Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 195 (2021) 404–411

Procedia Computer Science

www.elsevier.com/locate/procedia

XI Latin and American Algorithms, Graphs and Optimization Symposium

Cycle factors in randomly perturbed graphs

Julia Böttcher[a,1], Olaf Parczyk[a,1,2], Amedeo Sgueglia[a,1], Jozef Skokan[a,1]

[a]*Department of Mathematics, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, UK*

Abstract

We study the problem of finding pairwise vertex-disjoint copies of the ℓ-vertex cycle C_ℓ in the randomly perturbed graph model, which is the union of a deterministic n-vertex graph G and the binomial random graph $G(n,p)$. For $\ell \geq 3$ we prove that asymptotically almost surely $G \cup G(n,p)$ contains $\min\{\delta(G), \lfloor n/\ell \rfloor\}$ pairwise vertex-disjoint cycles C_ℓ, provided $p \geq C \log n/n$ for C sufficiently large. Moreover, when $\delta(G) \geq \alpha n$ with $0 < \alpha \leq 1/\ell$ and G and is not 'close' to the complete bipartite graph $K_{\alpha n,(1-\alpha)n}$, then $p \geq C/n$ suffices to get the same conclusion. This provides a stability version of our result. In particular, we conclude that $p \geq C/n$ suffices when $\alpha > n/\ell$ for finding $\lfloor n/\ell \rfloor$ cycles C_ℓ.

Our results are asymptotically optimal. They can be seen as an interpolation between the Johansson–Kahn–Vu Theorem for C_ℓ-factors and the resolution of the El-Zahar Conjecture for C_ℓ-factors by Abbasi.

Keywords: random graph; randomly perturbed graphs; cycle-factors.

1. Introduction and results

Given a graph H, deciding whether a graph F has an H-*factor*, i.e. the union of $\lfloor v(F)/v(H) \rfloor$ pairwise vertex-disjoint copies of H, is computationally hard [17] already when H is a triangle. Consequently, it is valuable to determine natural sufficient conditions on F which guarantee an H-factor. Two natural and prominent conditions of this type concern minimum degree conditions on the one hand, and edge densities in the setting of random graphs on the other hand. In this extended abstract we concentrate on the case that H is the ℓ-vertex cycle C_ℓ with $\ell \geq 2$; in the degenerate case $\ell = 2$ the cycle C_ℓ is a single edge and we obtain a perfect matching. We shall first summarise what is known in these two different settings, and then consider a well-studied combination of both, the so called randomly perturbed graph model.

Let us first consider the case when F is the binomial random graph $G(n,p)$, which is a graph on n vertices in which each edge is chosen independently with probability $p = p(n)$. In this case we are interested in the probability *threshold* $\hat{p} = \hat{p}(n,H)$ such that *asymptotically almost surely* (a.a.s.), that is, with probability tending to one as n

[1] Email: {j.boettcher|o.parczyk|a.sgueglia|j.skokan}@lse.ac.uk.

[2] Supported by the Deutsche Forschungsgemeinschaft (DFG, Grant PA 3513/1-1).

1877-0509
10.1016/j.procs.2021.11.049

tends to infinity, $G(n, p)$ contains an H-factor when $p = \omega(\hat{p})$, and a.a.s. $G(n, p)$ does not contain an H-factor when $p = o(\hat{p})$. For $H = C_\ell$ with $\ell \geq 3$, the celebrated theorem of Johansson, Kahn, and Vu [16] implies that the threshold is $\hat{p}(n, C_\ell) = n^{-(\ell-1)/\ell}(\log n)^{1/\ell}$, where in the case of a perfect matching ($\ell = 2$) the threshold $\log n/n$ has been known since the seminal work of Erdős and Rényi [12].

Turning to minimum degree conditions enforcing the existence of C_ℓ-factors, let G_α be any n-vertex graph of minimum degree at least αn for $0 \leq \alpha \leq 1$. By Dirac's Theorem [9], $\alpha \geq 1/2$ suffices for guaranteeing a perfect matching. Corrádi and Hajnal [8], on the other hand, showed that $\alpha \geq 2/3$ suffices for guaranteeing a C_3-factor. Abbasi [1] generalised this, confirming more generally a conjecture of El-Zahar, showing that any graph G with minimum degree $\delta(G) \geq \frac{n}{\ell} \cdot \lceil \frac{\ell}{2} \rceil$ contains a C_ℓ-factor. Note that this implies that the case of even ℓ and that of odd ℓ behave differently: for even ℓ we need $\delta(G) \geq n/2$, while for odd ℓ we need $\delta(G) \geq \frac{\ell+1}{2\ell} n$. This is not surprising, as in general the optimal minimum degree enforcing an H-factor depends on the chromatic number (or some variant, called the critical chromatic number). See the survey by Kühn and Osthus [20] for more details.

These results provide optimal minimum degree conditions. This can be easily seen by taking the complete bipartite graph with partition classes of sizes $n/2 - 1$ and $n/2 + 1$ for ℓ even, and the complete tripartite graph with classes of sizes $\frac{n}{\ell} - 1$, $\frac{\ell-1}{2\ell} n + 1$ and $\frac{\ell-1}{2\ell} n$ for ℓ odd.

Bohman, Frieze, and Martin [4] combined these two settings, introducing the *randomly perturbed graph* $G_\alpha \cup G(n, p)$, which is obtained by adding to a deterministic graph G_α on n vertices with minimum degree at least αn, a random graph graph $G(n, p)$ on the same vertex set. This model can be motivated as follows. We can alternatively think of $G(n, p)$ as a graph obtained by the following *random process*: start with the empty graph on n vertices and add random edges, one by one. Asking how many random edges we need to add to guarantee a certain property then corresponds to determining the threshold for this property. In fact, as formulated, this process does not generate $G(n, p)$ but rather the uniform random graph $G(n, M)$; but this is easy and standard to fix. It is then natural to modify this process by starting, instead of the empty graph, with some other deterministic n-vertex graph, for example a graph G_α with minimum degree αn. The question then is how this influences the number of random edges needed to enforce the considered property. When $\alpha > 0$ is small then this can be seen as asking how much the threshold for the property is influenced by the existence of low-degree vertices. When p is small then, in analogy to the *smoothed analysis of algorithms* introduced Spielman and Teng [22], this question can be seen as asking how 'atypical' extremal graphs for the property are.

With this in mind, for a fixed $\alpha > 0$ and H, we define the *perturbed threshold* for an H-factor as the $\hat{p} = \hat{p}(n, \alpha, H)$ such that:

(i) when $p = \omega(\hat{p}(n, \alpha, H))$, for any G_α we have $\lim_{n\to\infty} \mathbb{P}(G_\alpha \cup G(n, p)$ contains an H-factor$) = 1$, and

(ii) when $p = o(\hat{p}(n, \alpha, H))$, there exists a G_α such that $\lim_{n\to\infty} \mathbb{P}(G_\alpha \cup G(n, p)$ contains an H-factor$) = 0$.

Balogh, Treglown, and Wagner [3] proved a lower bound on $\hat{p}(n, \alpha, H)$ for any H, which is sharp for all H provided α is small enough. In our setting where $H = C_\ell$, their result states that for any $\ell \geq 3$ and for any $\alpha > 0$, there is a constant $C = C(\alpha, \ell)$ such that $G_\alpha \cup G(n, p)$ with $p \geq Cn^{-(\ell-1)/\ell}$ a.a.s. contains a C_ℓ-factor; their result gives the lower bound of $p \geq C/n$ also in the case $\ell = 2$, which was already proven in [4]. Compared to the threshold $n^{-(\ell-1)/\ell}(\log n)^{1/\ell}$ in $G(n, p)$ alone this saves a log-factor. By taking G_α to be the complete bipartite graph $K_{\alpha n,(1-\alpha)n}$, it is easy to see that this is optimal for $\alpha < 1/\ell$, so $\hat{p}(n, \alpha, C_\ell) = n^{-(\ell-1)/\ell}$ when $0 < \alpha < 1/\ell$.

The problem of determining the perturbed threshold for the remaining range of α (that is, $\alpha \in [1/\ell, 1/2)$ for ℓ even, and $\alpha \in [\frac{1}{\ell}, \frac{\ell+1}{2\ell})$ for ℓ odd) remained open. In fact, this 'intermediate regime' where α is not small but potentially far from the extremal bound in the deterministic setting has so far only infrequently been studied for randomly perturbed graphs. One exception is the work by Han, Morris, and Treglown [14] concerning clique-factors and proving in particular that $\hat{p}(n, \alpha, C_3) = n^{-1}$ for $\alpha \in (1/3, 2/3)$. We recently filled in the remaining gap $\alpha = 1/3$ and proved $\hat{p}(n, 1/3, C_3) = \log n/n$ in [6]. In this extended abstract we generalise these results to larger ℓ and determine the perturbed threshold $\hat{p}(n, \alpha, C_\ell)$ in all open cases for C_ℓ-factors with $\ell \geq 3$.

Theorem 1.1. *For any integer $\ell \geq 3$ and any $\alpha \geq 1/\ell$, there exists $C > 0$ such that for any n-vertex graph G_α with minimum degree $\delta(G_\alpha) = \alpha n$, the randomly perturbed graph $G \cup G(n, p)$ a.a.s. contains a C_ℓ-factor*

(a) *if $\alpha > 1/\ell$ and $p \geq C/n$, and also* (*b*) *if* $\alpha = 1/\ell$ *and* $p \geq C \log n/n$.

The bound on p in (*b*) is asymptotically optimal. To see this, take $p \leq \frac{1}{2}\frac{\log n}{n}$ and $G = G_\alpha$ to be the graph $K_{n/\ell, n-n/\ell}$. Let A and B be its partition classes with $|A| < |B|$, and observe this graph has minimum degree n/ℓ. By an easy first-moment calculation, a.a.s. there is at least a polynomial number of vertices in B that only have neighbours in A. In particular, any cycle containing one of such vertices must contain at least two vertices from A. However, if a C_ℓ-factor exists in $G \cup G(n, p)$, since $|A| = n/\ell$, for each copy of C_ℓ that has at least two vertices in A, there must be at least one copy of C_ℓ fully contained in B, and thus with all edges from $G(n, p)$. Again by an easy first-moment calculation, a.a.s. there are at most $O(\log^\ell n)$ copies of C_ℓ in $G(n, p)$ alone. Therefore a.a.s. a C_ℓ-factor does not exist in $G \cup G(n, p)$ for $p \leq \frac{1}{2}\frac{\log n}{n}$. Together with (*b*) this implies that $\hat{p}(n, 1/\ell, C_\ell) = \log n/n$.

Now we turn to the optimality of (*a*). For even $\ell \geq 4$ and $\alpha \in (1/\ell, 1/2)$, we consider $G = K_{\alpha n, (1-\alpha)n}$. It has minimum degree αn and there can be at most $\frac{\alpha n}{\ell/2} < \frac{n}{\ell}$ copies of C_ℓ using only edges of G and, therefore, we need at least a linear number of random edges. Together with (*a*) this implies that when ℓ is even, $\hat{p}(n, \alpha, C_\ell) = 1/n$ for $\alpha \in (1/\ell, 1/2)$. For odd $\ell \geq 3$, then the same graph shows optimality for $\alpha \in (1/\ell, 1/2)$, as G is bipartite and does not contain any odd cycle and again we need at least a linear number of random edges. For $\alpha \in [\frac{1}{2}, \frac{\ell+1}{2\ell})$, we consider the tripartite complete graph with one class of size $\left(\alpha - \frac{\ell-1}{2\ell}\right)n$ and two classes of sizes $\left(\frac{1}{2} - \frac{\alpha}{2} + \frac{\ell-1}{4\ell}\right)n$. This graph has minimum degree αn and, as above, there are at most $(\alpha - \frac{\ell-1}{2\ell})n < \frac{n}{\ell}$ copies of C_ℓ using only edges of G and we need a linear number of edges from $G(n, p)$. Together with (*a*) this implies that when ℓ is odd, $\hat{p}(n, \alpha, C_\ell) = 1/n$ for $\alpha \in (\frac{1}{2}, \frac{\ell+1}{2\ell})$. Table 1 summarises the resulting perturbed thresholds for cycle factors.

Table 1. The perturbed threshold $\hat{p} = \hat{p}(n, \alpha, C_\ell)$ for C_ℓ-factor in $G_\alpha \cup G(n, p)$, where $\delta(G_\alpha) \geq \alpha n$.

Even ℓ	α	$\alpha = 0$	$0 < \alpha < 1/\ell$	$\alpha = 1/\ell$	$1/\ell < \alpha < 1/2$	$1/2 \leq \alpha$
Odd ℓ	α	$\alpha = 0$	$0 < \alpha < 1/\ell$	$\alpha = 1/\ell$	$1/\ell < \alpha < \frac{\ell+1}{2\ell}$	$\frac{\ell+1}{2\ell} \leq \alpha$
	$\hat{p}$	$n^{-(\ell-1)/\ell}(\log n)^{1/\ell}$	$n^{-(\ell-1)/\ell}$	$n^{-1}\log n$	n^{-1}	0

It is, further, natural to ask how this behaviour changes when instead of a C_ℓ-factor we are interested in covering only a smaller percentage of the vertices with vertex disjoint C_ℓ-copies. To this end, we can prove that we can always find $\delta(G)$ pairwise vertex disjoint copies of C_ℓ in $G \cup G(n, p)$ when $p \geq C \log n/n$.

Theorem 1.2. *For any integer* $\ell \geq 3$, *there exists a* $C > 0$ *such that for any n-vertex graph G we can a.a.s. find* $\min\{\delta(G), \lfloor n/\ell \rfloor\}$ *pairwise disjoint copies of* C_ℓ *in* $G \cup G(n, p)$, *provided that* $p \geq C \log n/n$.

For $\ell = 3$ this is a perturbed version of a result of Dirac [10], that states that an n-vertex graph G with $\frac{1}{2}n \leq \delta(G) \leq \frac{2}{3}n$ contains at least $2\delta(G) - n$ pairwise vertex-disjoint triangles. For $\ell \geq 4$ an approximate version of this result for longer cycles follows from a more general result of Komlós [18]. More precisely, for any $\varepsilon > 0$ and large enough n, he showed that when ℓ is odd (respectively even) there are at least $2\delta(G) - (1 + \varepsilon)n$ (respectively $\frac{2}{\ell}\delta(G) - \varepsilon n$) pairwise vertex-disjoint copies of C_ℓ in any n-vertex graph G with $\frac{1}{2}n \leq \delta(G) \leq \frac{\ell+1}{2\ell}n$ (respectively $\delta(G) \leq \frac{1}{2}n$).

Moreover, we establish a stability version of Theorem 1.2: when the graph G has minimum degree linear in n and G is not 'close' to $K_{m,n-m}$ with $m = \min\{\delta(G), n/\ell\}$, then a.a.s $G \cup G(n, p)$ contains m pairwise vertex-disjoint copies of C_ℓ already at probability $p \geq C/n$. To formalise this we introduce the following notion of stability, where, for numbers a, b, c, we write $a \in b \pm c$ for $b - c \leq a \leq b + c$.

Definition 1.3. *For* $0 < \beta < \alpha < 1/2$ *we say that an n-vertex graph G is* (α,β)-stable *if there exists a partition of* $V(G)$ *into two sets A and B of size* $|A| = (\alpha \pm \beta)n$ *and* $|B| = (1 - \alpha \pm \beta)n$ *such that the minimum degree of the bipartite subgraph* $G[A, B]$ *of G induced by A and B is at least* $\alpha n/4$, *all but* βn *vertices from A have degree at least* $|B| - \beta n$ *into B, all but* βn *vertices from B have degree at least* $|A| - \beta n$ *into A, and* $G[B]$ *contains at most* βn^2 *edges.*

Roughly speaking, the stability condition with $\alpha = 1/\ell$ says that the size of B is roughly $(\ell - 1)$-times the size of A, there is a minimum degree condition between A and B, in each part all but few vertices see most of the other part, and B is almost independent. Moreover, for $0 < \alpha \leq 1/\ell$ and $m = \alpha n$ an integer, $K_{m,n-m}$ is $(\alpha, 0)$-stable. Using regularity method, we can prove the following stability result.

Theorem 1.4 (Stability Theorem)**.** *Fix an integer $\ell \geq 3$. For $0 < \beta < 1/(4\ell)$ there exist $\gamma > 0$ and $C > 0$ such that for any α with $4\beta \leq \alpha \leq 1/\ell$ the following holds. Let G be an n-vertex graph with minimum degree $\delta(G) \geq (\alpha - \gamma)n$ that is not (α,β)-stable. With $p \geq C/n$ a.a.s. the perturbed graph $G \cup G(n, p)$ contains $\min\{\alpha n, \lfloor n/\ell \rfloor\}$ pairwise vertex-disjoint disjoint copies of C_ℓ.*

Note that this immediately implies *(a)* of Theorem 1.1. Indeed, given an integer $\ell \geq 3$ and $\alpha > 1/\ell$, there is a small enough β such that any n-vertex graph G with minimum degree at least αn is not $(1/\ell,\beta)$-stable. We can then apply Theorem 1.4 on input β (and by taking $\alpha = 1/\ell$), which always gives a C_ℓ-factor. Moreover, when we restrict to graphs G that are not (α,β)-stable, then the constant C only depends on ℓ and β, but is independent of α. To deal with (α,β)-stable graphs for a small enough $\beta > 0$, we need the $\log n$-factor and we prove the following.

Theorem 1.5 (Extremal Theorem)**.** *Fix an integer $\ell \geq 3$. For $0 < \alpha_0 \leq 1/\ell$ there exist $\beta, \gamma > 0$ and $C > 0$ such that for any α with $\alpha_0 \leq \alpha \leq 1/\ell$ the following holds. Let G be an n-vertex graph with minimum degree $\delta(G) \geq (\alpha - \gamma)n$ that is (α,β)-stable. With $p \geq C \log n/n$ a.a.s. the perturbed graph $G \cup G(n, p)$ contains $\min\{\delta(G), \lfloor \alpha n \rfloor\}$ pairwise vertex-disjoint copies of C_ℓ.*

Together with Theorem 1.4 this implies *(b)* of Theorem 1.1. When the minimum degree is smaller, we prove the following result.

Theorem 1.6 (Sublinear Theorem)**.** *Fix an integer $\ell \geq 3$. There exists a $C > 0$ such that the following holds for any $1 \leq m \leq \frac{n}{64\ell^2}$ and any n-vertex graph G of minimum degree $\delta(G) \geq m$. With $p \geq C \log n/n$ a.a.s. the perturbed graph $G \cup G(n, p)$ contains m pairwise vertex-disjoint copies of C_ℓ.*

For this result we are not aware of a construction that justifies the $\log n$-term and it would be interesting to know if it can be omitted. Also, we did not optimise the upper bound on m stated in Theorem 1.6, as Theorem 1.4 and 1.5 cover anyway the cases of larger values of m and $p \geq C \log n/n$.

We remark that, in the randomly perturbed graph setting, several variations of the problem we investigated can be considered. For example, given an integer $\ell \geq 3$, $0 < \delta \leq 1/\ell$ and $\alpha \in (0, 1)$, one can ask for the threshold for the property that the randomly perturbed graph $G_\alpha \cup G(n, p)$ contains δn pairwise vertex-disjoint copies of C_ℓ, for any n-vertex graph G_α with minimum degree at least αn. The case $\delta = 1/\ell$ corresponds to cycle-factors and has been the core of our work, so we now focus on the case $0 < \delta < 1/\ell$. It is an easy corollary of our Theorem 1.4 that, given an integer $\ell \geq 3$, $0 < \delta < 1/\ell$, $\varepsilon > 0$, and any n-vertex graph G with minimum degree $\delta(G) \geq \delta n$, we can a.a.s. find $(\delta - \varepsilon)n$ pairwise vertex-disjoint copies of C_ℓ in $G \cup G(n, p)$ provided $p \geq C/n$, where C is a large enough constant depending only on ℓ and ε. In other words we get the following.

Corollary 1.7. *Given any integer $\ell \geq 3$, $\varepsilon > 0$, and $0 < \delta < 1/\ell$, there exists $C > 0$ such that for any n-vertex graph G with minimum degree $\delta(G) \geq (\delta + \varepsilon)n$ we can a.a.s. find δn pairwise disjoint copies of C_ℓ in $G \cup G(n, p)$, provided that $p \geq C/n$.*

This gives a lower bound on the threshold for any $\alpha > \delta$; moreover it is optimal for $\alpha < \frac{\ell\delta}{2}$ when ℓ is even and $\alpha < \frac{1+\delta}{2}$ when ℓ is odd (see the discussion after Theorem 1.2 for the explanations of these bounds). When $\alpha = \delta$, the threshold is $\log n/n$ as discussed in the first part of Theorem 1.1. When $0 \leq \alpha < \delta$, the deterministic graph does not help and the threshold in $G(n, p)$ was determined by Ruciński [21].

Theorem 1.2 follows from Theorem 1.4, 1.5, and 1.6. The proofs of Theorem 1.4 and 1.5 closely follow the corresponding proof for a triangle-factor in [6], once all lemmas are adjusted to the cycle setting. Therefore, we will only sketch their proofs in Section 2, together with a precise statement of each lemma, and we refer the reader to [6] for more details. It is not hard to derive the new lemmas from the corresponding ones in [6], and we skip their proof. Theorem 1.6 requires new ideas, thus we will give more details of the proof in Section 3, while a complete proof can be found in Section 4 of the full version of this extended abstract [7].

2. Sketch of the proofs of the Stability and Extremal Theorems and main lemmas

For simplicity we assume $\alpha = 1/\ell$ and that G is an n-vertex graph with minimum degree $\delta(G) \geq n/\ell$, with n being a multiple of ℓ, in which case both Theorems 1.4 and 1.5, give a C_ℓ-factor in $G \cup G(n, p)$. The proofs for smaller α

follow along the same lines. As ℓ is fixed, throughout all the section when we say *cycle*, this always refers to a cycle of length ℓ. We will use the Szemerédi regularity lemma for Theorem 1.4 and the concepts of regular and super-regular pairs in both proofs. We use the degree form of the regularity lemma in [19], and more details can be found there.

2.1. Extremal Theorem.

Let $0 < \beta \ll \varepsilon \ll d \ll 1$ and $C > 0$ be such that the following holds. Let G be a $(1/\ell, \beta)$-stable graph on n vertices with $\ell | n$ and $\delta(G) \geq n/\ell$. To cover vertices with cycles we will repeatedly use that in any set of size βn there is a path on $\ell - 1$ vertices with edges of $G(n, p)$. This holds a.a.s. with $p \geq C \log n/n$ using a standard application of Janson's inequality.

As G is $(1/\ell, \beta)$-stable, there exists a partition of $V(G)$ into $A \cup B$ where $|A| = (1/\ell \pm \beta)n$ and $|B| = (1 - 1/\ell \pm \beta)n$ such that all conditions in Definition 1.3 are satisfied. We can find a cycle factor in $G \cup G(n, p)$ in three steps. Firstly, we find a collection of disjoint cycles $\mathcal{F}_1$, such that after removing the cycles of $\mathcal{F}_1$, we are left with two sets $A_1 := A \setminus V(\mathcal{F}_1)$ and $B_1 := B \setminus V(\mathcal{F}_1)$ such that $|B_1| = (\ell - 1)|A_1|$. The way we find these cycles depends on the size of A and B. If $|B| \geq (\ell - 1)n/\ell$, then $|B| = (\ell - 1)n/\ell + m$ for some $0 \leq m \leq \beta n$ and we have to find m disjoint cycles entirely within B, just using the minimum degree $\delta(G[B]) \geq n/\ell - |A| = m$ and random edges. This can be done using our Theorem 1.6, and we let $\mathcal{F}_1$ be the family of disjoint cycles we get. Otherwise $|B| < (\ell - 1)n/\ell$ and $|A| = n/\ell + m$ for some $1 \leq m \leq \beta n$, and we let $\mathcal{F}_1$ be any family of $m/(\ell - 2)$ disjoint cycles in $G \cup G(n, p)$ each with $\ell - 1$ vertices in A and one vertex in B. Such a family can be found greedily: indeed during the process, there is always a vertex v in B, not yet contained in a cycle, with at least $d(v, A) - (\ell - 1)m/(\ell - 2) \geq \beta n$ uncovered neighbours in A in the graph G and thus there is a path on $\ell - 1$ vertices in its neighbourhood in $G(n, p)$, that completes to a cycle in $G \cup G(n, p)$. Notice that the minimum degree of $G[A_1, B_1]$ is still linear in n and all but few vertices of A_1 and B_1 have high degree to the other part (in fact they see all but few vertices in the other part).

In the second step we want to cover those vertices in A_1 and B_1 that do not see all but $10\beta n$ vertices from the other side. We will cover them (and some other vertices) with two collections of cycles $\mathcal{F}_2$ and $\mathcal{F}_3$ respectively, where each cycle has one vertex in A_1 and $(\ell - 1)$ vertices in B_1 so that we still have $|B_2| = (\ell - 1)|A_2|$, where $A_2 := A_1 \setminus V(\mathcal{F}_2 \cup \mathcal{F}_3)$ and $B_2 := B_1 \setminus V(\mathcal{F}_2 \cup \mathcal{F}_3)$. That can be done greedily as above, just using the minimum degree condition in $G[A_1, B_1]$ and random edges, because there are only few vertices that do not have high degree. Notice that after this each vertex from A_2 and B_2 sees all but $10\beta n$ vertices from the other side, because we covered all vertices of smaller degree, and these sets are still large, because we only removed few cycles.

Finally we split B_2 arbitrarily into $\ell - 1$ subsets B_2^i of equal size, for $i = 1, \ldots, \ell - 1$, and we remark that $|A_2| = |B_2^1| = \cdots = |B_2^{\ell-1}|$. It is straightforward to check that (A_2, B_2^1) and $(A_2, B_2^{\ell-1})$ are (ε, d)-super-regular pairs, as each vertex has large degree to the other part. Using random edges between B_2^i and B_2^{i+1} for $0 < i < \ell - 1$ and the following Lemma, we can cover $A_2 \cup B_2^1 \cup \cdots \cup B_2^{\ell-1}$ with a cycle factor $\mathcal{F}_4$.

Lemma 2.1. *For any $0 < d < 1$ there exists an $\varepsilon > 0$ and a $C > 0$ such that the following holds. Let $\ell \geq 3$ be an integer and $V, U_1, U_2, \ldots, U_{\ell-1}$ be sets of size n such that (V, U_1) and $(V, U_{\ell-1})$ are (ε, d)-super regular pairs, and for each $0 < i < \ell - 1$ let $G(U_i, U_{i+1}, p)$ be a random bipartite graph with $p \geq C \log n/n$. Then a.a.s. there exists a C_ℓ-factor.*

We conclude by observing that the collection of cycles $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{F}_4$ gives a C_ℓ-factor in $G \cup G(n, p)$.

2.2. Stability Theorem.

Let $0 < \varepsilon \ll \gamma \ll d \ll \beta < 1/(4\ell)$ and $C > 0$. Let G be a graph on n vertices with $\ell | n$ and $\delta(G) \geq (1/\ell - \gamma)n$ that is not $(1/\ell, \beta)$-stable. We apply the regularity lemma to G and obtain the reduced graph R, whose vertices are the clusters and there is an edge between two clusters if they give an (ε, d)-regular pair in G. By adapting ideas from [2] we proved the following stability result in [6].

Lemma 2.2 (Lemma 4.4 in [6])**.** *For any $0 < \beta < \frac{1}{12}$ there exists a $d > 0$ such that the following holds for any $0 < \varepsilon < d/4$, $4\beta \leq \alpha \leq \frac{1}{3}$, and $t \geq \frac{10}{d}$. Let G be an n vertex graph with minimum degree $\delta(G) \geq (\alpha - \frac{1}{2}d)n$ that is not (α, β)-stable and let R be the (ε, d)-reduced graph for some (ε, d)-regular partition $V_0, \ldots, V_t$ of G. Then R contains a matching M of size $(\alpha + 2d)t$.*

Using Lemma 2.2 and that the reduced graph inherits a minimum degree condition from G, we can cover $V(R)$ with pairwise disjoint cherries $K_{1,2}$ and matching edges $K_{1,1}$, such that there are not too many cherries. In the following, by *extended cherry*, we mean a cherry $K_{1,2}$ in R together with $\ell - 3$ additional clusters. Just by rearranging the clusters and distributing the remaining clusters we can cover most of $V(R)$ with pairwise disjoint copies of extended cherries and matching edges $K_{1,1}$, such that we still have many matching edges. The uncovered clusters are added to V_0. We make each edge super-regular (both in cherries and matching edges) by removing some vertices and adding them to V_0, while keeping all clusters of the same size. Then we remove a few more vertices from all but the centre cluster of each extended cherry and add them to V_0, in order to make each centre cluster of an extended cherry bigger than the other clusters. Finally, by moving a few more vertices to V_0, we can assume that for all extended cherries and matching edges in R the number of vertices in the clusters together is divisible by ℓ (and thus $|V_0|$ is divisible by ℓ as well). We can do all this such that V_0 does not get too large.

We start by covering V_0 with a collection of cycles $\mathcal{F}_1$. For this we use that with $p \geq C/n$ we can a.a.s. assume that we can find short paths in $G(n, p)$ with vertices in predefined sets that are not too small and that for any regular pair we can find a cycle using one of its edges and the other vertices within predefined sets that are not too small using $\ell - 1$ edges from $G(n, p)$. Although covering V_0 could be done greedily just using the minimum degree condition and random edges, we do this more carefully in such a way that the total number of vertices in the clusters of each super-regular extended cherry or matching edge remains a multiple of ℓ (to avoid divisibility issues later), and that none of the centre cluster of an extended cherry gets significantly smaller. Notice this can be guaranteed because we cover $V(R)$ without using not too many extended cherries and, therefore, every vertex from V_0 has high degree into the non-centre clusters. By constructing another collection of cycles $\mathcal{F}_2$, we modify just the extended cherries to ensure the sizes of the clusters are as required by the following lemma.

Lemma 2.3. *For any $0 < \delta' \leq d < 1$ there exist $\delta_0, \delta, \varepsilon$ with $\delta' \geq \delta_0 > \delta > \varepsilon > 0$ such that given an integer $\ell \geq 3$, there exists $C = C(\ell)$ such that the following holds. Let $V, U_1, U_2, \dots, U_{\ell-1}$ be sets of size $|V| = n$ and $(1 - \delta_0)n \leq |U_1| = \cdots = |U_{\ell-1}| \leq (1 - \delta)n$, where $|V| + |U_1| + \cdots + |U_{\ell-1}| \equiv 0 \pmod{\ell}$. Further assume (V, U_1) and $(V, U_{\ell-1})$ are (ε, d)-super regular pairs and let $G(V, p)$ and $G(U_i, U_{i+1}, p)$ for each $0 < i < \ell - 1$ be random graphs with $p \geq C/n$. Then a.a.s. there exists a C_ℓ-factor.*

From Lemma 2.3 we also derive the following result about the existence of a cycle factor in a super-regular edge, again with the help of the random edges.

Lemma 2.4. *For any $0 < d < 1$ there exist $\varepsilon > 0$ and $C > 0$ such that the following holds for any integer $\ell \geq 3$ and any sets U, V of size $|V| = n$ and $3n/4 \leq |U| \leq n$, where $|V| + |U| \equiv 0 \pmod{\ell}$. If (U, V) is an (ε, d)-super-regular pair and $G(U, p)$ and $G(V, p)$ are random graphs with $p \geq C/n$, then a.a.s. there exists a C_ℓ-factor.*

We use Lemma 2.3 on each extended cherry and Lemma 2.4 on each matching edge to cover the remaining vertices with a collection of cycles $\mathcal{F}_3$. Together $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ gives a C_ℓ-factor in $G \cup G(n, p)$.

3. Proof of the Sublinear Theorem

Let $\ell \geq 3$ be an integer. Let $1 \leq m \leq \frac{n}{64\ell^2}$ and G be an n-vertex graph with minimum degree $\delta(G) \geq m$. We let $p \geq C \log n/n$, with C large enough such that a.a.s. the applications of the propositions we state later hold and we a.a.s. have the following properties hold in $G(n, p)$:

(i) for any set of vertices U of size $n/(64\ell)$ there is a path on $\ell - 1$ vertices in $G(n, p)[U]$ and

(ii) for a given set of vertices U of size $n/2$ there are at least $\log^\ell n$ pairwise vertex-disjoint C_ℓ's in $G(n, p)[U]$.

Notice that (i) can be guaranteed using the Janson's inequality and the union bound, while (ii) follows from [15, Theorem 3.29]. We want to show that a.a.s. there exist m pairwise vertex-disjoint copies of C_ℓ in $G \cup G(n, p)$.

Any vertex v of large degree in G can easily be covered by a cycle using (i). If there are enough of these vertices, we can already claim m cycles, otherwise we first ignore these vertices and cover them later. For this, let V' be the set of vertices from G of degree at least $\frac{n}{64\lfloor \ell/2 \rfloor}$. If $|V'| \geq m$, then we can greedily find m disjoint cycles in $G \cup G(n, p)$, each containing exactly one vertex from V'. Indeed, as long as we have less than m cycles, there is a vertex $v \in V'$

not yet contained in a cycle. Then there are at least $\frac{n}{64\lfloor \ell/2 \rfloor} - \ell m \geq \frac{n}{32\ell} - \frac{n}{64\ell} \geq \frac{n}{64\ell}$ vertices $U \subseteq N_G(v)$ not covered by cycles and we can find a path on $\ell - 1$ vertices within $G(n, p)[U]$ using property (i), that gives us a cycle on ℓ vertices containing v. Otherwise, $|V'| < m$ and we remove V' from G to obtain $G' = G[V \setminus V']$. Note that we have $v(G') = n - |V'| \geq n/2$, minimum degree $\delta(G') \geq m - |V'| = m'$, and maximum degree $\Delta(G') < \frac{n}{64\lfloor \ell/2 \rfloor} \leq \frac{v(G')}{32\lfloor \ell/2 \rfloor}$.

Now the split the proof in three ranges for the value of m':

$$m' < \log^\ell n, \quad \log^\ell n \leq m' \leq M\sqrt{v(G')}, \quad \text{and} \quad M\sqrt{v(G')} \leq m' \leq \frac{n}{64\ell^2}$$

where M is the constant given by Propositon 3.2 below with input ℓ. If $m' < \log^\ell n$, then we a.a.s. find m' cycles C_ℓ's in $G(n, p)$ using (ii). If $\log^\ell n \leq m' \leq M\sqrt{v(G')}$, then we also have $\log^\ell v(G') \leq \log^\ell n \leq m' \leq M\sqrt{v(G')}$, so we apply the following proposition to G' with $\gamma = 1/(32\lfloor \ell/2 \rfloor)$, and a.a.s. find at least m' vertex-disjoint cycles in $G \cup G(n, p)[V \setminus V']$.

Proposition 3.1. *Let $\ell \geq 3$ be an integer. For any $M \geq 1$ and $0 < \gamma < 1/2$ there exists $C > 0$ such that for any $\log^\ell n \leq m \leq M\sqrt{n}$ and any n-vertex graph G with maximum degree $\Delta(G) \leq \gamma n$ and minimum degree $\delta(G) \geq m$ the following holds. With $p \geq C \log n/n$ there are a.a.s. at least m disjoint C_ℓ's in $G \cup G(n, p)$.*

Finally, if $M\sqrt{v(G')} \leq m' \leq \frac{n}{64\ell^2}$ then we also have $M\sqrt{v(G')} \leq m' \leq \frac{n}{64\ell^2} \leq \frac{v(G')}{16\lceil \ell/2 \rceil}$, and, given the choice of M, we can apply the following proposition to G' and again a.a.s. find at least m' vertex-disjoint cycles in $G \cup G(n, p)[V \setminus V']$.

Proposition 3.2. *Let $\ell \geq 3$ be an integer. There exist $M = M(\ell) \geq 1$ and $C = C(\ell) > 0$ such that for any $M\sqrt{n} \leq m \leq \frac{n}{16\lceil \ell/2 \rceil}$ and any n-vertex graph G with maximum degree $\Delta(G) < \frac{n}{32\lfloor \ell/2 \rfloor}$ and minimum degree $\delta(G) \geq m$ the following holds. With $p \geq C \log n/n$ there are a.a.s. at least m disjoint C_ℓ's in $G \cup G(n, p)$.*

Now, after we found m' disjoint cycles, we can greedily add cycles by using the $m - m'$ vertices from V' and a path in their neighbourhood until we have m cycles. Analogous to above, as long as we have less than m cycles, each available vertex v from V' has at least $\frac{n}{64\ell}$ neighbours not covered by cycles, and with (i) we get a cycle in $G \cup G(n, p)$ containing v. That completes the proof of Theorem 1.6.

The proofs of Propositions 3.1 and 3.2 are available in Section 4 of the full version [7]. However we discuss here the main ideas. For Proposition 3.1 we rely on the following lemma that we proved in our previous work [6, Lemma 7.3] and that allows us to find many large enough pairwise vertex-disjoint stars in G. Before stating it, we need to introduce some notation. With $g_K \geq 2$ an integer, a *star* K is a graph on $g_K + 1$ vertices with one vertex of degree g_K (this vertex is called the *centre*) and the other vertices of degree one (these vertices are called *leaves*).

Lemma 3.3 (Lemma 7.3 in [6])**.** *For every $0 < \gamma < 1/2$ and integer $s > 0$ there exists an $\varepsilon > 0$ such that for n large enough and any m with $2/\varepsilon \leq m \leq \sqrt{n}$ the following holds. In every n-vertex graph G with minimum degree $\delta(G) \geq m$ and maximum degree $\Delta(G) \leq \gamma n$ there exists a family $\mathcal{K}$ of vertex-disjoint stars in G such that every $K \in \mathcal{K}$ has g_K leaves with $\varepsilon m \leq g_K \leq \varepsilon\sqrt{n}$ and*

$$\sum_{K \in \mathcal{K}} g_K^2 \geq s\varepsilon^2 nm.$$

With $\mathcal{K}$ being the family of stars given by Lemma 3.3, we show that a.a.s. at least m stars of $\mathcal{K}$ can be completed to cycles using edges of $G(n, p)$. When $\ell = 3$, it suffices to find one random edge within the set of leaves of a star K, for at least m different stars $K \in \mathcal{K}$ (see the proof of Proposition 7.1 in [6] for details). However, when $\ell > 3$, we cannot find enough paths of length $\ell - 2$ within the sets of leaves and, instead, proceed differently. For each star $K \in \mathcal{K}$, we split the set of its leaves into two sets $A_{K,2}$ and $A_{K,\ell}$ each of size $g_K/2$. Then we find pairwise disjoint sets $A_{K,3}, \ldots, A_{K,\ell-1}$, each of size $g_K/2$, such that for $i = 3, \ldots, \ell - 1$, every vertex from $A_{K,i}$ has at least one neighbour in $A_{K,i-1}$ in the random graph $G(n, p)$. This can be done using expansion properties of $G(n, p)$, and we can also guarantee that the sets are pairwise disjoint for every $K \in \mathcal{K}$. Finally, analogously to the C_3 case, a.a.s. we find an edge between $A_{K,\ell}$ and $A_{K,\ell-1}$ for at least m different stars K. Using the property of the new sets, by working backwards from $A_{K,\ell-1}$ to $A_{K,2}$, and adding two more edges from the star K, we obtain at least m pairwise vertex-disjoint copies of C_ℓ.

On the other hand, for Proposition 3.2 and $m > M\sqrt{n}$, we cannot hope to find many large enough disjoint stars, and we need a different approach. When ℓ is even, the proof is easy and follows from upper bounds on the extremal number of C_ℓ, which is the maximum number of edges in an n-vertex graph that does not contain a copy of C_ℓ. We find at least m cycles greedily in G, as any C_ℓ-free graph contains at most $\frac{3}{4}n^{3/2}$ edges if $\ell = 4$ [11], and at most $O(n^{1+2/\ell})$ edges if $\ell > 4$ [5].

For odd ℓ, we use that $m \geq M\sqrt{n}$ is large enough to find an edge within the neighbourhood of each vertex, already with probability $q = \frac{C\log n}{m^2}$. When $\ell = 3$, we let $s = \lceil \frac{n}{m} \rceil$ and $t = \lceil \frac{m^2}{2n} \rceil$ and we find s cycles C_3 in each of t rounds. More precisely, in each round we find s vertices $v_1, \ldots, v_s$ and pairwise disjoint sets of neighbours $B_1, \ldots, B_s$ each of size $\lceil m/16 \rceil$. Then we simply reveal edges with probability q and get an edge of $G(n, q)$ within each of $B_1, \ldots, B_s$. As $tq \leq C\frac{\log n}{n}$ we can repeat this for t rounds and find $ts \geq m$ pairwise vertex-disjoint cycles C_3 (see the proof of Proposition 7.2 in [6] for details).

When $\ell > 3$, we still find s cycles C_ℓ in t rounds, but this time, in each round and for each $i = 1, \ldots, s$, we do the following. Instead of the vertex v_i, we construct a path in G on $\ell - 2$ vertices $v_{i,2}, v_{i,3}, \ldots, v_{i,\ell-1}$ and two sets $B_{i,2}$ and $B_{i,\ell-1}$ of at least $\lceil m/32 \rceil$ neighbours of $v_{i,2}$ and $v_{i,\ell-1}$, respectively. We find this path using the dependent random choice technique, which is a powerful tool that, for example, gives upper bounds on extremal numbers of bipartite graphs (see the survey [13]). After having done that for each $i = 1, \ldots, s$, we find an edge of $G(n, q)$ between $B_{i,2}$ and $B_{i,\ell-1}$, that gives a cycle C_ℓ. As in the case $\ell = 3$, we can perform t rounds and find $ts \geq m$ pairwise vertex-disjoint cycles C_ℓ.

References

[1] Abbasi, S., "The solution of the El-Zahar problem," Ph.D. thesis, Rutgers University (1998).

[2] Balogh, J., F. Mousset and J. Skokan, *Stability for vertex cycle covers*, Electron. J. Combin. **24** (2017), pp. 25, P3.56.

[3] Balogh, J., A. Treglown and A. Z. Wagner, *Tilings in randomly perturbed dense graphs*, Combin. Probab. Comput. **28** (2019), pp. 159–176.

[4] Bohman, T., A. Frieze and R. Martin, *How many random edges make a dense graph hamiltonian?*, Random Structures Algorithms **22** (2003), pp. 33–42.

[5] Bondy, J. A. and M. Simonovits, *Cycles of even length in graphs*, J. Combinatorial Theory Ser. B **16** (1974), pp. 97–105.

[6] Böttcher, J., O. Parczyk, A. Sgueglia and J. Skokan, *Triangles in randomly perturbed graphs* (2020), arXiv:2011.07612. URL https://arxiv.org/abs/2011.07612

[7] Böttcher, J., O. Parczyk, A. Sgueglia and J. Skokan, *Cycle factors in randomly perturbed graphs* (2021), arXiv2103.06136. URL https://arxiv.org/abs/2103.06136

[8] Corrádi, K. and A. Hajnal, *On the maximal number of independent circuits in a graph*, Acta Math. Acad. Sci. Hungar. **14** (1963), pp. 423–439.

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

[10] Dirac, G. A., *On the maximal number of independent triangles in graphs*, Abh. Math. Sem. Univ. Hamburg **26** (1963), pp. 78–82.

[11] Erdős, P., A. Rényi and V. T. Sós, *On a problem of graph theory*, Studia Sci. Math. Hungar. **1** (1966), pp. 215–235.

[12] Erdős, P. and A. Rényi, *On the existence of a factor of degree one of a connected random graph*, Acta Math. Acad. Sci. Hungar. **17** (1966), pp. 359–368.

[13] Fox, J. and B. Sudakov, *Dependent random choice*, Random Structures Algorithms **38** (2011), pp. 68–99.

[14] Han, J., P. Morris and A. Treglown, *Tilings in randomly perturbed graphs: bridging the gap between Hajnal-Szemerédi and Johansson-Kahn-Vu*, Random Structures Algorithms **58** (2021), pp. 480–516.

[15] Janson, S., T. Łuczak and A. Ruciński, "Random graphs," Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000, xii+333 pp.

[16] Johansson, A., J. Kahn and V. Vu, *Factors in random graphs*, Random Structures Algorithms **33** (2008), pp. 1–28.

[17] Kirkpatrick, D. G. and P. Hell, *On the completeness of a generalized matching problem*, in: *Conference Record of the Tenth Annual ACM Symposium on Theory of Computing (San Diego, Calif., 1978)*, ACM, New York, 1978 pp. 240–245.

[18] Komlós, J., *Tiling Turán theorems*, Combinatorica **20** (2000), pp. 203–218.

[19] Komlós, J. and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, in: *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, János Bolyai Math. Soc., Budapest, 1996 pp. 295–352.

[20] Kühn, D. and D. Osthus, *Embedding large subgraphs into dense graphs*, in: *Surveys in Combinatorics*, London Math. Soc. Lecture Note Ser. **365**, Cambridge Univ. Press, 2009 pp. 137–167.

[21] Ruciński, A., *Matching and covering the vertices of a random graph by copies of a given graph*, Discrete Math. **105** (1992), pp. 185–197.

[22] Spielman, D. A. and S.-H. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, J. ACM **51** (2004), pp. 385–463.