Received: 16 February 2019 | Revised: 7 April 2020 | Accepted: 18 April 2020

DOI: 10.1002/jgt.22576

ARTICLE

WILEY

Finding any given 2-factor in sparse pseudorandom graphs efficiently

Jie Han[1] | **Yoshiharu Kohayakawa**[2] | **Patrick Morris**[3,4] | **Yury Person**[5]

[1]Department of Mathematics, University of Rhode Island, Kingston, Rhode Island

[2]Instituto de Matemáticae Estatística, Universidade de São Paulo, São Paulo, Brazil

[3]Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

[4]Berlin Mathematical School, Berlin, Germany

[5]Institut für Mathematik, Technische Universität, Ilmenau, Germany

Correspondence
Jie Han, Department of Mathematics, University of Rhode Island, 5 Lippitt Road, Kingston, RI 02881.
Email: jie_han@uri.edu

Funding information
Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Numbers: 2013/03447-6, 2014/18641-5; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Numbers: 310974/2013-5, 311412/2018-1, 423833/2018-9; Leverhulme Trust, Grant/Award Number: SAS-2017-052\9

Abstract

Given an n-vertex pseudorandom graph G and an n-vertex graph H with maximum degree at most two, we wish to find a copy of H in G, that is, an embedding $\varphi: V(H) \to V(G)$ so that $\varphi(u)\varphi(v) \in E(G)$ for all $uv \in E(H)$. Particular instances of this problem include finding a triangle-factor and finding a Hamilton cycle in G. Here, we provide a deterministic polynomial time algorithm that finds a given H in any suitably pseudorandom graph G. The pseudorandom graphs we consider are (p, λ)-bijumbled graphs of minimum degree which is a constant proportion of the average degree, that is, $\Omega(pn)$. A (p, λ)-bijumbled graph is characterised through the discrepancy property: $|e(A, B) - p|A||B|| < \lambda\sqrt{|A||B|}$ for any two sets of vertices A and B. Our condition $\lambda = O(p^2 n/\log n)$ on bijumbledness is within a log factor from being tight and provides a positive answer to a recent question of Nenadov. We combine novel variants of the absorption-reservoir method, a powerful tool from extremal graph theory and random graphs. Our approach builds on our previous work, incorporating the work of Nenadov, together with additional ideas and simplifications.

KEYWORDS

2-factors, absorbing meth, expander graphs

1 | INTRODUCTION

A pseudorandom graph of edge density p is a deterministic graph which shares typical properties of the corresponding random graph $G(n, p)$. These objects have attracted considerable attention in computer science and mathematics. Thomason [48,49] was the first to introduce a quantitative notion of a pseudorandom graph by defining so-called (p, λ)-jumbled graphs G which satisfy $\left| e(U) - p\binom{|U|}{2} \right| \le \lambda |U|$ for every vertex subset $U \subseteq V(G)$. Ever since, there has been a great deal of investigation into the properties of pseudorandom graphs and this is still a very active area of modern research.

The most widely studied class of jumbled graphs are the so-called (n, d, λ)-graphs, which were introduced by Alon in the 1980s. These graphs have n vertices, are d-regular and their second largest eigenvalue in absolute value is at most λ. An (n, d, λ)-graph satisfies the *expander mixing lemma* [11] allowing good control of the edges between any two sets of vertices A and B:

$$\left| e(A, B) - \frac{d}{n}|A||B| \right| < \lambda\sqrt{|A||B|}, \tag{1}$$

where $e(A, B) = e_G(A, B)$ denotes the number of pairs[1] $(a, b) \in A \times B$ so that ab is an edge of G. An illuminating survey of Krivelevich and Sudakov [37] provides a wealth of applications.

There are three interesting regimes in the study of pseudorandom graphs and the class of (n, d, λ)-graphs is versatile enough to capture the essence of all of these regimes. In the first, one assumes $\lambda = \varepsilon n$, where n is the number of vertices in a graph G and $\varepsilon > 0$ is an arbitrary fixed parameter. In this regime one can control edges between sets of linear sizes. This is tightly connected to the theory of quasirandom graphs [15] and the applications of the regularity lemma of Szemerédi [34]. The second regime is when d is constant and $\lambda < d$. This class then contains (nonbipartite) expanders [27] and Ramanujan graphs [41], which are prominent objects of study throughout mathematics and computer science. The third regime (sparse graphs) concerns λ being $o(n)$, often some power of n, where one has better control on the distribution of edges between truly smaller sets. This case has been investigated more recently and made amenable to some tools from extremal combinatorics [1,2,19,24,25,26,31,36,37,38,44].

The focus of this paper will be on conditions under which certain spanning or almost spanning structures are forced in sparse pseudorandom graphs. Our main motivation comes from probabilistic and extremal combinatorics, in particular the problem of universality. A graph G is called $\mathcal{F}$-universal for some family $\mathcal{F}$ if any member $F \in \mathcal{F}$ can be embedded into G. This problem attracted a lot of attention [6,7,8,9], especially for the case where $\mathcal{F}$ is a class of bounded degree spanning subgraphs. In this case we say an n-vertex graph G is Δ-universal if it contains all graphs on at most n vertices of maximum degree Δ. A large part of the focus of the study has been on the universality properties of $G(n, p)$ [9,18,20,22,23,29]. It is also natural to investigate the universality properties of (n, d, λ)-graphs as was suggested by Krivelevich, Sudakov and Szabó in [38]. In this setting of sparse pseudorandom graphs, a general result on universality has been proved only recently in [1]. Let us comment that the case of dense graphs is well understood since the blow-up lemma of Komlós et al [32] establishes that pseudorandom

[1]Note that edges in $A \cap B$ are counted twice.

graphs of linear minimum degree contain any given bounded degree spanning structure. A little later, the second and fourth author established jointly with Allen et al in [1], a variant of a blow-up lemma for regular subgraphs of pseudorandom graphs. This provides a machinery, complementing the results of Conlon et al [19] and allowing to transfer many results about dense graphs to sparse graphs in a unified way. However, these results are very general and thus do not establish tight conditions for special cases of spanning structures.

Much more is known for questions about finding one particular spanning structure in a pseudorandom graph and the most prominent spanning structures which were considered in the last fifteen years include perfect matchings, studied by Alon et al in [37], Hamilton cycles studied by Krivelevich and Sudakov [36], clique-factors [24,25,38,44] and powers of Hamilton cycles [2].

The problem of when a triangle-factor[2] appears in a given (n, d, λ)-graph has been a prominent question and is an instructive insight into the behaviour of pseudorandom graphs. It is easy to infer from the expander mixing lemma that if $\lambda \leq 0.1d^2/n$, then any (n, d, λ)-graph contains a triangle (in fact, every vertex lies in a triangle). An ingenious construction of Alon [4] provides an example of a triangle-free (n, d, λ)-graph with $\lambda = \Theta(n^{1/3})$ and $d = \Theta(n^{2/3})$, which is essentially as dense as possible, considering the previous comments. This example can be bootstrapped, as is done in [38], to the whole possible range of $d = d(n)$, giving K_3-free (n, d, λ)-graphs with $\lambda = \Theta(d^2/n)$. Further examples of (near) optimal dense pseudorandom triangle-free graphs have since been given [17, 35]. On the other hand, Krivelevich et al [38] proved that (n, d, λ)-graphs with $\lambda = o(d^3/(n^2 \log n))$ contain a triangle-factor if $3|n$ and they made the following intriguing conjecture, which is one of the central problems in the theory of spanning structures in (n, d, λ)-graphs.

Conjecture 1.1 (Conjecture 7.1 in Krivelevich et al [38]). *There exists an absolute constant $c > 0$ such that if $\lambda \leq cd^2/n$, then every (n, d, λ)-graph G on $n \in 3\mathbb{N}$ vertices has a triangle-factor.*

This conjecture is supported by their result [38] that $\lambda \leq 0.1d^2/n$ implies the existence of a fractional triangle-factor. Furthermore, a recent result of three of the authors [25, 26] states that, under the condition $\lambda \leq (1/600)d^2/n$, any (n, d, λ)-graph G with n sufficiently large contains a family of vertex-disjoint triangles covering all but at most $n^{647/648}$ vertices of G, thus a "near-perfect" triangle-factor. A very recent, remarkable result of Nenadov [44] infers that $\lambda \leq cd^2/(n \log n)$ for some constant $c > 0$ is sufficient to yield a triangle-factor. Considering the triangle-free constructions mentioned above, we see that Nenadov's result is within a log factor of the optimal conjectured bound. Nenadov also raised the question in [44] of whether a similar condition would imply the existence of any given 2-factor[3] in a pseudorandom graph. The purpose of this study is to give a positive answer to Nenadov's question, casting the question in terms of 2-universality and showing that we can efficiently find a given subgraph of maximum degree 2 in polynomial time.

To state our result we will switch[4] to working with (p, λ)-bijumbled graphs (introduced in [31]), which give a convenient, slight variant of Thomason's jumbledness. Bijumbled graphs G satisfy the property

$$|e(A, B) - p|A||B|| < \lambda\sqrt{|A||B|} \tag{2}$$

[2]That is, vertex-disjoint copies of K_3 covering all the vertices.

[3]A 2-factor is a 2-regular spanning subgraph.

[4]Nenadov also worked in this broader class of pseudorandom graphs.

for all $A, B \subseteq V(G)$. In particular it is easy to see by the expander mixing lemma (1) that an (n, d, λ)-graph is $(d/n, \lambda)$-(bi)jumbled. Moreover, the two concepts are closely linked as a (p, λ)-(bi)jumbled graph is almost pn-regular, in that almost all vertices have degree close to pn.

Before the current paper, the best result towards 2-universality in pseudorandom graphs is due to Allen, Böttcher, Hàn and two of the authors [2]. There, they proved that there exists an $\varepsilon > 0$ such that $(p, \varepsilon p^{5/2} n)$-bijumbled graphs of minimum degree $\Omega(pn)$ contain a square[5] of a Hamilton cycle and hence are 2-universal. The proof is algorithmic, leading to an efficient procedure. Here we weaken the requirement on λ to match that of Nenadov and obtain the following.

Theorem 1.2. *For all $\delta > 0$, there exist constants $\varepsilon > 0$ and n_0 such that, for any $p \in (0, 1]$, the following holds. For any $n \geq n_0$ and any given potential 2-factor F (ie, family of disjoint cycles whose lengths sum up to n), there is a polynomial time algorithm which finds a copy of F in any (p, λ)-bijumbled graph G on n vertices with $\lambda \leq \varepsilon p^2 n / \log n$ and minimum degree $\delta(G) \geq \delta pn$.*

In particular, Theorem 1.2 implies that such a (p, λ)-bijumbled graph G is 2-universal. Indeed, given a graph F' on at most n vertices with $\Delta(F') \leq 2$, we find a supergraph F of F' on n vertices, so that all but at most one of the components of F are cycles. It is possible that F may have either one isolated vertex or a single edge but since we can easily embed a single vertex/edge into a bijumbled graph G altering its minimum degree only a little, it suffices to concentrate on the case that F is a 2-factor.

We remark that the minimum degree condition in Theorem 1.2 is weak and natural. Indeed some minimum degree condition is necessary as otherwise one could have isolated vertices and the bijumbled definition (2) guarantees that almost all vertices satisfy the minimum degree condition in any case. Finally, we mention that we do not try to optimise the running time of our algorithm and are satisfied with being able to provide a deterministic algorithm which is efficient in that it runs in polynomial time. Indeed, the problem of establishing the existence of certain 2-factors (eg, for triangle-factors [30] and Hamilton cycles [28]) in graphs is known to be NP-complete and many proofs of existence of spanning structures in certain graph classes adopt probabilistic methods.

1.1 | Proof method

Our proof uses the absorption-reservoir method, which has been a powerful tool in proving the existence of certain spanning structures and is often superior to the aforementioned blow-up lemmas. The basic idea of the method is to carefully define an "absorbing structure" which can contribute to the desired spanning structure in many ways. One then finds such an absorbing structure in the host (hyper-)graph and, after putting this to one side, finds almost all of the desired spanning structure in the remainder of the host graph. The absorbing structure then provides the flexibility to "clean up" and complete the spanning structure. The beginnings of this method date back to the early 1990s, but the breakthrough in the wide applicability of these

[5]A square of a graph H is obtained by connecting its vertices at distance at most two through edges. The existence of a square of a Hamilton cycle implies 2-universality as one can greedily find vertex-disjoint cycles of arbitrary lengths (see, eg [21]).

methods, however, was first established by Rödl et al [46, 47] in their study of Hamiltonicity in hypergraphs. There, the method was used to study dense hypergraphs but the methods have since been adapted to other settings (see eg [2, 3, 39]).

In his work on spanning trees in random graphs [42, 43], Montgomery ingeniously wove sparse "robust" bipartite graphs (which we call *sparse templates*) into the absorption-reservoir method. The first use of sparse templates for the absorption in the context of pseudorandom graphs was recently given by the current authors in [24]. Here, we again use this idea and introduce for the first time, an efficient version of this new type of absorption, which may be of independent interest. To explicitly generate a sparse template we use bounded degree bipartite graphs with strong expansion properties. Such graphs are known as *concentrators* [10, 27].

Our general proof approach here builds on the ideas from our paper [24], derandomising additional arguments at various places and adapting the method to handle different types of 2-factors. To deal with triangles we also require a different absorbing-type argument, namely, the argument due to Nenadov [44], for which we replace certain nonalgorithmic arguments.

2 | PROOF OF THEOREM 1.2

The following three theorems will establish our main result. Note that the nonalgorithmic version of Theorem 2.1 was proved in [44].

Theorem 2.1 (Theorem 1.2, Nenadov [44]). *For every $\delta > 0$ there exists a constant $\varepsilon > 0$ such that, for any $p \in (0, 1]$, a (p, λ)-bijumbled graph G on $n \in 3\mathbb{N}$ vertices with $\lambda \le \varepsilon p^2 n/\log n$ and minimum degree $\delta(G) \ge \delta pn$ contains a triangle-factor, which can be found with a deterministic polynomial time algorithm.*

Theorem 2.2. *For every $\delta > 0$ and $L \in \mathbb{N}$ there exist constants $\varepsilon_0 = \varepsilon_0(\delta, L) > 0$ and $n_0 = n_0(\delta, L)$ such that for any $0 < \varepsilon < \varepsilon_0$ the following holds. Let G be a (p, λ)-bijumbled graph on $n \ge n_0$ vertices with $p \in (0, 1/2], \lambda \le \varepsilon p^2 n$ and minimum degree $\delta(G) \ge \delta pn$. Then in polynomial time, one can find any family of vertex-disjoint cycles with lengths in the interval $[4, L]$ whose lengths sum up to at most n.*

Theorem 2.3. *For every $\delta > 0$ there exist constants $L \in \mathbb{N}, \varepsilon_1 > 0$ and n_0 such that the following holds. For any $p \in (0, 1/3]$ and $0 < \varepsilon < \varepsilon_1$, let G be a (p, λ)-bijumbled graph on $n \ge n_0$ vertices with $\lambda \le \varepsilon p^2 n$ and minimum degree $\delta(G) \ge \delta pn$. Then in polynomial time, one can find any family of vertex-disjoint cycles with lengths in the interval $[L + 1, n]$ whose lengths sum up to at most n.*

Now we can quickly derive Theorem 1.2.

Proof of Theorem 1.2. We consider three (not mutually exclusive) cases

(1) there is subset of at least $n/2$ vertices of F which induce a collection of vertex-disjoint triangles in F,
(2) there is a subset of at least $n/4$ vertices of F which induce a collection of vertex-disjoint cycles with lengths in the interval $[4, L]$ where L is some absolute constant determined by Theorem 2.3 above,

(3) there is a vertex subset of at least $n/4$ vertices of F which induce a collection of vertex-disjoint cycles with lengths in the interval $[L+1, n]$.

For a given 2-factor F on at most n vertices, we are in one of the three cases defined above. Let F_1, F_2, and F_3 denote the subgraphs of F, so that all triangles constitute F_1, all cycles with lengths in $[4, L]$ constitute the subfamily F_2 and all cycles of length at least $L+1$ are F_3. We set $n_i := v(F_i)$ for each $i \in [3]$.

If we are in the first case ($n_1 \geq n/2$) we partition the vertex set V of G into three parts $V_1 \dot\cup V_2 \dot\cup V_3$, so that $|V_1| = n/2$ and $|V_3| = |V_4| = n/4$ and each $G[V_i]$ remains a (p, λ)-bijumbled graph. Moreover, every vertex $v \in V$ satisfies $\deg(v, V_i) \geq \delta p|V_i|/2$ for any $i \in [3]$. Clearly, one could achieve this via a random partition and it will be possible to derandomise this approach (see Corollary 3.8). If $n_2 \geq n_3$, then we first apply Theorem 2.3 to embed F_3 via some embedding φ_3 into $G[V_3]$. Then Theorem 2.2 asserts that F_2 can be embedded into $G_2' := G[(V_2 \dot\cup V_3)\backslash\varphi_3(V(F_3))]$, since G_2' is itself a (p, λ)-bijumbled graph with minimum degree at least $\delta pn/4$. Finally, we apply Theorem 2.1 to embed F_1 into the remaining graph (which is again (p, λ)-bijumbled graph with minimum degree at least $\delta pn/2$). If $n_3 \geq n_2$ then we first embed F_2, then F_3 and, finally, F_1. The other cases $n_2 \geq n/4$ and $n_3 \geq n/4$ are treated analogously. $\square$

2.1 | Structure of the paper

It remains to prove Theorems 2.1 to 2.3. We will only consider the case $p \leq 1/3$, since the dense case can be treated fairly easily by the algorithmic version of the blow-up lemma due to Komlós, Sárközy and Szemerédi [33]. In Section 3 we collect some notation and useful tools and algorithms for our study. In the subsequent two sections we prove the first two theorems (Theorems 2.2 and 2.3) and in Section 6 we replace one nonalgorithmic argument from [44] with a constructive proof.

Throughout we use the shorthand (p, λ)-graphs to refer to (p, λ)-bijumbled graphs, we write log to denote the natural logarithm and we omit floor and ceiling signs in order not to clutter the arguments. The final section closes with some problems left for further study.

3 | AUXILIARY RESULTS

3.1 | Simple statements about (p, λ)-bijumbled graphs

In this section, we collect some useful properties of (p, λ)-graphs. We will use the following notation. Given a graph $G = (V, E)$, we denote by $\deg(v, U)$ the number of neighbours of $v \in V$ in $U \subseteq V$. A u-v-path is a path P with end vertices u and v, and we call the other vertices of P the *inner vertices*. For vertex subsets A, B, an A-B path is a u-v path for some vertices $u \in A$ and $v \in B$. The length of a path is the number of its edges. Finally we will denote by $C_\ell(1, \ldots, 1, K)$, the graph that consists of a path P of length $\ell - 2$, whose end vertices have exactly K distinct common neighbours outside of $V(P)$, for some $K \in \mathbb{N}$. We start with the following remark which follows directly from the definition (2).

Remark 3.1. If $\varepsilon > 0$ and A and B are subsets of a (p, λ)-graph with $\lambda \le \varepsilon p^2 n$, such that $|A||B| \ge 4\varepsilon^2 p^2 n^2$, then $e(A, B) \ge \frac{p|A||B|}{2}$.

Next, we show that bijumbled graphs cannot be too sparse.

Proposition 3.2. *Given $\varepsilon \in (0, 1)$, there exists $n_0 \in \mathbb{N}$ such that if $G = (V, E)$ is a (p, λ)-graph on $n \ge n_0$ vertices with $\lambda \le \varepsilon p^2 n$ and $\varepsilon p \le 1/2$, then $p \ge (\varepsilon^2 n)^{-1/3}/4$.*

Proof. Let $S \subseteq V$ be a set of at least $n/2$ vertices. Then there is a vertex $v \in S$ whose degree in $G[S]$ is at most $2p|S|$. Indeed, we have $\sum_{v\in S}\deg(v, S) = 2e(G[S]) \le p|S|^2 + \lambda|S|$, which implies that the average degree in $G[S]$ is at most $p|S| + \lambda \le 2p|S| \le 2pn$.

We consecutively find vertices $v_1, \ldots, v_t$ with $t = n/(2 + 4pn)$ such that setting $V_i := V\backslash\{v_1, \ldots, v_{i-1}\}$, we have $\deg(v_i, V_i) \le 2p|V_i| \le 2pn$. Thus setting $U := \{v_i : i \in [t]\}$ and $W := V\backslash(U \cup \bigcup_{i\in[t]} N(v_i))$, we have that $|W| \ge n - t(1 + 2pn) = n/2$ and $e_G(U, W) = 0 \ge p|U||W| - \lambda\sqrt{|U||W|}$.

It follows that $\varepsilon p^2 n \ge \lambda \ge p\sqrt{tn/2}$. Thus, $2\varepsilon^2 p^2 n \ge t = n/(2 + 4pn)$, which implies $p^2 \ge (\varepsilon^{-2}/4)\min\{1/(4pn), 1/2\}$. Rearranging we get $p \ge \min\{(\varepsilon^2 n)^{-1/3}/3, 1/(2\sqrt{2}\varepsilon)\} \ge (\varepsilon^2 n)^{-1/3}/4$ for n sufficiently large. □

The following fact also concerns the edge distribution of bijumbled graphs.

Fact 3.3. Let $\varepsilon > 0$ and G be a (p, λ)-graph on n vertices with $p \in (0, 1]$ and $\lambda \le \varepsilon p^2 n$.

(i) If U is a set of vertices, then there are at most $4\varepsilon^2 p^2 n^2/|U|$ vertices w in G with $|N_G(w) \cap U| < p|U|/2$.
(ii) Given an integer t and vertex sets $U_1, \ldots, U_t, W$ such that $|W| > \sum_{i=1}^{t} 4\varepsilon^2 p^2 n^2/|U_i|$, we can find a vertex $w \in W$ such that $|N_G(w) \cap U_i| \ge p|U_i|/2$ for all $i \in [t]$, in polynomial time.

Proof. Let U' be the set of vertices w such that $|N_G(w) \cap U| < p|U|/2$. From (2) we have $|U'|\, p|U|/2 > e(U', U) \ge p|U||U'| - \lambda\sqrt{|U||U'|}$. The conclusion follows from rearranging.

By the first part of the fact, W clearly contains a desired vertex. We find it by screening the degree of any vertex of W into each U_i, which takes polynomial time. □

Next, given three sets A, B, and C, we show how to find an A-B-path of given length such that the inner vertices are from C.

Proposition 3.4. *Let $\varepsilon > 0$, $\ell \in \mathbb{N}$ and G be a (p, λ)-graph on n vertices with $p \in (0, 1]$, $\lambda \le \varepsilon p^2 n$ and $\varepsilon pn \ge 1$. If A and B are sets of at least $2^{\ell-1}\varepsilon pn$ vertices and C is a set of at least $2^{\ell-1}\varepsilon n$ vertices, then in polynomial time, we can find an A-B-path P of length ℓ whose inner vertices lie in C.*

Proof. If $\ell = 1$ then we have $e(A, B) > p|A||B| - \lambda\sqrt{|A||B|} \ge \sqrt{|A||B|}(p\varepsilon pn - \lambda) \ge 0$, namely, there is an edge with one end in A and the other in B. We can find such an edge by searching the neighbourhoods of vertices in A one by one. We proceed now inductively and we assume that $\ell \ge 2$ and the assumption holds for $\ell - 1$.

By Fact 3.3 (ii) we find a vertex $a \in A$ with degree at least $p|C|/2$ into C in polynomial time. Applying our inductive hypothesis to $N(a) \cap C, B \backslash \{a\}$ and $C \backslash \{a\}$ we find an $(N(a) \cap C)$-$(B \backslash \{a\})$-path of length $\ell - 1$ with inner vertices in C, which together with a yields the desired path of length ℓ. □

We will use copies of $C_\ell(1, ..., 1, K)$ in our absorbing structure. The following simple fact asserts that we can find these copies in any large enough set of vertices.

Fact 3.5. Let $\varepsilon > 0, K \in \mathbb{N}$ and let G be a (p, λ)-graph on n vertices with $p \in (0, 1]$ and $\lambda \leq \varepsilon p^2 n$. Let $\varepsilon p^2 n \geq K/4$, $\ell \geq 4$ and U be a set of at least $2^\ell \varepsilon n$ vertices. Then we can find a copy of $C_\ell(1, ..., 1, K)$, and thus also a copy of C_ℓ in U, in polynomial time.

Proof. Let U_1' be the set of vertices $v \in U$ with $|N(v) \cap U| < p|U|/2$. Since $|U| \geq 2^\ell \varepsilon n$, Fact 3.3 (i) implies that $|U_1'| \leq \varepsilon p^2 n/4$. We fix a vertex $u_1 \in U \backslash U_1'$, that is, $\deg(u_1, U) \geq p|U|/2$. Let U_2' be the set of vertices $v \in U$ with $|N(v) \cap (N(u_1) \cap U)| < p|N(u_1) \cap U|/2$. Since $|N(u_1) \cap U| \geq p|U|/2 \geq 8\varepsilon pn$, Fact 3.3 (i) implies that $|U_2'| \leq \varepsilon pn/2$. Thus, we have $|U_1' \cup U_2'| \leq \varepsilon pn$.

We choose an arbitrary vertex $u_2 \in U \backslash (U_1' \cup U_2' \cup \{u_1\})$. If $\ell = 4$ then we clearly find a copy of $C_\ell(1, 1, 1, K)$ in U, because $|N(u_1) \cap N(u_2) \cap U| \geq p^2|U|/4 \geq K + 1$. If $\ell \geq 5$, then we first set aside a set W of K vertices from the common neighbourhood of u_1 and u_2. Now due to the fact that $|(N(u_i) \cap U) \backslash (W \cup \{u_1, u_2\})| \geq 2^{\ell-2} \varepsilon pn$ for $i = 1, 2$ and $|U \backslash (W \cup \{u_1, u_2\})| \geq 2^{\ell-2} \varepsilon n$, we find by Proposition 3.4 a path of length $\ell - 4$ between $N(u_1) \cap U$ and $N(u_2) \cap U$, which together with u_1, u_2, and W, forms a copy of $C_\ell(1, ..., 1, K)$.

For the running time, by Fact 3.3 (ii), we can find u_1 and u_2 in polynomial time and the rest of the proof runs in polynomial time because we use Proposition 3.4. □

The following lemma asserts that we can (greedily) find almost spanning paths in (p, λ)-graphs.

Lemma 3.6. *Let $\varepsilon > 0$ and G be a (p, λ)-graph on n vertices with $p \in (0, 1/2]$ and $\lambda \leq \varepsilon p^2 n$. If U is a vertex subset of size greater than εn, then we can find any path of length $\ell \leq |U| - \varepsilon n$ in U in polynomial time.*

Proof. By Fact 3.3 there is a vertex $u \in U$ of degree at least $p|U|/2$ in U. This gives us a path of length 0. Assume now that we found inductively a path $P_t = u_0 u_1 ... u_t$ of length $t \leq \lfloor |U| - \varepsilon n \rfloor - 1$ such that $\deg(u_t, U \backslash V(P_t)) \geq p|U \backslash V(P_t)|/2$. Then by Fact 3.3 (i), as $|U \backslash V(P_t)| \geq \varepsilon n$ and using that $p \leq 1/2$, we have that there exists a vertex $u_{t+1} \in N(u_t) \cap (U \backslash V(P_t))$ with $\deg(u_{t+1}, U \backslash V(P_t)) \geq p|U \backslash V(P_t)|/2$ and the induction step is complete.

Since the proof is a repeated application of Fact 3.3 (i), Fact 3.3 (ii) implies that the running time is polynomial. □

3.2 | Partitioning vertex sets

At various points in our proof, we will wish to partition our vertex set in such a way that every vertex maintains good degree to all parts of the partition. This can be easily achieved probabilistically by choosing a random partition. However this idea can also be derandomised and achieved computationally efficiently. We use the following theorem of Alon and Spencer.

Theorem 3.7 (Theorem 16.1.2 in Alon and Spencer[13]). *Let $(a_{ij})_{i,j=1}^n$ be an $n \times n$ 0/1-matrix. Then one can find, in polynomial time, $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 1\}$ such that for every $1 \le i \le n$, it holds that $|\sum_{j=1}^n \varepsilon_j a_{ij}| \le \sqrt{2n\log(2n)}$.*

Corollary 3.8. *Let $k \in \mathbb{N}\varepsilon, \beta, \delta > 0$ and $p \in (0, 1]$. Then there exists $n_0 \in \mathbb{N}$ such that for any (p, λ)-graph G on $n \ge n_0$ vertices such that $\lambda \le \varepsilon p^2 n$, the following holds. Let $U, W \subseteq V(G)$ be subsets of vertices such that $|U| \ge \beta n$ and for all $w \in W$, $\deg(w, U) \ge \delta p|U|$. Then in polynomial time, we can find $s := 2^k$ sets $U_1, \ldots, U_s \subseteq U$ such that $U = U_1 \dot\cup \cdots \dot\cup U_s$ forms an equipartition[6] of U and for all $w \in W$ and $i \in [s]$, $\deg(w, U_i) \ge \delta p|U_i|/2$.*

Proof. We apply Theorem 3.7 to the adjacency matrix of G, where we add an all one row and an all one column and impose that row i is all zero if $i \notin W$ and column j is all zero if $j \notin U$. We let $U_b' = \{j \in U : \varepsilon_j = (-1)^b\}$, for $b = 1, 2$. The last row of the matrix guarantees that $||U_1'| - |U_2'|| \le \sqrt{2(n+1)\log(2n+1)} =: g(n)$. The other rows guarantee that the vertices in W have good degree to both sets, so that after moving some vertices from one of the sets to another to balance $|U_1'|$ and $|U_2'|$, we have that for all $w \in W$, $\deg(w, U_i') \ge \delta p|U|/2 - 2g(n)$.

We can now apply the above procedure to each U_i', with the new minimum degrees. Repeating this k times, we end up with $U_1, \ldots, U_s$ as an equipartition of U such that for any $w \in W$, $\deg(w, U_i) \ge \delta p|U|/s - 2kg(n)$. Owing to Proposition 3.2, we are done because for sufficiently large n, $2kg(n) \le \delta\beta pn/(2s) \le \delta p|U|/(2s)$. □

3.3 | A connecting lemma

The lemma below allows us to close many paths (whose ends are "well-connected" into a large set) into cycles using short paths of a fixed prescribed length. In the following lemma a v-v-path refers to a cycle through v whose inner vertices are all the vertices of the cycle not equal to v.

Lemma 3.9 (Multiple connection lemma). *For every $0 < \beta, \delta' \le 1, \ell \ge 3$ there exists $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that for all $\varepsilon \in (0, \varepsilon_0)$ and $n \ge n_0$ the following holds. Let G be a (p, λ)-graph on n vertices with $p \in (0, 1]$ and $\lambda \le \varepsilon p^2 n$. Let U be a vertex subset of size at least βn and $(a_i, b_i)_{i\in[r]}$ a system of pairs of vertices in G, so that every vertex occurs at most twice in $(a_1, \ldots, a_r, b_1, \ldots, b_r)$ and U is disjoint from $\bigcup_i\{a_i, b_i\}$. If $r \le |U|/(8\ell)$ and $\deg(a_i, U), \deg(b_i, U) \ge \delta' p|U|$ for all $i \in [r]$ then the following holds. In polynomial time, we can find a family $\mathcal{Q}$ of length ℓ a_i-b_i-paths Q_i, whose inner vertices are pairwise disjoint and lie in U.*

Proof. Fix $\varepsilon_0 \le \delta'\beta 2^{-(\ell+6)}/\ell$. Firstly, using Corollary 3.8, in polynomial time, we can split U into $U = U_1 \dot\cup U_2$ such that $|U_1| = |U_2| = |U|/2$ and $\deg(a_i, U_b), \deg(b_i, U_b) \ge \delta' p|U|/4$ for

[6]Due to divisibility constraints, we formally mean here that the the sizes of the sets differ by at most one, and so $|U_i| \in \{\lfloor |U|/s\rfloor, \lceil |U|/s\rceil\}$ for each i.

all i and $b = 1, 2$. We will build our paths algorithmically in two phases, first using vertices of U_1 and then vertices of U_2. We initiate by letting $\mathcal{Q}' = \emptyset$, $U_1' = U_1$ and $U_2' = U_2$. We will use $\mathcal{Q}'$ to denote our intermediate family of paths and U_1', U_2' the remaining sets of vertices that we can use. Note that throughout we will have $|V(\mathcal{Q})| \le r\ell \le |U|/8$, and thus $|U_1'|, |U_2'|$ will always have size at least $|U|/4$.

We proceed as follows. If there is an $i \in [r]$ such that $\deg(a_i, U_1'), \deg(b_i, U') \ge \delta_P' |U|/8$, then using Proposition 3.4, in polynomial time we find a length $\ell - 2$ path P_i from a vertex in $N(a_i) \cap U_1'$ to a vertex in $N(b_i) \cap U_1'$ using vertices in U_1'. Add $Q_i := a_i\text{-}P_i\text{-}b_i$ to $\mathcal{Q}$ and delete the vertices of P_i from U_1'. At the end of this phase, let $I \subseteq [r]$ be the remaining indices. Since each vertex appears at most twice in $(a_i, b_i)_{i \in [r]}$, by Fact 3.3 (i), we have that

$$|I| \le \frac{4\varepsilon^2 p^2 n^2}{|U_1'|} \le \frac{4\varepsilon^2 p^2 n^2}{\beta n/4} \le 16\varepsilon_0 p^2 n \le \delta' p^2 |U|/(8\ell) \le \delta' |U|/(8\ell),$$

where we used $|U_1'| \ge |U|/4 \ge \beta n/4$, and $\varepsilon \le \varepsilon_0 \le \delta' \beta 2^{-(\ell+6)}/\ell$. Now we run the process again, using U_2 in place of U_1. As $|V(\mathcal{Q}') \cap U_2| \le \delta' p |U|/8$ throughout, we can proceed greedily by the degree assumptions and complete the family $\mathcal{Q}$. Note that in each step, we need to screen the degrees of the remaining pair a_i and b_i. The application of Proposition 3.4 then runs in polynomial time and so the whole algorithm runs in polynomial time. □

3.4 | An explicit template

A *template* T with *flexibility* $m \in \mathbb{N}$ is a bipartite graph on $7m$ vertices with vertex parts I and $J = J_1 \dot\cup J_2$, such that $|I| = 3m$, $|J_1| = |J_2| = 2m$, and for any $\bar{J} \subseteq J_1$, with $|\bar{J}| = m$, the induced graph $T[V(T) \backslash \bar{J}]$ has a perfect matching. We call J_1 the *flexible set* of vertices for the template.

Sparse templates, with maximum degree smaller than some absolute constant, are very useful in absorption arguments and can be used to design robust absorbing structures. Montgomery first introduced the use of such templates when applying the absorbing method in his work on spanning trees in random graphs [42, 43]. Ferber et al [22] followed the same argument as Montgomery (with some small adjustments) when studying the 2-universality of the random graph. Kwan [40] also used sparse templates to study random Steiner triple systems, generalising the template to a hypergraph setting and using it to define an absorbing structure for perfect matchings. Further applications were given by Ferber and Nenadov [23] in their work on universality in the random graph, recently by the current authors in [24] which was the first use of the method in the context of pseudorandom graphs, and by Nenadov and Pehova [45] who used the method to study a variant of the Hajnal-Szeméredi Theorem. The final three papers mentioned all adapt the method to give absorbing structures which output disjoint copies of a fixed graph H (a partial H-factor), however the different absorbing structures used are interestingly all significantly distinct.

It is not difficult to prove the existence of sparse templates for large enough m probabilistically; see for example [42, Lemma 2.8]. As we wish to give a completely algorithmic proof, in this section we show how to build a template T efficiently. We use the following result of Lubotzky et al [41].

Theorem 3.10 (Lubotzky et al [41]). *For primes $p, q \equiv 1 \pmod 4$ such that p is a quadratic residue modulo q, one can construct an explicit $(p+1)$-regular Ramanujan graph G in polynomial time (in q) with $(q^3 - q)/2$ vertices.*

A Ramanujan graph, by definition, is a d-regular graph all of whose eigenvalues (other than d and, if bipartite, $-d$) are in absolute value at most $2\sqrt{d-1}$. We will in fact use a bipartite Ramanujan graph constructed as follows. Consider the graph G provided by Theorem 3.10—take V_1 and V_2 as two identical copies of $V(G)$, and join $v_1 \in V_1$ and $v_2 \in V_2$ if and only if the preimages of v_1 and v_2 in $V(G)$ form an edge of G. It is clear that this bipartite Ramanujan graph is still d-regular and satisfies the expander mixing lemma (1) for all $A \subseteq V_1$ and $B \subseteq V_2$, where n is the number of vertices in each part, and $\lambda = 2\sqrt{d-1}$.

Proposition 3.11. *Let $d \geq 144/\alpha^2$. Let G be a bipartite d-regular Ramanujan graph on vertex set $V_1 \dot\cup V_2$, with $|V_1| = |V_2| = n$. Suppose $U \subseteq V_1$ and $W \subseteq V_2$ are vertex subsets of $V(G)$ such that $|U| = |W| = \alpha n$ and $\deg(w, U) \geq \alpha d/3$ for any $w \in W$ and $\deg(u, W) \geq \alpha d/3$ for any $u \in U$. Then $G[U, W]$ contains a perfect matching.*

Proof. We will verify Hall's condition for $G[U, W]$. First we claim that it suffices to prove the condition $|N(X) \cap W| \geq |X|$, for sets $X \subseteq U$ of size $|X| \leq \lceil |U|/2 \rceil = \lceil \alpha n/2 \rceil$. Indeed by symmetry, one can then conclude that Hall's condition holds for subsets of W which have size smaller than $\lceil \alpha n/2 \rceil$. Now for $X \subseteq W$ such that $|X| > \alpha n/2$, if $|N(X) \cap W| < |X|$, then setting Y' to be a subset of $W \backslash N(X)$ of size $\alpha n - |X| + 1$ we have that $|Y'| \leq \lceil \alpha n/2 \rceil$ and so from above we can conclude that $|N(Y') \cap U| \geq |Y'|$. This contradicts the definition of Y' as we must have that $N(Y')$ intersects X and hence Y' intersects $N(X)$.

So it remains to prove that, for $X \subseteq U$ with $|X| \leq \lceil \alpha n/2 \rceil$, taking $Y = N(X) \cap W$ we have that $|Y| \geq |X|$. Assume to the contrary that $|Y| < |X|$. We first assume that $|X| \leq \alpha n/6$. By the degree condition, we obtain that $e(X, Y) \geq |X|\, \alpha d/3$. In contrast, by (1), we have

$$e(X, Y) \leq \frac{d}{n}|X||Y| + \lambda\sqrt{|X||Y|} < \frac{\alpha d}{6}|X| + 2\sqrt{d|X||Y|}.$$

Putting these together, we get $2\sqrt{d|X||Y|} \geq \alpha d|X|/6$. By $d \geq 144/\alpha^2$, this implies $|Y| \geq |X|$, a contradiction. Next we assume that $\alpha n/6 < |X| \leq \lceil \alpha n/2 \rceil$. By $|W \backslash Y| \geq \alpha n/2$, (1) and the fact that $\alpha^2 d \geq 144$ we have that

$$e(X, W\backslash Y) \geq \left(\frac{d}{n}\sqrt{|X||W\backslash Y|} - \lambda\right)(\sqrt{|X||W\backslash Y|}) > \alpha^2 dn/12 - 2\sqrt{d\alpha^2 n^2/12} > 0,$$

contradicting the definition of Y. □

Lemma 3.12. *Let $p \equiv 1 \pmod 4$ be a prime such that $p \geq 68000$. For a sufficiently large integer m, a template with flexibility m and maximum degree $d := p + 1$ can be constructed in polynomial time.*

Proof. It follows from the Siegel-Walfisz theorem [50] that for large x, the density in $[x]$ of the set of primes which are $1 \pmod{4p}$ is of the order $\Theta(1/\log x)$. Therefore for all sufficiently large m we can pick a prime $q \equiv 1 \pmod{4p}$ between $(21m)^{1/3}$ and $1.01(21m)^{1/3}$. Thus $20m \leq q^3 - q \leq 22m$. Using quadratic reciprocity to infer that p is a quadratic residue modulo q, we have by Theorem 3.10 that we can construct in polynomial time a bipartite d-regular Ramanujan graph $G = (X \cup Y, E)$ with $10m \leq |X| = |Y| \leq 11m$ and second eigenvalue $\lambda \leq 2\sqrt{d}$. We first show that for any set $U \subseteq X$ (or Y) of size at least $3m/2$, there are at most $34000m/d$ vertices v in Y (or X) such that $\deg(v, U) < d/10$. Indeed, denote by B the set of such vertices v. Clearly we have $e(U, B) < d|B|/10$. In contrast, by (1), we have

$$\frac{d|B|}{10} > e(U, B) \geq \frac{d}{11m}|B||U| - \lambda\sqrt{|B||U|} \geq \frac{3d|B|}{22} - 2\sqrt{d|B| \cdot 11m}.$$

This implies that $2d|B|/55 < 2\sqrt{11d|B|\ m}$, that is, $|B| < 33275m/d < 34000m/d$, as claimed.

Now take arbitrary sets $V_1'' \subseteq X$, $V_2'' \subseteq Y$ such that $|V_1''| = 3m$ and $|V_2''| = 2m$. Next, we sequentially delete vertices from V_1'' and V_2'' as follows.

- Initiate with $V_i' := V_i''$ for $i = 1, 2$.
- If there is a vertex $v \in V_1'$ such that $\deg(v, V_2') < d/10$, then delete v from V_1',
- If there is a vertex $v \in V_2'$ such that $\deg(v, V_1') < d/10$, then delete v from V_2'.

Note that since $|V_i''| - 34000m/d \geq 3m/2$, by our claim above, at most $34000m/d$ vertices will be deleted from each set. Denote by V_1' and V_2' the resulting sets. Next, since there are at most $34000m/d$ vertices that have degree less than $d/10$ to $V_i{}'$, $i = 1, 2$, respectively, we can add vertices to V_1' and V_2' and obtain V_1 and V_2 such that $|V_1| = 3m$, $|V_2| = 2m$ and $\deg(v, V_i) \geq d/10$ for any $v \in V_{3-i}$, $i = 1, 2$. Finally, we pick J_1 as a set of $2m$ vertices in $Y \backslash V_2$ which have degree at least $d/10$ to V_1.

We claim that $T = G[V_1 \cup V_2 \cup J_1]$ is the desired template with flexible set J_1. It remains to check the property of T. For this, take any set J' of m vertices in J_1 and consider $G[V_1, V_2 \cup J']$. Since the assumptions of Proposition 3.11 are satisfied with $\alpha = 3m/|X| \in [3/11, 3/10]$, $G[V_1, V_2 \cup J']$ has a perfect matctehing and we are done. For the running time, note that in each of the steps above, it is enough to query the neighbourhood of a vertex, which can be done in constant time. So the overall running time is polynomial in m. □

4 | PROOF OF THEOREM 2.2

In [24] an *absorbing structure* for cliques was defined. Here we generalise it for cycles as follows. Assume that $T = (I, J_1 \cup J_2, E)$ is a bipartite template with flexibility m, maximum degree $\Delta(T) \leq K$ and flexible set J_1. It will be convenient to identify T with its edges which may be viewed as the corresponding subset of tuples $(i, j) \in [3m] \times [4m]$, hence we will also think of I as $[3m]$, J_1 as $[2m]$, J_2 as $[2m + 1, 4m] := \{2m + 1, \ldots, 4m\}$ and $J = J_1 \cup J_2$.

An *absorbing structure for cycles of length* $s + 2$ is a tuple $\mathcal{S} = (T, \mathcal{P}_1, A, \mathcal{P}_2, Z, Z_1)$ which consists of the template T with flexibility m, the two sets $\mathcal{P}_1$ and $\mathcal{P}_2$ of vertex-disjoint paths of fixed length s and three vertex sets A, Z, and Z_1 with $Z_1 \subseteq Z$. Furthermore, the sets $V(\mathcal{P}_1), V(\mathcal{P}_2), A$ and Z are pairwise disjoint and with the labelling $Z_1 = \{z_1, ..., z_{2m}\}$, $Z_2 = \{z_{2m+1}, ..., z_{4m}\}$ (so that $Z := Z_1 \cup Z_2$), $\mathcal{P}_1 := \{P^1, P^2, ..., P^{3m}\}$, $A = \{a_{ij} : (i, j) \in E(T)\}$ and $\mathcal{P}_2 = \{P_{ij}: (i, j) \in E(T)\}$, the following holds in G for $(i, j) \in E(T)$:

- a_{ij} is adjacent to the ends of P^i, that is, closes a cycle on $s + 2$ vertices,
- each a_{ij} is adjacent to the ends of P_{ij},
- each z_j is adjacent to the ends of P_{ij}.

It is well known that finding a maximum matching in a graph can be done in polynomial time. Using this and unpacking the definition of the absorbing structure leads to the following fact.

Fact 4.1. The absorbing structure $\mathcal{S} = (T, \mathcal{P}_1, A, \mathcal{P}_2, Z, Z_1)$ has the property that, for any subset $\bar{Z} \subseteq Z_1$ with $|\bar{Z}| = m$, the removal of $\bar{Z}$ leaves a graph with a C_{s+2}-factor, which can be found in polynomial time.

Proof. By the property of the template $T \subseteq [3m] \times [4m]$, there is a perfect matching M in $[3m] \times ([4m] \backslash \bar{J}) \cap T$ with $\bar{J} := \{j: z_j \in \bar{Z}\}$. Furthermore, we can find M in polynomial time.

Then for each edge $(i, j) \in M$, we take the $(s + 2)$-cycles on $\{a_{ij}\} \cup P^i$ and $\{z_j\} \cup P_{ij}$; for the edges $(i, j) \in E(T) \backslash M$, we take the $(s + 2)$-cycle on $\{a_{ij}\} \cup P_{ij}$. This gives the desired C_{s+2}-factor. □

The following lemma is a variant of Lemma 2.7 from [24].

Lemma 4.2. *Let* $K := 68042$. *For every* $\delta > 0, \ell \geq 4$ *and* $\alpha \in (0, \alpha(\ell)]$ *(where* $\alpha(\ell) := 1/(60\ell(K + 2))$*) there exists* $\varepsilon_0 > 0$ *such that for all* $\varepsilon \in (0, \varepsilon_0)$ *there is an* $n_0 \in \mathbb{N}$ *such that the following holds for all* $n \geq n_0$. *Let* G *be a* (p, λ)*-graph with* n *vertices,* $p \in (0, 1/3]$, $\lambda \leq \varepsilon p^2 n$, $\delta(G) \geq \delta pn$ *and suppose* $m = \alpha n$. *Then in polynomial time we can find an absorbing structure* $\mathcal{S} = (T, \mathcal{P}_1, A, \mathcal{P}_2, Z, Z_1)$ *for cycles of length* ℓ *with flexibility* m *in* G. *Further, one can find in polynomial time a set* $W \subseteq V(G) \backslash V(\mathcal{S})$, *with* $|W| = n/4$ *and* $\deg(v, W) \geq \delta p|W|/8$ *for all vertices* v *of* G.

Proof. First we choose $\varepsilon_0 = \min\{\delta/(400K\ell), 2^{-(\ell+6)}, \alpha\}$ and let $\varepsilon \in (0, \varepsilon_0)$. Then we take n_0 large enough. Therefore, owing to Proposition 3.2, quantities p^2n and pn are large as well.

We consider a partition of $V(G) = V_1 \dot\cup V_2 \dot\cup V_3 \dot\cup V_4$ with $|V_1| = |V_2| = |V_3| = |V_4| = n/4$, such that

$$\deg(v, V_i) \geq \delta p|V_i|/2 \tag{3}$$

for all $i \in [4]$ and $v \in V$, as given by Corollary 3.8. We fix $W = V_4$ and thus the conditions on W are satisfied. We now build our absorbing structure using vertices of $V(G) \backslash W$. Throughout the proof, we denote the intermediate partial absorbing structure by $\mathcal{S}'$. Note that an absorbing structure for cycles of length ℓ with flexibility m which uses a template

T has at most $3\ell m(\Delta(T)+2)$ vertices, and thus, due to the condition on α and the fact that we will have $\Delta(T) \le K$, we will have that $|V(\mathcal{S}')| \le n/20$ throughout the proof.

Let $T \subseteq [3m] \times [4m]$ be a bipartite template with flexibility m and flexible set $J_1 = [2m]$ such that $\Delta(T) \le K$, as provided by Lemma 3.12. Pick an arbitrary collection of $3m$ vertex-disjoint copies of $C_\ell(1, \ldots, 1, K)$ in V_1 (using Fact 3.5). For the ith copy of $C_\ell(1, \ldots, 1, K)$, we label the corresponding path on $\ell - 2$ edges by P^i (so that the ends of P^i have K common neighbours), and we set $\mathcal{P}_1 := \{P^1, P^2, \ldots, P^{3m}\}$. Then we label $A = \{a_{ij}: (i,j) \in E(T)\}$ as the vertices in the classes of K vertices in the copies of $C_\ell(1, \ldots, 1, K)$ such that each a_{ij} is connected to the ends of P^i, that is, forms a copy of C_ℓ (we may then discard some extra vertices, according to the degree of i in T).

We will pick $Z = \{z_1, \ldots, z_{4m}\}$ and $\mathcal{P}_2 = \{P_{ij} : (i,j) \in E(T)\}$ satisfying the definition of the absorbing structure as follows. We choose Z in two phases, where all but at most $\varepsilon p^2 n$ vertices for Z will be chosen in the first phase. We first use vertices in V_1. We recursively do the following. We pick the smallest index $j \in [4m]$ (as long as there exists such an index) so that $|N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')| \ge \delta pn/10$ for all i such that $(i,j) \in T$ (there are at most K such i). We pick as z_j an arbitrary vertex in $V_2 \setminus (V(\mathcal{S}') \cup B_j)$, where B_j is the set of vertices z in G such that $|(N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')) \cap N_G(z)| < \delta p^2 n/20$ for some i with $(i,j) \in E(T)$. Since $|N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')| \ge \delta pn/10$ and $\Delta(T) \le K$, Fact 3.3 (i) with $U = N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')$ implies that $|B_j| \le 40K\delta^{-1}\varepsilon^2 pn \le n/8$, and so such a choice always exists.

Having chosen z_j, our next aim is to construct vertex-disjoint paths P_{ij} of length $\ell - 2$, for each $(i,j) \in E(T)$, so that the endpoints of P_{ij} are adjacent to both a_{ij} and z_j. For this purpose, we would like to pick two vertices y_1, y_2 in $U_{ij} := (N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')) \cap N_G(z_j)$, which are supposed to be the ends of the path P_{ij} which we are going to construct. Since $z_j \notin B_j$, we have $|U_{ij}| \ge \delta p^2 n/20$. Letting $V_1' := V_1 \setminus (V(\mathcal{S}') \cup U_{ij})$, we have that $|V_1'| \ge n/8$. From Remark 3.1, we get that $e(V_1', U_{ij}) \ge p|U_{ij}||V_1'|/2$. We consider two cases. If $\ell = 4$ then, since there is a vertex w from V_1' of degree at least $p|U_{ij}|/2 \ge \delta p^3 n/40 \ge 2$ into U_{ij} (by Proposition 3.2), there is a path P_{ij} of length 2 with ends (labeled as) y_1 and y_2 in U_{ij}. If $\ell \ge 5$, then by Fact 3.3 (i) and the choice of ε, we can find two vertices y_1 and $y_2 \in U_{ij}$, whose degrees into V_1' are at least $pn/30$. Proposition 3.4 then yields the existence of a path of length $\ell - 4$ with ends in $N(y_1) \cap V_1'$ and $N(y_2) \cap V_1'$. Together with y_1 and y_2 this provides us with the desired path P_{ij}.

It remains still to deal with the situation (second phase), when there are no remaining appropriate indices $j \in [4m]$. Let $\tilde{J} \subseteq [4m]$ be the set of those indices j such that for some $\{i,j\} \in T$ we have $|N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')| < \delta pn/10$. Since $|V_1 \setminus V(\mathcal{S}')| \ge n/5$ we have with Fact 3.3 (i) and $\Delta(T) \le K$ that $|\tilde{J}| \le K(20\varepsilon^2 p^2 n) \le \varepsilon p^2 n$. To finish the embedding, we will use vertices in V_3 as well. At any point we will have that $|V(\mathcal{S}') \cap V_3| \le K|\tilde{J}|\ \ell \le \delta pn/40$. From (3) we get $\deg(v, V_3 \setminus V(\mathcal{S}')) \ge \delta pn/10$ for all vertices $v \in V(G)$ throughout the process and we can proceed as in the two paragraphs above, using V_3 in place of V_1.

Now we analyse the running time. Firstly, we pick the copies of $C_\ell(1, \ldots, 1, K)$ by Fact 3.5. Second, to find a desired $j \in [4m]$, we check $|N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')|$ for all vertices a_{ij}; with such a j, to choose z_j, we search through the vertices z not in $V(\mathcal{S}')$ and check $|(N_G(a_{ij}, V_1) \setminus V(\mathcal{S}')) \cap N_G(z)|$ for at most K such i's. By Fact 3.3 (ii), this takes polynomial time. At last, we pick the desired path P_{ij} of length $\ell - 2$. If $\ell = 4$, then we find the vertex $w \in V_1'$ with degree 2 to U_{ij} and hence P_{ij} in polynomial time. If $\ell \ge 5$, we find y_1 and y_2 and then apply Proposition 3.4, which runs in polynomial time. Therefore, the overall running time is polynomial since partitioning as is done by Corollary 3.8 works in polynomial time as well. □

Now we are ready to prove Theorem 2.2.

Proof of Theorem 2.2. Let $K = 68042$. Let $L_0 := \min\{2^k : k \in \mathbb{N}, 2^k > L\}$ and fix $\varepsilon \le \varepsilon_0 := \min\{\delta/(8000KL2^{(L_0+6)}), \varepsilon_{4.2}, \varepsilon_{3.9}\}$, where $\varepsilon_{4.2}$ is as asserted by Lemma 4.2 on input $\delta' := \delta/2L_0, \alpha(L)$ and $\varepsilon_{3.9}$ is as asserted by Lemma 3.9 on input $\beta = 1/(60L(K+2)), \delta'$ and L. Let n_0 be large enough. First, using Corollary 3.8, we find a partition of the vertex set of G into sets $V_1 \dot\cup V_2$ such that $|V_1| = n/L_0$ and every vertex $v \in V(G)$ satisfies

$$\deg(v, V_i) \ge \delta p |V_i|/2, \tag{4}$$

for $i \in [2]$. Here V_2 is taken to be the union of all other sets in the equipartition given by Corollary 3.8, thus $|V_2| = (L_0 - 1)n/L_0$. Let F be a collection of cycles of lengths in the interval $[4, L]$, whose lengths sum up[7] to n. There is (at least) one length $\ell \in [4, L]$ such that F contains at least $n/((L-3)\ell)$ cycles C_ℓ. We write $F = F' \dot\cup F_\ell$, where F_ℓ consists of cycles of length ℓ from F, while F' contains all other cycles. We will embed F into G in two stages. First, we greedily embed F' into $G[V_2]$. This is possible since

$$|V(F')| \le \frac{(L-4)n}{L-3} \le \frac{(L_0-4)n}{L_0-3} = \frac{(L_0-1)n}{L_0} - \frac{3n}{L_0(L_0-3)} = |V_2| - \frac{3n}{L_0(L_0-3)}$$

and since any set of at least $3n/(L_0(L_0-3))$ vertices in G contains a cycle of any length from the interval $[4, L]$ (see Fact 3.5).

In the second stage we are left with a vertex set $U \supseteq V_1$ such that $|V(F_\ell)| = |U|$ and $\delta(G[U]) \ge \delta pn/(2L_0) \ge \delta' p|U|$, due to (4). All that remains to do is to find a C_ℓ-factor in $G[U]$. We are thus in a position to apply Lemma 4.2 to $G[U]$, where one can check that the conditions there are satisfied with respect to $|U|$ and δ'. Thus, in polynomial time we can construct an absorbing structure $\mathcal{S} = (T, \mathcal{P}_1, A, \mathcal{P}_2, Z, Z_1)$ for cycles of length ℓ with flexibility $m = \alpha|U|$, where $\alpha := 1/(60L(K+2)) \le \alpha(\ell)$, and a vertex set $W \subseteq V(G)\backslash V(\mathcal{S})$, with $|W| = |U|/4$, such that for any vertex v in G, we have $\deg(v, W) \ge \delta' p|U|/8$. Let $U_0 \subseteq (U \backslash V(\mathcal{S}))$ be the set of vertices u such that $\deg(u, Z_1) \le p|Z_1|/2$. By Fact 3.3 (i), we have that $|U_0| \le 4\varepsilon^2 p^2 |U|^2/|Z_1| = 2\varepsilon^2\alpha^{-1}p^2|U|$. We first incorporate the vertices of U_0 into cycles of length ℓ using vertices of $W \backslash U_0$ by applying Lemma 3.9 (in polynomial time) to the pairs $\{(u, u) : u \in U_0\}$. Let $\mathcal{C}_1$ be the set of disjoint cycles produced by this process.

Now we greedily apply Fact 3.5 to find vertex-disjoint cycles C_ℓ in $G[U \backslash (V(\mathcal{S}) \cup V(\mathcal{C}_1))]$, until we are left with a set U_1 of cardinality at most $2^\ell \varepsilon n$. What remains is to find a C_ℓ-factor in $G[U_1 \dot\cup V(\mathcal{S})]$. Recall that $\deg(u, Z_1) \ge p|Z_1|/2$ for every $u \in U_1$. The assumptions of Lemma 3.9 are met (in particular $|Z_1| \gg |U_1|$), and therefore, applying it to the pairs of vertices $\{(u, u) : u \in U_1\}$ (to find paths through Z_1) we find a family $\mathcal{C}_2$ of $|U_1|$ vertex-disjoint cycles C_ℓ that cover all of U_1 (and some subset of Z_1). Next, we greedily find, applying

[7] We can assume that F has n vertices as if not, we can take a supergraph by adding 4-cycles repeatedly. We can then remove up to three vertices from G without affecting the properties of G as in the statement of Theorem 2.2.

Fact 3.5, $(m - |U_1|(\ell - 1))/\ell$ cycles C_ℓ in $Z_1 \backslash V(\mathcal{C}_2)$, so a set Z_1' of exactly[8] m vertices of Z_1 remains uncovered. But then, letting $Z_1'' = Z_1 \backslash Z_1'$, Fact 4.1 guarantees the existence of a C_ℓ-factor on $V(\mathcal{S}) \backslash Z_1''$. This then gives us a copy of F in G.

Note that we applied Fact 3.5 linearly many times. Moreover, we applied Corollary 3.8, Lemma 3.9, Fact 4.1, and Lemma 4.2 constantly many times. So we conclude that we can indeed find a copy of F in polynomial time. □

5 | PROOF OF THEOREM 2.3

Before proving Theorem 2.3, let us sketch some of the ideas that arise in the proof. First, we will apply Lemma 4.2 to show the existence of an absorbing structure $\mathcal{S} = (T, \mathcal{P}_1, A, \mathcal{P}_2, Z, Z_1)$ for cycles of length 4 with flexibility $m = \lfloor \gamma n \rfloor$, with $\gamma \leq \alpha(4) = 1/(240(K+2))$, as defined in Lemma 4.2. Recall that Fact 4.1 guarantees that no matter which m vertices of Z_1 we remove, on the rest of the vertices of $\mathcal{S}$ we can find a C_4-factor (which will contain exactly $3m + |E(T)|$ copies of C_4). Let us relabel the $r := 3m + |E(T)|$ paths of length two in $\mathcal{P}_1 \cup \mathcal{P}_2$ as $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_r\}$, let $Q_h = a_h b_h c_h$ for each $h \in [r]$ and let $Y = Z \,\dot\cup\, A$. Now the property of the absorbing structure can be rephrased as follows. After removing exactly m vertices, Z', from $Z_1 \subseteq Y$, there is a perfect matching between $\mathcal{Q}$ and $Y \backslash Z'$ such that if $Q_h \in \mathcal{Q}$ is matched with $y \in Y$, then $a_h y c_h b_h$ forms a copy of C_4. In what follows, the idea is to omit an edge (eg, $a_h b_h$) from each of these C_4 to get paths of length three which we will connect to longer paths. The key point is that we can do this by only omitting edges in the length two paths from $\mathcal{Q}$. Thus we can simply connect vertices from paths in $\mathcal{Q}$ through short connecting paths. Eventually, this will lead to a longer path that will contribute to our factor and although we do not know exactly what these paths will be (as it depends on the choice of matching to $y \in Y$), the lengths of the paths and the vertices not in Y are fixed. More precisely, we will group the paths in $\mathcal{Q}$ according to the desired lengths of the cycle and connect the ones in the same group, for example, connect a_h with b_{h-1} and connect b_h with a_{h+1}. At the end of the proof, by Fact 4.1 we can match every remaining vertex $y \in Y$ to one of the Q_h's, such that $a_h y c_h b_h$ forms a copy of P_3 which will contribute to some longer path which in turn is part of a cycle in F.

Proof of Theorem 2.3. Let $K = 68042$. Let $L \geq 8000K$ and fix $\gamma := 1/(600(K+2)) \leq \alpha(4)$ with $\alpha(4)$ defined in Lemma 4.2. Next, choose $\varepsilon \leq \varepsilon_1 := \min\{\delta/(1600000K), \varepsilon_{4.2}, \varepsilon_{3.9}\}$, where $\varepsilon_{4.2}$ is as asserted by Lemma 4.2 on input δ, $\alpha = \gamma$, $\ell = 4$, and $\varepsilon_{3.9}$ is as asserted by Lemma 3.9 on input $\beta = \gamma$, $\delta' := \delta/16$ and $\ell = 3$. Let n_0 be large enough. Let F be a graph whose components are cycles of length greater than L. We can assume that $v(F) \geq n - L$, otherwise we can instead consider a supergraph by adding cycles of length $L + 1$. Let F consist of t cycles of lengths $l_1 \geq \cdots \geq l_t$, and let $l = \sum_i^t l_i$. Note that $t \leq n/L$ and $n - L \leq l \leq n$. We will exhibit an algorithm which finds $F \subseteq G$.

Let $m = \gamma n$. Apply Lemma 4.2 to get an absorbing structure $\mathcal{S} = (T, \mathcal{P}_1, A, \mathcal{P}_2, Z, Z_1)$ for cycles of length 4 with flexibility m and a vertex set $W \subseteq V(G) \backslash V(\mathcal{S})$, with $|W| = n/4$, such that for any vertex v in G, we have $\deg(v, W) \geq \delta p |W|/8$. Label the vertices and paths of $\mathcal{S}$ as in the discussion above. In particular, recall that $r := 3m + |E(T)|$. Let

[8]Note that this is possible due to divisibility conditions. Indeed it is clear that $\ell | (|U_1| + |V(\mathcal{S})|)$ as we look to find the remaining C_ℓ-factor on $U_1 \cup V(\mathcal{S})$. Also, Fact 4.1 guarantees that $\ell | (|V(\mathcal{S})| - m)$ and so we can conclude that $\ell | (m + |U_1|)$ and hence $\ell | (m - |U_1|(\ell - 1))$ as required.

$m' := \varepsilon n$, and let $Z' \subseteq Z_1$ be an arbitrary subset of size $m + 2m' + 4t$. Let $V_0 \subseteq (V(G) \backslash V(\mathcal{S})) \cup (Z_1 \backslash Z')$ be the set of vertices v such that $\deg(v, Z') \le \frac{p|Z'|}{2}$. Write $V_0 := \{v_1, v_2, \ldots, v_{|V_0|}\}$. By Fact 3.3 (i), we have that $|V_0| \le 4\varepsilon^2\gamma^{-1}p^2 n$. We find nonnegative integers q_{ij}, $i \in [t], j \in [3]$ such that the following holds:

- $6q_{i1} + 3q_{i2} + 3q_{i3} \le l_i - 10$, for each $i \in [t]$,
- $\sum_{i=1}^{t} q_{i1} = r$, $\sum_{i=1}^{t} q_{i2} = |V_0|$, and $\sum_{i=1}^{t} q_{i3} = m'$.

Such a choice can be achieved easily since $r = 3m + |E(T)|$ *and* $6r + 3|V_0| + 3m' \ll l - 15t$.

We now run the first phase of our algorithm

(1) We arbitrarily partition the set $\{\{a_h, b_h, c_h\}, h \in [r]\}$ into t subsets of sizes $q_{11}, q_{21}, \ldots, q_{t1}$ and partition V_0 into t subsets of sizes $q_{12}, \ldots, q_{t2}$.
(2) For $i \in [t]$, we fix an arbitrary linear order of the q_{i1} triples of vertices and q_{i2} vertices of V_0, and insert two new vertices x_i^1, x_i^2 not in $W \cup V_0 \cup V(\mathcal{S})$ to the ordering, one to the beginning, one to the end. Apply Lemma 3.9 to the pairs $\{b_{h-1}, a_h\}$ of consecutive elements from each group simultaneously (we view each single vertex v in the ordering as $v = a_h = b_h$), and get disjoint length three paths through $W \backslash V_0$ joining the pairs. This is possible because the number of pairs we connect is at most $2t + r + |V_0| \le 2n/L + 3m(1 + K) + 4\varepsilon^2\gamma^{-1}p^2 n \le 2n/L + 3(K + 2)\gamma n < n/120$, and every vertex has degree at least $\delta p|W|/8 - |V_0| \ge \delta p|W|/9$ to $W \backslash V_0$, and $|W \backslash V_0| \ge n/5$.

For each $i \in [t]$, we obtain a sequence of paths on (in total) $5q_{i1} + 3q_{i2} + 3$ vertices (they will become a single path of length $6q_{i1} + 3q_{i2} + 3$ after absorbing *exactly* q_{i1} vertices from Z_1). Next we will greedily find paths for each $i \in [t]$ which will comprise the majority of the remainder of the cycles.

(1) Fix U to be the vertices in $(V(G) \backslash (V(\mathcal{S}))) \cup (Z_1 \backslash Z')$ which were not used in the paths chosen in the first phase. For $i \in [t]$, we repeatedly find a path of length exactly $l_i - 6q_{i1} - 3q_{i2} - 3q_{i3} - 9$ in the uncovered vertices of U using Lemma 3.6 (for this observe that there are at least $\ge \varepsilon n$ unused vertices from U by the choice of the parameters). Denote the endpoints of the path by x_i^3 and x_i^4.
(2) Arbitrarily choose m' vertices from U (it could happen that there are more vertices in U but only if F has less than n vertices), partition and label them in such a way that for each i there are q_{i3} vertices $u_{i,1}, \ldots, u_{i,q_{i3}}$.
(3) Apply Lemma 3.9 to find paths of length 3 to connect the following set of pairs

$$\bigcup_{i=1}^{t} \left\{ \left(x_i^2, x_i^3\right), \left(x_i^4, u_{i,1}\right), (u_{i,1}, u_{i,2}), \ldots, \left(u_{i,q_{i3}}, x_i^1\right) \right\}$$

with inner vertices from Z'. Note that this is possible as all the vertices of the pairs have good degree to Z' and the number of pairs to connect is $2t + \sum_i q_{i3} = 2t + m'$, which is much less than $m = \gamma n$.
(4) In the previous step we used exactly $2m' + 4t$ vertices of Z' in length 3 paths. Thus the set $Z'' \subseteq Z_1$ of unused vertices has size exactly m. By Fact 4.1 we can find a

C_4-factor on $V(\mathcal{S})\backslash(Z_1\backslash Z'')$ in polynomial time. Note that the paths $a_j y_j c_j b_j$ for each C_4 on $\{y_j, a_j, b_j, c_j\}$ will complete the cycles of length exactly

$$(6q_{i1} + 3q_{i2} + 3) + (l_i - 6q_{i1} - 3q_{i2} - 3q_{i3} - 9) + 3q_{i3} + 6 = l_i$$

for each $i \in [t]$. Thus, we have found a copy of F in G.

Note that we can compute the values of q_{ij} greedily in time $O(n)$. Each of Lemma 3.9, Fact 4.1, and Lemma 4.2 runs in polynomial time and we use them at most twice. Finally, we applied Lemma 3.6 t times and so the overall running time is polynomial. □

Let us mention here that one could also define an absorbing structure specifically for the longer cycles we build in Theorem 2.3, connecting edges into paths according to the adjacencies of a template. Although this alternative structure would be easier to describe and would remove some of the technicalities in the above proof, we chose to instead work from the absorbing structure used for finding factors which involve short cycles, for the sake of brevity.

6 | A PROOF OF THEOREM 2.1

Nenadov's proof is algorithmic, except the proof of [44, Lemma 3.5], in which he used a Hall-type result for hypergraphs due to Haxell. Here we give an alternative proof of this lemma, which moreover provides a polynomial time algorithm.

We first need to recall some definitions from [44]. Let K_4^- be the unique graph with four vertices and five edges. Define an ℓ-chain as a graph obtained by sequentially identifying ℓ copies of K_4^- on vertices of degree 2. Note that an ℓ-chain contains exactly $\ell + 1$ vertices such that the removal of any one of them results in a graph that has a triangle-factor. These vertices are called *removable*.

We say that a triangle in G *traverses* three chains D_1, D_2 and D_3 if it intersects all of them at some removable vertices. Observe that if D_1, D_2 and D_3 are disjoint chains in G and there exists a triangle in G traversing them, then $G[V(D_1) \cup V(D_2) \cup V(D_3)]$ contains a triangle-factor.

Here we state [44, Lemma 3.5] and give an alternative (algorithmic) proof.

Lemma 6.1 (Lemma 3.5 in Nenadov [44]). *Let G be a (p, λ)-bijumbled graph on n vertices with $\lambda \le \varepsilon p^2 n$ for some $\varepsilon \in (0, 1/16]$. Suppose we are given disjoint ℓ-chains $D_1', \ldots, D_t' \subseteq G$ for some $t, \ell \in \mathbb{N}$ such that ℓ is even, $t \ge 2000$ and $400\lambda/p^2 \le t(\ell + 1) \le n/24$. Then for any subset $W \subseteq V(G)\backslash\bigcup_{i\in[t]} V(D_i')$ of size $|W| \ge n/4$ there exist disjoint $(\ell/2)$-chains $D_1, \ldots, D_{2t} \subseteq G[W]$ with the following property: for every $L \subseteq [2t]$ there exists $L' \subseteq [t]$ such that*

$$G\left[\bigcup_{i\in L} V(D_i) \cup \bigcup_{i\in L'} V(D_i{}')\right]$$

contains a triangle-factor, which can be found in polynomial time.

Proof. We set $\varepsilon := 1/16$. Note that a similar calculation as in the proof of Fact 3.3 (i) shows that the number of vertices which have at most $\varepsilon pt\ell$ neighbours in a set of size at least $t(\ell+1)/8 \geq 50\lambda/p^2$ is at most

$$\frac{\lambda^2 t(\ell+1)/8}{(1/8-\varepsilon)^2 p^2 t^2(\ell+1)^2} \leq \frac{\lambda}{3200(1/8-\varepsilon)^2} < \lambda/2. \tag{5}$$

Given ℓ-chains $D_1', \ldots, D_t'$, we partition them arbitrarily into four groups of almost equal sizes, $\mathcal{D}_1, \ldots, \mathcal{D}_4$. Note that for $\mathcal{D}_3$ and $\mathcal{D}_4$, since each of them contains at least $t(\ell+1)/4$ removable vertices, by (5) the number of vertices of G that have degree less than $\varepsilon pt\ell \leq \varepsilon pn/24$ to either of their removable vertices is at most λ. Now we greedily pick $2t(\ell/2)$-chains $D_1, \ldots, D_{2t}$ in W but avoiding these bad vertices by [44, Lemma 3.2]. It remains to verify the 'absorption' property. Fix any subset $L \subseteq [2t]$ of $(\ell/2)$-chains D_i, $i \in L$. We first greedily find triangles traversing $(\ell/2)$-chains (and thus obtain triangle-factors on them) until $t/8$ of them are left. Indeed, this is possible since as long as there are more than $t/8$ of them left, we can greedily partition them into three groups of size roughly $t/24$. Because $(t/24)(\ell/2+1) > t(\ell+1)/48 > 2\lambda/p^2$, it follows from (2) (for a proof see, eg [44, Lemma 2.4]), we find a triangle with one vertex from each group. This triangle traverses the three chains containing it and thus there is a triangle-factor covering these three chains. So we can reduce the number of chains by 3.

We will match the remaining $t/8(\ell/2)$-chains with the ℓ-chains. We start with using ℓ-chains in $\mathcal{D}_1, \mathcal{D}_2$ and recursively find triangles traversing one $(\ell/2)$-chain and one ℓ-chain in $\mathcal{D}_1$, one ℓ-chain in $\mathcal{D}_2$. That is, as long as there exists a vertex v in one of the 'unmatched' chains that sends more than $\varepsilon pt\ell$ edges to the unused removable vertices in both $\mathcal{D}_1$ and $\mathcal{D}_2$, then we pick an edge (whose existence is asserted by (2)) from these neighbourhoods, namely, a triangle containing v. Note that when we stop, the vertices remaining unmatched have degree at most $\varepsilon pt\ell$ to the unused removable vertices of either in $\mathcal{D}_1$ or $\mathcal{D}_2$. Note that there are still roughly half of the chains in $\mathcal{D}_1$ and $\mathcal{D}_2$ left, which contain at least $(\ell+1)\cdot t/8$ removable vertices in both $\mathcal{D}_1$ and $\mathcal{D}_2$. Thus, by (5) there are at most λ vertices that send low degree to either of them, namely, at most $2\lambda/\ell$ $(\ell/2)$-chains are left unmatched. Now we can proceed to match the chains greedily by $\mathcal{D}_3$ and $\mathcal{D}_4$. This is possible because each time we match a chain, we consume $\ell+1$ removable vertices from $\mathcal{D}_3$ and $\mathcal{D}_4$, respectively, and so in total this will consume at most $(\ell+1)(2\lambda/\ell) = 2\lambda(1+1/\ell)$ removable vertices, which is much less than $\varepsilon pt\ell$.

For the running time, note that we used [44, Lemma 3.2] in the proof, but the desired chains can be constructed by depth-first search, which can be done in polynomial time. We also used [44, Lemma 2.4] to claim the existence of a triangle, but we can find this triangle by brute-force searching the neighbourhood of a vertex. Finally, it takes polynomial time to decide which v to use and to find the triangle containing v. □

7 | CONCLUDING REMARKS

In this paper, we answered the question of Nenadov [44] by providing a deterministic polynomial time algorithm, which finds any given 2-factor in a $(p, \varepsilon p^2 n/\log n)$-bijumbled graph on n vertices of minimum degree δpn (for any fixed $\delta > 0$), with $p > 0$ and some absolute parameter $\varepsilon = \varepsilon(\delta) > 0$. This is optimal up to the $O(\log n)$-factor. It also follows from the proof that the strongest condition hinges on the fact that a triangle might be present in a 2-factor (see Theorem 2.1). Indeed, it follows from the proof of Theorem 1.2 that the condition $\lambda \leq \epsilon p^2 n$ would suffice for a 2-factor of girth at least 4 and a solution to Conjecture 1.1 would imply that this condition would guarantee the existence of any 2-factor. The celebrated construction, due to Alon [4], of triangle-free pseudorandom graphs has been extended by Alon and Kahale [12] to graphs without odd cycles of length $2\ell + 1$. They constructed $(n, \Theta(n^{2/(2\ell+1)}), \Theta(n^{1/(2\ell+1)}))$-graphs of odd girth at least $2\ell + 3$. It is proved in [37, Proposition 4.12] that an (n, d, λ)-graph with $\lambda^{2\ell-1} \ll d^{2\ell}/n$ contains a copy of $C_{2\ell+1}$. Since $\lambda = \Omega(\sqrt{d})$ for, say $d \leq n/2$, we have the lower bound on $d = \Omega(n^{2/(2\ell+1)})$. As for even cycles, a theorem of Bondy and Simonovits [14], which does not require any bound on λ, states that $d \gg n^{1/\ell}$ already implies the existence of $C_{2\ell}$. It is thus a natural avenue to further investigate the (almost) optimal conditions of when a (p, λ)-bijumbled graph contains a given 2-factor of girth at least ℓ. When $\ell = n$, the best condition for (n, d, λ)-graphs is provided by the result of Krivelevich and Sudakov [36] which gives $\lambda \leq d(\log\log n)^2/(1000 \log n \log\log\log n)$, while another conjecture of these authors [36] states that $\lambda \leq cd$ should already be sufficient for some absolute $c > 0$. This conjecture would follow from the famous toughness conjecture of Chvátal [16], as shown by Alon [5].

ACKNOWLEDGMENTS

JH was supported by FAPESP (2014/18641-5 and 2013/03447-6). YK was partially supported by FAPESP (2013/03447-6) and CNPq (310974/2013-5, 311412/2018-1, 423833/2018-9). PM is supported by a Leverhulme Trust Study Abroad Studentship (SAS-2017-052\9). YP is supported by the Carl Zeiss Foundation. The cooperation of the authors was supported by a joint CAPES-DAAD PROBRAL project (Project no 430/15, 57350402 and 57391197). FAPESP is the São Paulo Research Foundation. CNPq is the National Council for Scientific and Technological Development of Brazil.

ORCID

Jie Han http://orcid.org/0000-0002-2013-2962

REFERENCES

1. P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person. Blow-up lemmas for sparse graphs. 2016.
2. P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Powers of Hamilton cycles in pseudorandom graphs*, Combinatorica **37** (2017), no. 4, 573–616.
3. P. Allen, J. Böttcher, Y. Kohayakawa, and Y. Person, *Tight Hamilton cycles in random hypergraphs*, Random Struct. Algorithms **46** (2015), no. 3, 446–465.
4. N. Alon, *Explicit Ramsey graphs and orthonormal labelings*, Electron. J. Combin. **1** (1994) R12, 8 pp.
5. N. Alon, *Tough ramsey graphs without short cycles*, J. Algebr. Comb. **4** (1995), no. 3, 189–195.
6. N. Alon, *Universality, tolerance, chaos and order*, An irregular mind. Szemerédi is 70. Dedicated to Endre Szemerédi on the occasion of his seventieth birthday, Springer, Berlin, 2010, pp. 21–37.
7. N. Alon and M. Capalbo, *Sparse universal graphs for bounded-degree graphs*, Random Struct. Algorithms. **31** (2007), no. 2, 123–133.
8. N. Alon, and M. Capalbo, Optimal universal graphs with deterministic embedding, Proceedings of the nineteenth annual ACM-SIAM symposium on discrete algorithms (SODA 2008), San Francisco, CA, January 20–22, 2008, Association for Computing Machinery (ACM); Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), New York, NY, 2008, pp. 373–378.
9. N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi. Universality and tolerance. *Proc 41st IEEE FOCS*, 2000, pp. 14–21.
10. N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi, *Near-optimum universal graphs for graphs with bounded degrees (extended abstract)*, Approximation, randomization, and

combinatorial optimization (Berkeley, CA, 2001), volume 2129 of Lecture Notes in Comput. Sci., Springer, Berlin, 2001, pp. 170–180.

11. N. Alon and F. R. K. Chung, *Explicit construction of linear sized tolerant networks*, Disc. Math. **72** (1988), no. 1–3, 15–19.
12. N. Alon and N. Kahale, *Approximating the independence number via the* ϑ*-function*, Math. Prog. **80** (1998), no. 3, 253–264.
13. N. Alon and J. H. Spencer, The probabilistic method, 4th ed., Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2016.
14. J. A. Bondy and M. Simonovits, *Cycles of even length in graphs*, J. Comb. Theory, Ser. B. **16** (1974), 97–105.
15. F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, Combinatorica **9** (1989), no. 4, 345–362.
16. V. Chvátal, *Tough graphs and Hamiltonian circuits*, Disc. Math. **5** (1973), no. 3, 215–228.
17. D. Conlon, *A sequence of triangle-free pseudorandom graphs*, Combin. Probab. Comput. **26** (2017), no. 2, 195–200.
18. D. Conlon, A. Ferber, R. Nenadov, and N. Škorić, *Almost-spanning universality in random graphs*, Random Struct. Algorithms **50** (2017), no. 3, 380–393.
19. D. Conlon, J. Fox, and Y. Zhao, *Extremal results in sparse pseudorandom graphs*, Adv. Math. **256** (2014), 206–290.
20. D. Dellamonica Jr., Y. Kohayakawa, V. Rödl, and A. Ruciński, *An improved upper bound on the density of universal random graphs*, LATIN 2012: Theoretical informatics, Springer, 2012, pp. 231–242.
21. G. Fan and H. A. Kierstead, *Hamiltonian square-paths*, J. Comb. Theory, Ser. B. **67** (1996), no. 2, 167–182.
22. A. Ferber, G. Kronenberg, and K. Luh, *Optimal threshold for a random graph to be 2-universal*, Trans. Am. Math. Soc. **372** (2019), no. 6, 4239–4262.
23. A. Ferber and R. Nenadov, *Spanning universality in random graphs*, Random Struct. Algorithms **53** (2018), no. 4, 604–637.
24. J. Han, Y. Kohayakawa, P. Morris, and Y. Person, *Clique-factors in sparse pseudorandom graphs*, European J. Comb. **82** (2019), 102999.
25. J. Han, Y. Kohayakawa, and Y. Person. Near-optimal clique-factors in sparse pseudorandom graphs. arXiv:1806.00493, submitted.
26. J. Han, Y. Kohayakawa, and Y. Person, *Near-perfect clique-factors in sparse pseudorandom graphs, Discrete mathematics days 2018*. Extended abstracts of the 11th "Jornadas de matemática discreta y algorítmica" (JMDA), Sevilla, Spain, June 27–29, 2018, Elsevier, Amsterdam, 2018, pp. 221–226.
27. S. Hoory, N. Linial, and A. Widgerson, *Expander graphs and their applications*, Bull. Am. Math. Soc., New Ser. **43** (2006), no. 4, 439–561.
28. R. M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (R. E. Miller, J. W. Thatcher, and J. D. Bohlinger, eds.), Springer, Boston, MA, 1972, pp. 85–103.
29. J. H. Kim and S. J. Lee, *Universality of random graphs for graphs of maximum degree two*, SIAM J. Disc. Math. **28** (2014), no. 3, 1467–1478.
30. D. G. Kirkpatrick and P. Hell, *On the complexity of general graph factor problems*, SIAM J. Comp. **12** (1983), no. 3, 601–609.
31. Y. Kohayakawa, V. e. Rödl, M. Schacht, P. Sissokho, and J. Skokan, *Turán's theorem for pseudo-random graphs*, J. Combin. Theory Ser. A. **114** (2007), no. 4, 631–657.
32. J. Komlós, G. N. Sárközy, and E. Szemerédi, *Blow-up lemma*, Combinatorica **17** (1997), no. 1, 109–123.
33. J. Komlós, G. N. Sárközy, and E. Szemerédi, *An algorithmic version of the blow-up lemma*, Random Struct. Algorithms **12** (1998), no. 3, 297–312.
34. J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi, *The regularity lemma and its applications in graph theory*, Theoretical aspects of computer science. Advanced lectures, Springer, Berlin, 2002, pp. 84–112.
35. S. Kopparty. Cayley graphs (lecture notes), 2011, available at http://sites.math.rutgers.edu/~sk1233/courses/graphtheory-F11/cayley.pdf
36. M. Krivelevich and B. Sudakov, *Sparse pseudo-random graphs are Hamiltonian*, J. Graph Theory. **42** (2003), no. 1, 17–33.
37. M. Krivelevich and B. Sudakov, *Pseudo-random graphs*, More sets, graphs and numbers, volume 15 of Bolyai Soc. Math. Stud., Springer, Berlin, 2006, pp. 199–262.

38. M. Krivelevich, B. Sudakov, and T. Szabó, *Triangle factors in sparse pseudo-random graphs*, Combinatorica **24** (2004), no. 3, 403–426.
39. D. Kühn and D. Osthus, *On Pósa's conjecture for random graphs*, SIAM J. Disc. Math. **26** (2012), no. 3, 1440–1457.
40. M. Kwan. Almost all Steiner triple systems have perfect matchings. 2016.
41. A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.
42. R. Montgomery. Embedding bounded degree spanning trees in random graphs. 2014.
43. R. Montgomery, *Spanning trees in random graphs*, Adv. Math. **356** (2019), 106793.
44. R. Nenadov, *Triangle-factors in pseudorandom graphs*, Bull. London Math. Soc. **51** (2019), no. 3, 421–430.
45. R. Nenadov and Y. Pehova. On a Ramsey-Turán variant of the Hajnal-Szemerédi theorem. arXiv preprint arXiv:1806.03530, 2018.
46. V. Rödl, A. Ruciński, and E.Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combin. Probab. Comput. **15** (2006), no. 1-2, 229–251.
47. V. Rödl, A. Ruciński, and E. Szemerédi, *An approximate Dirac-type theorem for k-uniform hypergraphs*, Combinatorica **28** (2008), no. 2, 229–260.
48. A. Thomason, *Pseudorandom graphs*, Random graphs' 85 (Poznań, 1985), volume 144 of North-Holland Math. Stud., North-Holland, Amsterdam, 1987, pp. 307–331.
49. A. Thomason, *Random graphs, strongly regular graphs and pseudorandom graphs*, Surveys in combinatorics 1987 (New Cross, 1987), volume 123 of London Math. Soc. Lecture Note Ser., Cambridge University Press, Cambridge, 1987, pp. 173–195.
50. A. Walfisz, *Zur additiven Zahlentheorie. II*, Math. Zeitschrift. **40** (1936), no. 1, 592–607.

How to cite this article: Han J, Kohayakawa Y, Morris P, Person Y. Finding any given 2-factor in sparse pseudorandom graphs efficiently. *J Graph Theory*. 2020;1–22. https://doi.org/10.1002/jgt.22576