{6k'},
\]
which proves Claim 27.

By a similar argument we have that if \( X_{i}' \) is the number of vertices of \( \bigcup_{j=1}^{m_{3}} S_{j} \) assigned to \( 2i \), then
\[
\mathbb{P} \left[ X_{i}' > \frac{1}{2k'} \left( \frac{m_{3}n}{m_{2}} + \frac{|B| - |A|}{2} \right) + \frac{\xi n}{6k'} \right] \leq \frac{1}{3k'},
\]
for every \( 1 \leq i \leq k' \). Taken together with Claim 27 this implies that with probability at least \( 1/3 \),
\[
X_{i} \leq \frac{1}{2k'} \left( \frac{m_{3}n}{m_{2}} + \frac{|B| - |A|}{2} \right) + \frac{\xi n}{6k'}
\]
and \( X_{i}' \leq \frac{1}{2k'} \left( \frac{m_{3}n}{m_{2}} + \frac{|B| - |A|}{2} \right) + \frac{\xi n}{6k'} \)
for every \( 1 \leq i \leq k' \), and hence there exists an assignment such that (11) holds.
For each $m_3 < j \leq m_2$, let $D_j$ be the union of the first $k_2$ small segments of $L_j$ and $S_j$ the union of the remaining small segments. We now assign the vertices of $\bigcup_{j=m_3+1}^{m_2} L_j$ using an algorithm similar to that for $\bigcup_{j=1}^{m_3} L_j$, but in reverse order. That is, we first choose $1 \leq i_0' \leq k'$ randomly and assign the vertices of $S_{m_3}$ using the sober algorithm, but with the roles of $A_i$ and $B_i$ exchanged for each $i$. Thus we assign the vertices of $B_{m_3}$ to $2i_0' - 1$, the vertices of $A_{m_3}$ to $2i_0'$, etc. Similarly we use the drunken algorithm to assign the vertices of $D_{m_2}$ (again with the roles of $A_i$ and $B_i$ exchanged for each $i$), and so on until we have assigned all the vertices up to $S_{m_3+1}$. (As before, the initial vertex of any application of an algorithm is the successor of the final vertex of the previous application of an algorithm.) Finally we use the $2i_2'$-seeking algorithm to assign the vertices of the last drunken segment $D_{m_3+1}$. (Recall that the final vertex of the $2i_2'$-seeking algorithm is always $2i_2'$.) Let $X'_i$ be the number of vertices of $\bigcup_{j=m_3+1}^{m_2} S_j$ assigned to to $2i - 1$ and $X_i^j$ the number assigned to $2i$. By using a proof analogous to that of Claim 27, we can ensure that

$$X_i' \leq \frac{1}{2k'} \left( \frac{(m_2 - m_3)n}{m_2} + \frac{|B| - |A|}{2} \right) + \frac{\xi n}{6k'} \tag{12}$$

and

$$X_i^j \leq \frac{1}{2k'} \left( \frac{(m_2 - m_3)n}{m_2} + \frac{|A| - |B|}{2} \right) + \frac{\xi n}{6k'}$$

for each $1 \leq i \leq k'$.

Note that

$$| \bigcup_{j=1}^{m_2} D_j | \leq m_2 \times \frac{\xi m_1}{6k'm_2} \times (n/m_1 + \beta n) \leq \xi n/3k',$$

and hence in total we assign at most

$$X_i + X_i' + | \bigcup_{j=1}^{m_2} D_j | \leq \frac{1}{2k'} \left( \frac{m_3 n}{m_2} + \frac{|A| - |B|}{2} \right) + \frac{1}{2k'} \left( \frac{(m_2 - m_3)n}{m_2} + \frac{|B| - |A|}{2} \right) + \frac{2\xi n}{3k'} \leq \frac{1}{2k'} (1 + 4\xi/3)n$$

vertices to $2i - 1$ and at most $(1 + 4\xi/3)n/2k'$ vertices to $2i$, for each $1 \leq i \leq k'$.

**Completing the proof.** This completes our definition of $f_2$. We now check that $f_2$ is a homomorphism. By properties (c) and (d) it suffices to show for each $i$ that whenever $A_i$ and $B_i$ (or $A_i$ and $A_{i+1}$) are assigned to vertices $1 \leq i', j' \leq 2k'$ of $C'$, then $i'j'$ is an edge of $C' \cup \{c'\}$. Observe first that the sober, drunken and seeking algorithms all assign vertices in such a way that $i'j'$ is an edge of $C'$. Further, recall that the initial vertex of any application of an algorithm is the successor of the final vertex of the previous application of an algorithm. So if, for example, the vertices of $B_i$ are assigned to the final vertex $2i^*$ where $B_i$ is the final segment assigned in an application one of the algorithms, then the vertices of $A_{i+1}$ will be assigned to the initial vertex $2i^* + 1$ in the next application of an algorithm.

The only pair this argument does not deal with is the pair $(B_j, A_{j+1})$, where $B_j$ is the last small segment of $D_{m_3}$ (and hence $A_{j+1}$ is the first small segment of $D_{m_3+1}$, and therefore the last to be assigned). Now by the definition of the seeking algorithm, the vertices of $B_j$ are assigned to the vertex $2i'_0$ of $C'$ and the vertices of $A_{j+1}$ are assigned to the vertex $2i'_2$ of $C'$. Recalling that $c' = (2i'_1)(2i'_2)$ we have that $f_2$ is indeed a homomorphism.

Now consider $f = f_1 \circ f_2$. Since $f_1$ and $f_2$ are both homomorphisms we have that $f$ is a homomorphism. Property (b) implies that condition $(\beta_1)$ holds. By (c), every edge $xy$ not in $H[S]$ lies in a pair $(A_i, B_i)$ for some $i$. Thus, by definition of our three algorithms, $xy$ is mapped to an
edge \((2j - 1)2j\) by \(f_2\) for some \(j\). By definition of \(f_1\), \((2j - 1)(2j)\) is mapped to \((2j' - 1)(2j')\) by \(f_1\) for some \(j'\). Therefore, \(f\) satisfies \((\beta_3)\).

To see that condition \((\beta_2)\) also holds, recall that \(f_1\) assigns to each vertex \(2i - 1\) of \(C\) (and also to \(2i\)) exactly \(\lceil (n_{2i-1} + n_{2i})k_1/n \rceil\) vertices of \(C'\). Hence \(f\) assigns at most
\[
\frac{1}{2k'}(1 + 4\xi/3)(n_{2i-1} + n_{2i})k_1/n \leq (1 + 4\xi/3)(n_{2i-1} + \xi n/10) + \frac{1}{2k'}(1 + 4\xi/3)n
\]
vertices of \(H\) to the vertex \(2i - 1\) of \(C\), for each \(1 \leq i \leq k\). Similarly \(f\) assigns at most \(n_{2i} + \xi n\) vertices of \(H\) to the vertex \(2i\).

\[\square\]

9. Completing the Proof

In this section we use Lemmas 23 and 25 to prove Theorem 8. We use the following definition and lemma from [2]; these allow us to prove that \(H\) embeds into \(G\) by checking some relatively simple conditions.

**Definition 28.** Let \(H\) be a graph on \(n\) vertices, let \(R\) be a graph on \([k]\), and let \(R' \subseteq R\). We say that a partition \(V(H) = (W_i)_{i \in [k]}\) of \(H\) is \(\varepsilon\)-compatible with an integer partition \((n_i)_{i \in [k]}\) of \(n\) if the following holds. For \(i \in [k]\) let \(S_i\) be the set of vertices in \(W_i\) with neighbours in some \(W_j\) with \(ij \notin E(R')\) and \(i \neq j\). Set \(S := \bigcup_{i \in [k]} S_i\) and \(T_i := N_H(S) \cap (W_i \setminus S)\). Then for all \(i, j \in [k]\) we have that
\[
\begin{align*}
(\gamma_1) \quad & |W_i| = n_i; \\
(\gamma_2) \quad & xy \in E(H) \text{ for } x \in W_i \text{ and } y \in W_j \text{ implies that } ij \in E(R); \\
(\gamma_3) \quad & |S_i| \leq \varepsilon n_i \text{ and } |T_i| \leq \varepsilon \cdot \min\{n_j | i \text{ and } j \text{ are in the same component of } R'\}.
\end{align*}
\]

The partition \(V(H) = (W_i)_{i \in [k]}\) is \(\varepsilon\)-compatible with a partition \(V(G) = (V_i)_{i \in [k]}\) of a graph \(G\) and \(R' \subseteq R\) if \(V(H) = (W_i)_{i \in [k]}\) is \(\varepsilon\)-compatible with \((|V_i|)_{i \in [k]}\) and \(R' \subseteq R\).

**Lemma 29** ([2], Lemma 3.12). For all \(d, \Delta, r > 0\) there is a constant \(\varepsilon = \varepsilon(d, \Delta, r)\) such that the following holds. Let \(G\) be a graph on \(n\) vertices and suppose that \((V_i)_{i \in [k]}\) is a partition of \(V(G)\). Suppose \(R\) is an \((\varepsilon, d)\)-reduced graph of \(G\) on \(V_1, \ldots, V_k\) and that \(R'\) is a subgraph of \(R\) whose connected components have size at most \(r\). Assume that \((V_i, V_j)_G\) is an \((\varepsilon, d)\)-super-regular pair for every edge \(V_iV_j \in E(R')\). Further, let \(H\) be a graph on \(n\) vertices with maximum degree \(\Delta(H) \leq \Delta\) that has a vertex partition \(V(H) = (W_i)_{i \in k}\) which is \(\varepsilon\)-compatible with \(V(G) = (V_i)_{i \in [k]}\) and \(R' \subseteq R\). Then \(H \subseteq G\).

Lemma 29 is a consequence of the Blow-up lemma of Komlós, Sárközy and Szemerédi [16]. We now prove Theorem 8.

**Proof of Theorem 8.** Firstly, note that it suffices to prove Theorem 8 under the addition assumption that \(\eta \ll 1\). We choose \(\beta, n_0\) as well as additional constants \(d, \varepsilon, \xi, \lambda\) as follows: Choose \(d \ll \nu\) as required by Lemma 23 also ensuring that \(d \ll 1/\Delta\). Then take \(\varepsilon \ll \varepsilon(d, \Delta, 2)\) as in Lemma 29, ensuring also that \(\varepsilon \ll d\). Finally choose
\[
1/n_0 \ll \beta \ll \lambda \ll \xi \ll \varepsilon,
\]
as required by Lemmas 23 and 25.

Apply Lemma 23 to \(G\) to obtain an integer \(k\) such that \(\xi \ll 1/k \ll \varepsilon\), a partition \((n_i)_{i=1}^{2k}\) of \(n\) and integers \(1 \leq i_1 \neq j_1, i_2 \neq j_2 \leq k\). Suppose \(C\) is the cycle \(12\ldots(2k)1\) with the chord \(c = (2i_2)(2j_2)\). Next we apply Lemma 25 with the partition \((n_i)_{i=1}^{2k}\) of \(n\) as input to obtain a set \(S \subseteq V(H)\) and a homomorphism \(f : H \to C \cup \{c\}\), such that
\[
(i) \quad |S| \leq \xi n;
\]
(ii) $|f^{-1}(i)| \leq n_i + \xi n$ for all $1 \leq i \leq 2k$;

(iii) Every edge which is not in $H[S]$ is mapped to the edge $(2i - 1)2i$, for some $1 \leq i \leq k$.

Let $W_i := f^{-1}(i)$ and $n_i := |f^{-1}(i)|$ for each $i$, and note that $(n_i)_{i=1}^k$ is a partition of $n$. Condition (ii) together with Lemma 23 imply that there is a partition $A_1', B_1', A_2', B_2', \ldots, A_k', B_k'$ of $V(G)$ and a spanning subgraph $G'$ of $G$ which satisfy conditions $(\alpha_1)$–$(\alpha_5)$.

Relabel these clusters $V_1, \ldots, V_k$ such that $V_{2i-1} := A_i'$ and $V_{2i} := B_i'$ for all $1 \leq i \leq k$. So $|V_i| = n_i'$ for all $1 \leq i \leq 2k$. Let $R$ be the $(\varepsilon, d)$-reduced graph of $G'$ on $V_1, \ldots, V_k$ with the maximal number of edges. Hence $(\alpha_2)$–$(\alpha_5)$ imply that $R$ contains the Hamilton cycle $C' = V_1V_2 \ldots V_{2k}V_1$ and the chord $c' := V_{2i}V_{2j}$ (we view $C' \cup \{c'\}$ as a copy of $C \cup \{c\}$ in $R$). Let $R'$ be the spanning subgraph of $R$ containing precisely the edges $V_{2i-1}V_{2i}$ for $1 \leq i \leq k$. Note that $(\alpha_2)$ implies that $(V_{2i-1}V_{2i})_C$ is an $(\varepsilon, d)$-super-regular pair for all $1 \leq i \leq k$.

We now check that the partition $V(H) = (W_i)_{i=1}^{2k}$ is $\varepsilon$-compatible with the partition $(V_i)_{i=1}^{2k}$ and $R' \subseteq R$. We defined $(V_i)_{i=1}^{2k}$ so that $|V_i| = |W_i|$ for each $1 \leq i \leq 2k$ and hence condition $(\gamma_1)$ of Definition 28 holds. Condition $(\gamma_2)$ holds since $f : H \to C \cup \{c\}$ is a homomorphism and $C \cup \{c\}$ is a subgraph of $R$. Note that for all $1 \leq i \leq 2k$, Lemma 23 implies that

$$
\varepsilon n_i' \geq \varepsilon (n_i - 2k\xi n) \geq \varepsilon (n/3k - 2k\xi n) \geq \varepsilon n/4k \gg \xi n \geq |S|.
$$

Furthermore, $|N_H(S) \cap (W_i \setminus S)| \leq \Delta |S| \leq \Delta \xi n \ll \varepsilon n/4k \leq \varepsilon n'_i$ for all $1 \leq i, j \leq 2k$. Thus, condition $(\gamma_3)$ holds. Hence, Lemma 29 implies that $G'$ (and therefore $G$) contains $H$, as desired.

\[\square\]

ACKNOWLEDGEMENTS

We would like to thank József Balogh for helpful discussions. We would also like to thank Daniela Kühn and Deryk Osthus for their comments on the manuscript.

REFERENCES


Fiachra Knox Andrew Treglown
School of Mathematics Faculty of Mathematics and Physics
University of Birmingham Malostranské Náměstí 25
Birmingham 188 00 Prague
B15 2TT Czech Republic
UK
E-mail addresses: knoxf@maths.bham.ac.uk, treglown@kam.mff.cuni.cz