Packing and counting arbitrary Hamilton cycles in random digraphs

Ferber, Asaf; Long, Eoin

DOI:
10.1002/rsa.20796

Packing and counting arbitrary Hamilton cycles in random digraphs

Asaf Ferber* Eoin Long†

February 22, 2018

Abstract

We prove packing and counting theorems for arbitrarily oriented Hamilton cycles in $\mathcal{D}(n,p)$ for nearly optimal p (up to a $\log^c n$ factor). In particular, we show that given $t = (1-o(1))np$ Hamilton cycles $C_1, \ldots, C_t$, each of which is oriented arbitrarily, a digraph $D \sim \mathcal{D}(n,p)$ w.h.p. contains edge disjoint copies of $C_1, \ldots, C_t$, provided $p = \omega(\log^3 n/n)$. We also show that given an arbitrarily oriented n-vertex cycle C, a random digraph $D \sim \mathcal{D}(n,p)$ w.h.p. contains $(1 \pm o(1))n!p^n$ copies of C, provided $p \geq \log^{1+o(1)} n/n$.

1 Introduction

A *Hamilton cycle* in a graph is a cycle passing through every vertex of the graph exactly once, and a graph is *Hamiltonian* if it contains a Hamilton cycle. For digraphs, a Hamilton cycle is a cycle passing through every vertex of the graph exactly once, with edges oriented cyclically. Hamiltonicity is one of the most central notions in graph theory, and has been intensively studied by numerous researchers in recent decades.

One of the first, and probably the most celebrated, sufficient conditions for Hamiltonicity in graphs was established by Dirac [6] in 1952. He proved that every graph on n vertices, $n \geq 3$, with minimum degree at least $n/2$ is Hamiltonian. Ghouila-Houri [13] proved an analogue of Dirac's theorem for digraphs, showing that any digraph of minimum semi-degree at least $n/2$ contains an oriented Hamilton cycle (the semi-degree of a digraph G, denoted $\delta^0(G)$, is the minimum of all the in- and out-degrees of vertices of G).

Instead of studying "consistently oriented" Hamilton cycles in digraphs, it is natural to consider Hamilton cycles with arbitrary orientations. This problem goes back to the 80s where Thomason [23] showed that each regular tournament contains every orientation of a Hamilton cycle. Later on, Häggkvist and Thomason [15] showed an approximate analog of the result of Ghouila-Houri [13] while proving that $\delta^0(G) \geq n/2 + n^{5/6}$ is sufficient to guarantee every orientation of a Hamilton cycle appears in G. Very recently, this problem has been settled completely by DeBiasio, Kühn, Molla, Osthus and Taylor [4]. They showed that $\delta^0(G) \geq n/2$ is enough for all cases other than an *anti-directed* Hamilton cycle, where for the latter, Debiaso and Molla showed in [5] that $\delta^0(G) \geq n/2 + 1$ is enough (an anti-directed Hamilton cycle is a cycle with no two consecutive edges having the same orientation).

*Department of Mathematics, MIT, USA. Email: ferbera@mit.edu. Research supported in part by an NSF grant.
†Mathematical Institute, University of Oxford, Oxford, UK. Email: long@maths.ox.ac.uk.
Supported in part by ERC Starter Grant 633509.

In this paper we restrict our attention to the sparse setting, that is, to *random directed graphs*. Let $\mathcal{D}(n,p)$ be the probability space consisting of all directed graphs on vertex set $[n]$ in which each possible arc is added with probability p independently at random.

One of the first results regarding Hamilton cycles in random directed graphs was obtained by McDiarmid in [21]. He showed (among other things) by using an elegant coupling argument that

$$\Pr\left(G \sim \mathcal{G}(n,p) \text{ is Hamiltonian}\right) \leq \Pr\left(D \sim \mathcal{D}(n,p) \text{ is Hamiltonian}\right).$$

Combined with the result of Bollobás [2] it follows that a typical $D \sim \mathcal{D}(n,p)$ is Hamiltonian for $p \geq \frac{\ln n + \ln\ln n + \omega(1)}{n}$. Frieze [11] later proved that $p = \frac{\ln n + \omega(1)}{n}$ is the correct threshold for the appearance of a Hamilton cycle in $D \sim \mathcal{D}(n,p)$.

While Frieze's result gives a better bound than McDiarmid's coupling argument, the former is much more flexible (for some further applications, see [7]). For example, given an arbitrary oriented Hamilton cycle C, it follows immediately from McDiarmid's proof that

$$\Pr\left(G \sim \mathcal{G}(n,p) \text{ is Hamiltonian}\right) \leq \Pr\left(D \sim \mathcal{D}(n,p) \text{ contains a copy of } C\right).$$

In contrast, the result of Frieze is tailored to "consistently oriented" Hamilton cycles and gives no improvement on the obtained bound of $p = \frac{\ln n + \ln\ln n + \omega(1)}{n}$ for general orientations. It may be interesting to find the exact threshold for the appearance of an arbitrary oriented Hamilton cycle and we conjecture the following:

Conjecture 1.1. *Let C be a Hamilton cycle oriented arbitrarily, then a digraph $D \sim \mathcal{D}(n,p)$ w.h.p. contains a copy of C, provided that $p = \frac{\ln n + \omega(1)}{n}$.*

Another recent result worth mentioning was given by Ferber, Nenadov, Peter, Noever and Škoric in [9]. Here it was proven using the "absorption method" that $D \sim \mathcal{D}(n,p)$ is w.h.p. Hamiltonian even if an adversary deletes roughly one half of the in- and out-degrees of all the vertices, provided that $p \geq \log^C(n)/n$ for some constant $C > 0$.

Here we deal with the problems of counting and packing arbitrary oriented Hamilton cycles in $D \sim \mathcal{D}(n,p)$, for edge-densities $p \geq \log^C(n)/n$. The analogous problems regarding the "consistently oriented" Hamilton cycles has been recently treated by Kronenberg and the authors in [8]. However, the proof method there is inapplicable to the arbitrary oriented case.

Enhancing a recent "online sprinkling" technique introduced by Ferber and Vu [10], we manage to tackle these two problems. Our first theorem gives an asymptotically optimal result for packing arbitrarily oriented Hamilton cycles in $D \sim \mathcal{D}(n,p)$.

Theorem 1.2. *Let $\epsilon > 0$ and $p(n) \in (0,1]$. Let $t = (1-\epsilon)np$ and suppose that $C_1, \ldots, C_t$ are n-vertex cycles with arbitrary orientations. Then w.h.p. $D \sim \mathcal{D}(n,p)$ contains edge disjoint copies of $C_1, \ldots, C_t$, provided $p \gg \log^3 n/n$.*

Our second result shows that given an arbitrarily oriented Hamilton cycle C, w.h.p. $D \sim \mathcal{D}(n,p)$ contains the "correct" number of copies of C.

Theorem 1.3. *Suppose that C is an arbitrarily oriented n-vertex cycle. Then w.h.p. a digraph $D \sim \mathcal{D}(n,p)$ contains $(1 \pm o(1))^n n! p^n$ distinct copies of C, provided $p \gg (\log\log n)\log n/n$.*

Before closing the introduction, let us mention that packing and counting Hamilton cycles in the undirected setting has been extensively studied by numerous researchers. In fact, both of these problems are now completely resolved (see [14, 18, 19, 20] and their references). In particular, as conjectured by Frieze and Krivelevich in [12], it is now known that for all p, a typical $G \sim G(n,p)$ contains $\lfloor \delta(G)/2 \rfloor$ edge-disjoint Hamilton cycles, which is clearly best possible (for a summary of all previous work we refer the reader to [18]). Note that in this paper we only find $(1-\varepsilon)\delta^0(D(n,p))$ edge-disjoint, arbitrary oriented Hamilton cycles, and only for $p \geq \log^C /n$. Therefore, it would be very interesting to obtain analogous statement for the directed setting, even if only for the 'consistently oriented' Hamilton cycle.

The main difficulty when working in the directed case, is that the so called Posá rotation-extension technique (see [22]) does not work in its simplest form and therefore one should find more creative ways for generating Hamilton cycles. This will be discussed in more details in later sections.

Notation: Given a directed graph (*digraph*) D, we write $V(D)$ for the vertex set of D and $E(D)$ for the edge set D. Given $v \in V(D)$ we write $N^+(v) = \{u \in V(D) : \overrightarrow{vu} \in E(D)\}$, the *out-neighbourhood* of v, and let $d^+(v) = |N^+(v)|$, the *outdegree* of v in D. Similarly define $N^-(v)$ and $d^-(v)$. Let $\delta^0(D)$ denote the *semi-degree* of D, given by $\delta^0(D) = \min_{v\in V(D), *\in\{+,-\}} d^*(v)$. Given $n \in \mathbb{N}$, let D_n denote the *complete directed graph* (or *complete digraph*) on n vertices, consisting of all possible $n(n-1)$ directed edges.

A *path* P of length k is a $(k+1)$-vertex digraph with k edges, given by $P := v_0v_1\dots v_k$ where for each $i \in [0, k-1]$ either $\overrightarrow{v_iv_{i+1}}$ or $\overleftarrow{v_iv_{i+1}}$ is an edge of P. Given $\sigma : [0, k-1] \to \{+,-\}$, we say that P is a *σ-path*, if for all $i \in [0, k-1]$ the edge $\overrightarrow{v_iv_{i+1}}$ lies in P whenever $\sigma(i) = +$, and $\overleftarrow{v_iv_{i+1}}$ lies in P whenever $\sigma(i) = -$. In this case we write $\sigma(P) = \sigma$. In a similar way, for $\sigma : [0, k-1] \to \{+,-\}$ a *σ-cycle* $C := v_1 \dots v_k v_1$ is a k-vertex digraph with k edges, each of the form $\overrightarrow{v_iv_{i+1}}$ or $\overleftarrow{v_iv_{i+1}}$, where each appears according to the sign of $\sigma(i)$. Given a cycle C and a subpath P of C, let P^c denote the path induced by the edges of C which do not lie in P, called the *complement* of P in C.

Given a digraph D, we write $\mathcal{D}(D,p)$ for the probability space of random subdigraphs of D obtained by including each edge of D independently with probability p. For a graph G, we write $\mathcal{G}(G,p)$ for the analogous distribution on subgraphs of G. In the special case when $D = D_n$ we simply write $\mathcal{D}(n,p)$. Similarly for graphs we write $\mathcal{G}(n,p)$. Given a sequence of n-vertex digraphs $\{D_n\}$ or n-vertex graphs $\{G_n\}$ we will say that an event holds *with high probability* (w.h.p.) for $\mathcal{D}(D_n,p)$ or $\mathcal{G}(G_n,p)$ if the event holds with probability at least $1-\epsilon(n)$, where $\epsilon(n)$ is some function tending to 0 with n. Occasionally this will be abbreviated to say $\mathcal{D}(D,p)$ holds with high probability for an n-vertex digraph D when the sequence is implicit. Similarly we say *with very high probability* (w.v.h.p.) to mean with probability $1 - n^{-\omega(1)}$.

2 Tools

2.1 Chernoff's inequalities

Throughout the paper we will make extensive use of the following well-known bound on the upper and lower tails of the Binomial distribution, due to Chernoff (see for example Appendix A in [1]).

Lemma 2.1 (Chernoff's inequality). *Let $X \sim Bin(n,p)$ and let $\mathbb{E}(X) = \mu$. Then*

- $\mathbb{P}(X < (1-a)\mu) < e^{-a^2\mu/2}$ *for every $a > 0$;*
- $\mathbb{P}(X > (1+a)\mu) < e^{-a^2\mu/3}$ *for every $0 < a < 3/2$.*

We also make use of the following simple lemma.

Lemma 2.2. *Let $X \sim Bin(n, p)$. Then, for every k we have*

$$\Pr(X \geq k) \leq \left(\frac{enp}{k}\right)^k.$$

Proof. Clearly,

$$\Pr(X \geq k) \leq \binom{n}{k} p^k \leq \left(\frac{enp}{k}\right)^k$$

as desired. □

2.2 A concentration inequality

A filtration $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_N$ of a measurable space Ω is an increasing sequence of σ-algebras of Ω. A sequence of random variables $0 \equiv X_0, X_1, \ldots, X_N$ is said to be a submartingale with respect to the filtration $\{\mathcal{F}_i\}_{i \in [N]}$ if each X_i is $\mathcal{F}_i$-measurable and

$$\mathbb{E}(X_i | \mathcal{F}_{i-1}) \leq X_{i-1} \text{ for all } i \in [N].$$

The next result gives a concentration bound for submartingales (see Theorem 7.3 in the survey of Chung and Lu [3], taking $\phi_i = a_i = 0$).

Theorem 2.3. *Suppose $0 \equiv X_0, X_1, \ldots, X_N$ is a submartingale with respect to the filtration $\{\mathcal{F}_i\}_{i \in [N]}$ and satisfies*

$$\mathbb{V}ar(X_i | \mathcal{F}_{i-1}) \leq \sigma; \qquad \textit{and} \qquad X_i - \mathbb{E}(X_i | \mathcal{F}_{i-1}) \leq M.$$

Then $\mathbb{P}(X_N \geq m) \leq e^{-m^2/2(N\sigma + Mm/3)}$.

We will make use of the following simple corollary.

Corollary 2.4. *Suppose that $A_1, \ldots, A_N$ are a sequence of events in a probability space $(\Omega, \mathbb{P})$. Suppose that for all $i \in [N]$ we have $\mathbb{P}(A_i | I_1, \ldots, I_{i-1}) \leq q$, where I_j is the indicator random variable for the event A_j. Then letting E_m denote the event that at least $qN + m$ of the events $A_1, \ldots, A_N$ occur, we have $\mathbb{P}(E_m) \leq e^{-m^2/2(Nq + m/3)}$.*

Proof. Let $\mathcal{F}_i$ be the σ-algebra generated by the sets $\{A_1, \ldots, A_i\}$ for each $i \in [N]$, so that $\{\mathcal{F}_i\}_{i \in [N]}$ is a filtration. Set $X_i := \sum_{j \leq i} \big(I_i - q\big)$ for all $i \in [N]$, and $X_0 \equiv 0$. Clearly X_i is $\mathcal{F}_i$-measurable for all $i \in [N]$ and

$$\mathbb{E}(X_i | \mathcal{F}_{i-1}) = \mathbb{E}(I_i | \mathcal{F}_{i-1}) - q + X_{i-1} \leq X_{i-1} \text{ for all } i \in [N],$$

showing that $X_0, \ldots, X_N$ is a submartingale with respect to $\{\mathcal{F}_i\}_{i \in [N]}$. We also have

$$\mathbb{V}ar(X_i | \mathcal{F}_{i-1}) = \mathbb{V}ar(I_i | \mathcal{F}_{i-1}) \leq q \text{ and } X_i - \mathbb{E}(X_i | \mathcal{F}_{i-1}) = I_i - \mathbb{E}(I_i | \mathcal{F}_{i-1}) \leq 1 \text{ for all } i \in [N].$$

Taking $M = 1$ and $\sigma = q$, Theorem 2.3 gives

$$\mathbb{P}\big(E_m\big) = \mathbb{P}(X_N > m) \leq e^{-m^2/2(Nq + m/3)},$$

as required. □

2.3 Completing paths into a Hamilton cycle

The following lemma is the main result of this subsection and will be used to complete paths into Hamilton cycles in $D \sim \mathcal{D}(n,p)$.

Lemma 2.5. *Let $\sigma \in \{+,-\}^{n-1}$. Suppose that G is an n-vertex digraph with $\delta^0(G) \geq \left(1 - \frac{1}{2\log n}\right) n$, with $n \geq n_0$. Then with probability $1 - n^{-\omega(1)}$ a digraph $D \sim \mathcal{D}(G,p)$ contains a σ-path Q between* any *two distinct vertices in D, provided $p = \omega(\log n/n)$.*

In order to prove Lemma 2.5 we make use of a result due to Hefetz, Krivelevich and Szábo [16] and the coupling idea of McDiarmid [21]. Given a graph G on n vertices, let us consider the following two properties (obtained from the ones in [16] by choosing $d = \log^{0.1} n$):

$(P1)$ For every $S \subset V(G)$ with $|S| \leq \frac{n}{\log n}$ we have $|N(S) \setminus S| \geq |S| \log^{0.1} n$;

$(P2)$ There is an edge between any two disjoint subsets $A, B \subseteq V(G)$ such that $|A|, |B| \geq \frac{n \log\log n}{\log n}$.

The following theorem is proven in [16].

Theorem 2.6. *Every sufficiently large graph G satisfying $(P1)$ and $(P2)$ is Hamiltonian connected. That is, for every $u, v \in V(G)$, there is a Hamiltonian path in G with u, v as its endpoints.*

Using Theorem 2.6 we now prove that given a graph G with high minimum degree, a random subgraph of it is Hamiltonian connected with very high probability.

Lemma 2.7. *Let G be a graph on $n \geq n_0$ vertices with $\delta(G) \geq n - n/\log n$. Then, with probability $1 - n^{-\omega(1)}$ a graph $H \sim \mathcal{G}(G,p)$ is Hamiltonian connected, provided that $p = \omega(\log n/n)$.*

Proof. By Theorem 2.6 it is enough to show that w.v.h.p. H satisfies both $(P1)$ and $(P2)$. Let us start with $(P1)$. Note that every vertex $v \in V(G)$ has $\mathbb{E}[d_H(v)] = (1 - o(1))np$. Therefore, by Lemma 2.1 and the union bound we obtain that the probability that there exists a vertex of degree not in $(1 \pm \frac{1}{2})np$ is at most $ne^{-\Theta(np)} = n^{-\omega(1)}$. Thus w.v.h.p. we find that for all $v \in V(G)$ we have $d_H(v) \geq np/2 \gg \log n$, and in particular $(P1)$ holds for all $|S| \leq \log^{0.9} n$.

We now bound the probability that $(P1)$ fails for some set $S \subseteq V(H)$ of size $s \in [\log^{0.9} n, n/\log n]$, i.e. that $|N_H(S)| < s \log^{0.1} n$. If this is the case, $T = S \cup N(S)$ is a subset with $t = |T| \leq 2s \log^{0.1} n$ containing at least $(\sum_{v \in S} d_H(v))/2 \geq snp/3 \geq tnp/6 \log^{0.1} n$ edges. However, any $X \subseteq V(H)$ satisfies $\mathbb{E}[e_H(X)] \leq |X|^2 p/2$. By Lemma 2.2 and the union bound we obtain that the probability for having such a set of size at most $2(n/\log n)\log^{0.1} n \leq n/\log^{0.8} n$ is at most

$$\begin{aligned}
\sum_{2\log n \leq t \leq n/\log^{0.8} n} \binom{n}{t} \left(\frac{et^2p/2}{tnp/6\log^{0.1} n}\right)^{snp/3} &\leq \sum_{2\log n \leq t \leq n/\log^{0.8} n} \left(\frac{en}{t}\right)^t \left(\frac{10t\log^{0.1}(n)}{n}\right)^{s\log n} \\
&\leq \sum_{2\log n \leq t \leq n/\log^{0.8} n} \left(30\log^{0.1}(n)\right)^{s\log n} (t/n)^{s\log n - t} \\
&\leq n(\log^{0.2} n)^{s\log n}(\log^{-0.8} n)^{s\log n/2} = n^{-\omega(1)}.
\end{aligned}$$

To prove $(P2)$, it is enough to show that the probability for having two subsets A, B of size exactly $n\log\log n/\log n$ with $e(A,B) = 0$ is $n^{-\omega(1)}$. Indeed, this probability is upper bounded by

$$\binom{n}{n\log\log n/\log n}^2 (1-p)^{(1-o(1))\left(\frac{n\log\log n}{\log n}\right)^2} \leq \left(\frac{e\log n}{\log\log n}\right)^{\frac{2n\log\log n}{\log n}} e^{-(1-o(1))\left(\frac{n\log\log n}{\log n}\right)^2 p} = n^{-\omega(1)}.$$

This completes the proof of the lemma. $\square$

We are now ready to prove Lemma 2.5, using a beautiful coupling idea of McDiarmid [21].

Proof of Lemma 2.5. Let G be a digraph with $\delta^0(G) \geq n - n/2\log n$. Delete edges of G which do not also appear with the opposite orientation in G, i.e. delete $\overrightarrow{uv} \in E(G)$ if $\overrightarrow{vu} \notin E(G)$. Abusing notation, let G denote the resulting digraph and note that $\delta^0(G) \geq n - n/\log n$. We also let G' denote the underlying graph of G, with $uv \in E(G')$ if and only if $\overrightarrow{uv}$ and $\overrightarrow{vu} \in E(G)$. Also let $t = |E(G')|$.

To prove the lemma let us fix an arbitrary ordering of the (undirected) edges of $E(G')$, say $e_1, \dots, e_t$, where $e_j = \{u_j, v_j\}$ for all $j \in [t]$. For each $i \in [0,t]$, consider the following random process to generate a subdigraph Γ_i of G. Toss $t+i$ independent Bernoulli coins, $C_1^{e_1}, C_2^{e_1}, \dots, C_1^{e_i}, C_2^{e_i}$ and $D^{e_{i+1}}, \dots, D^{e_t}$, each of which appears as heads with probability p. Then construct Γ_i according to the following rule

- For $j \in [i]$, adjoin $\overrightarrow{u_j v_j}$ to Γ_i if and only if $C_1^{e_j}$ appears as heads;
- For $j \in [i]$, adjoin $\overrightarrow{v_j u_j}$ to Γ_i if and only if $C_2^{e_j}$ appears as heads;
- For $j \in [i+1,t]$ adjoin *both* $\overrightarrow{u_j v_j}$ and $\overrightarrow{v_j u_j}$ to Γ_i if and only if D^{e_j} appears as heads.

Let $\mathcal{D}_i$ denote the resulting distribution on the subdigraphs of G.

Now let us fix two distinct vertices $u, v \in V(G)$ and $\sigma \in \{+,-\}^{n-1}$. Given a random subdigraph Γ of G let $E(\Gamma, u, v)$ denote the event that 'Γ contains a σ-path starting at u and ending at v'. The key inequality in our proof is that, for all $i \in [t]$,

$$\Pr_{\Gamma_i \sim \mathcal{D}_i}\left(E(\Gamma_i, u, v)\right) \geq \Pr_{\Gamma_{i-1} \sim \mathcal{D}_{i-1}}\left(E(\Gamma_{i-1}, u, v)\right). \tag{1}$$

In particular, by telescoping these inequalities, this gives

$$\Pr_{D \sim \mathcal{D}(G,p)}\left(E(D, u, v)\right) = \Pr_{\Gamma_t \sim \mathcal{D}_t}\left(E(\Gamma_t, u, v)\right) \geq \Pr_{\Gamma_0 \sim \mathcal{D}_0}\left(E(\Gamma_0, u, v)\right). \tag{2}$$

The equality here holds since $\mathcal{D}_t$ is exactly the distribution $\mathcal{D}(G,p)$. However, to generate a random digraph Γ according to $\mathcal{D}_0$ we simply select a random (undirected) subgraph H of G' and let Γ consist of all the directed edges corresponding to edges in H. In particular, $E(\Gamma, u, v)$ occurs if and only if H has a Hamilton path from u to v. As the minimum degree of G' is at least $n - n/\log n$ and $p = \omega(\log n/n)$, by Lemma 2.7 we have $\Pr_{\Gamma_0 \sim \mathcal{D}_0}\left(E(\Gamma_0, u, v)\right) = 1 - n^{-\omega(1)}$. Combined with (2), by the union bound, this gives $\Pr_{D \sim \mathcal{D}(G,p)}(\cap_{u,v} E(D, u, v)) \geq 1 - n^2 n^{-\omega(1)} = 1 - n^{-\omega(1)}$, as desired.

It remains to prove (1). To see this, note that we can couple random digraphs generated via $\mathcal{D}_{i-1}$ and $\mathcal{D}_i$. Indeed, we can use the same coins to generate the directed edges corresponding to $e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_t$ in both Γ_{i-1} and Γ_i and call the resulting subdigraph $\widetilde{\Gamma}$. Then use D^{e_i} to finish generating Γ_{i-1} and $C_1^{e_i}$ and $C_2^{e_i}$ to finish generating Γ_i. After exposing $\widetilde{\Gamma}$, there are three scenarios:

(*a*) $\widetilde{\Gamma}$ contains a σ-path from u to v not involving e_i, or

(*b*) $\widetilde{\Gamma}$ does not contain such a path even if we add both directions of e_i, or

(c) if we add some (possibly either) orientation of e_i to $\widetilde{\Gamma}$ it contains a σ-path from u to v, but does not otherwise.

Note that in (a) and (b) there is nothing to prove, as Γ_i contains the required σ-path if and only if Γ_{i-1} does. In case (c) D^{e_i} must appear as heads for Γ_{i-1} to have the desired path, which occurs with probability p (conditional on $\widetilde{\Gamma}$). However, it is clear that Γ_i contains a σ-path from u to v with at least this probability (perhaps more if both $\overrightarrow{u_iv_i}$ and $\overrightarrow{v_iu_i}$ guarantee a σ-path). $\square$

3 Packing arbitrarily oriented Hamilton cycles in $\mathcal{D}(n,p)$

In this section we prove Theorem 1.2. The proof naturally splits into two pieces. In the first piece, which appears in the next subsection, we will describe and analyse a simple randomized embedding algorithm to generate long paths of some fixed orientation. Then, in subsection 3.2 by repeatedly running this embedding algorithm in $\mathcal{D}(n,p)$ we will find a large subpath from each cycle C_i. Combined with an additional argument to close each of these paths to a cycle, this will prove Theorem 1.2.

3.1 A randomized algorithm for embedding oriented paths

Let D be an n-vertex digraph with $\delta^0(D) \geq n - \Delta$. Also let $P = v_1 \cdots v_\ell$ be a σ-path, for some arbitrary $\sigma : [\ell - 1] \to \{+,-\}$. Our aim in this section is to describe a randomized algorithm which w.h.p. finds a copy $Q := x_1 \cdots x_\ell$ of P in D over ℓ rounds. Let us fix a parameter p_{ex}, with $p_{ex} \ll p$. Throughout the algorithm, whenever we 'expose an edge', we mean that we toss a biased coin with heads probability p_{ex}, then regard the edge as present if the coin comes up heads (and indeterminate otherwise).

Path embedding algorithm:

1. To begin, select a vertex $x_1 \in V(D)$ uniformly at random and set $Q_1 = x_1$.
2. For $1 \leq i \leq \ell - 1$: suppose we are in round i and that we have now found $Q_i = x_1 \cdots x_i$, and aim to extend it to Q_{i+1} by finding x_{i+1}. Let $R_i = V(D) \setminus V(Q_i)$ and select an ordering of R_i uniformly at random, say $y_1, \ldots, y_{n-i}$.
3. To find x_{i+1} proceed as follows. First expose x_iy_1 with an orientation corresponding to $\sigma(i)$, with probability p_{ex}. If this pair is exposed as an edge <u>and</u> is an edge of D, set $x_{i+1} = y_1$ and $Q_{i+1} = Q_ix_{i+1}$. Otherwise expose x_iy_2 with an orientation corresponding to $\sigma(i)$, with probability p_{ex}. Again, if the exposed pair appears as an edge <u>and</u> is an edge of D, set $x_{i+1} = y_2$ and $Q_{i+1} = Q_ix_{i+1}$. Continue with this process until we either find x_{i+1} and Q_{i+1}, or run out of vertices in R_i. If this second case occurs, terminate the algorithm and declare a failure. If there is no failure and $i < \ell - 1$ return to 2. for round $i + 1$. Otherwise, proceed to 4.
4. Output $Q := Q_\ell$.

To analyze the algorithm, we will be interested in the following events:

$$F = \text{"the algorithm fails"};$$
$$E_{\overrightarrow{uv}} = \text{"the edge } \overrightarrow{uv} \text{ is exposed during the algorithm"};$$

$$A_{u,v} = \text{“}\{u,v\} \cap \big(V(Q) \setminus \{x_1, x_\ell\}\big) = \emptyset\text{”}.$$

The following lemma collects a number of key properties of this embedding process.

Lemma 3.1. *Let D be a digraph with $\delta^0(D) \geq n - \Delta$ and let P be a path of length ℓ. Suppose p_{ex} and ℓ satisfy $\log n/(n-\ell-\Delta) \ll p_{ex} \ll \min\left\{\frac{(n-\ell)^2}{n^2\Delta}, \frac{1}{(n\Delta)^{1/2}}\right\}$. Then running the path embedding algorithm with p_{ex} to find a copy of P in D, we have:*

(i) $\Pr(F) = o\left(n^{-2}\right)$.

(ii) $\Pr(E_{\overrightarrow{uv}}) \leq \frac{1+o(1)}{np_{ex}}$ *for every pair* $\{u,v\} \in \binom{V(D)}{2}$.

(iii) $\Pr(A_{u,v}) \leq (1+o(1))\left(\frac{n-\ell}{n}\right)^2$ *for every pair* $\{u,v\} \in \binom{V(D)}{2}$.

Proof. We first prove *(i)*. Note that the algorithm only ends in failure if for some $i \in [\ell-1]$ edges of orientation $\sigma(i)$ in $E(D)$ between x_i and all vertices of R_i were exposed, but none appeared as an edge. Using $|R_i| \geq n-\ell$, we see that

$$\Pr(F) \leq \ell(1-p_{ex})^{n-\ell-\Delta} \leq ne^{-(n-\ell-\Delta)p_{ex}} = o(n^{-2}),$$

where the last inequality holds since $p_{ex} = \omega\left(\frac{\log n}{n-\ell-\Delta}\right)$.

To see *(ii)* and *(iii)* it is helpful to think of the algorithm as proceeding in a slightly different, but equivalent way. First select a random subdigraph G of D_n, where each directed edge of D_n appears independently in G with probability p_{ex}. Now simply run the original algorithm to find a copy of P, but this time instead of exposing edges with probability p_{ex}, we add the edge if the corresponding edge is present in G. Clearly this gives an identical distribution on paths which appear as Q.

Now we claim that for all vertices $u_1, u_2, v_1, v_2 \in V(D)$ with $u_i \neq v_i$ for $i = 1,2$

$$\Pr\left(E_{\overrightarrow{u_1v_1}}\right) \leq \Pr\left(E_{\overrightarrow{u_2v_2}}\right) + (8\Delta+12)p_{ex}, \tag{3}$$

and that

$$\Pr(A_{u_1,v_1}) \leq \Pr(A_{u_2,v_2}) + (8\Delta+12)p_{ex}. \tag{4}$$

To see this, first note that if both u_1 and u_2 have the same in and out-neighbourhoods in D, and v_1 and v_2 have the same in and out-neighbourhoods in D, then $\Pr\left(E_{\overrightarrow{u_1v_1}}\right) = \Pr\left(E_{\overrightarrow{u_2v_2}}\right)$ and $\Pr(A_{u_1,v_1}) = \Pr(A_{u_2,v_2})$. The key observation to proving (3) and (4) is that conditional on a high probability event, we can assume that this 'same neighbourhood' property holds.

Concretely, let

$$S = \left\{z \in V(D) : \text{ at least one of the edges } \overrightarrow{yz}, \overrightarrow{zy} \text{ is not in } D \text{ for some } y \in \{u_1,u_2,v_1,v_2\}\right\}.$$

That is, S is the set of vertices which are not in-neighbours or out-neighbours in D of at least one vertex from $\{u_1,u_2,v_1,v_2\}$. Now consider the following event

$$B = \text{“no edge } \overrightarrow{yz} \text{ or } \overrightarrow{zy} \text{ appears in } G, \text{ where } y \in \{u_1,u_2,v_1,v_2\} \text{ and } z \in S \cup \{u_1,u_2,v_1,v_2\}\text{”}.$$

Note that conditional on B, by symmetry of the neighbourhoods of u_1, u_2, v_1 and v_2, the path embedding algorithm is equally likely to expose the $\overrightarrow{u_1v_1}$ and the edge $\overrightarrow{u_2v_2}$, i.e. $\Pr\left(E_{\overrightarrow{u_1v_1}}|B\right) = \Pr\left(E_{\overrightarrow{u_2v_2}}|B\right)$. But this gives

$$\Pr\left(E_{\overrightarrow{u_1v_1}}\right) = \Pr\left(E_{\overrightarrow{u_1v_1}}|B\right)\Pr(B) + \Pr\left(E_{\overrightarrow{u_1v_1}}|B^c\right)\Pr(B^c)$$

$$\le \Pr\left(E_{\overrightarrow{u_2v_2}}|B\right)\Pr(B) + \Pr(B^c)$$
$$\le \Pr\left(E_{\overrightarrow{u_2v_2}}\right) + \Pr(B^c). \tag{5}$$

Similarly, conditional on B, by symmetry, the pairs $\{u_1, v_1\}$ and $\{u_2, v_2\}$ are equally likely to be disjoint from $V(Q) \setminus \{x_1, x_\ell\}$ and $\Pr(A_{u_1,v_1}|B) = \Pr(A_{u_2,v_2}|B)$. Therefore, an identical calculation to (5) gives

$$\Pr(A_{u_1,v_1}) \le \Pr(A_{u_2,v_2}) + \Pr(B^c).$$

But $\Pr(B^c) \le (8\Delta + 12)p_{ex}$, as each $u \in V(D)$ has $\le \Delta$ non in-neighbours, $\le \Delta$ non out-neighbours in D and there are at most 12 edges between vertices in $\{u_1, u_2, v_1, v_2\}$ in D. This gives (3) and (4).

Now we can prove *(ii)*. For each $\overrightarrow{uv} \in E(D)$, let $C_{\overrightarrow{uv}}$ denote the event

$$C_{\overrightarrow{uv}} = \text{“}\overrightarrow{uv} \text{ gets exposed during the algorithm and } \overrightarrow{uv} \in E(G)\text{”}.$$

Clearly, we have $\Pr(C_{\overrightarrow{uv}}) = p_{ex} \times \Pr(E_{\overrightarrow{uv}})$. Let X denote the random variable which counts the number of edges in $G \cap D$ which get successfully exposed. Using (3), for any edge $\overrightarrow{uv} \in E(D)$ we have

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{\overrightarrow{xy} \in E(D)} \Pr(C_{\overrightarrow{xy}}) \\
&\ge \sum_{\overrightarrow{xy} \in E(D)} p_{ex} \times \left(\Pr(E_{\overrightarrow{uv}}) - (8\Delta + 12)p_{ex}\right) \\
&\ge \left(n(n-1) - \Delta n\right)p_{ex} \times \left(\Pr(E_{\overrightarrow{uv}}) - (8\Delta + 12)p_{ex}\right) \\
&= (1 - o(1))n^2 p_{ex} \times \left(\Pr(E_{\overrightarrow{uv}}) - (8\Delta + 12)p_{ex}\right).
\end{aligned} \tag{6}$$

However we always have $X \le \ell$, as each successfully exposed edge in $G \cap D$ completes a round of the algorithm and the algorithm consists of at most ℓ rounds. Combined with (6) this gives

$$\Pr(E_{\overrightarrow{uv}}) \le (1 + o(1))\frac{\ell}{n^2 p_{ex}} + (8\Delta + 12)p_{ex} = (1 + o(1))\frac{1}{np_{ex}},$$

since $\ell \le n$ and $p_{ex} = o\left((n\Delta)^{-1/2}\right)$. By applying (3) again, we conclude that $\Pr(E_{\overrightarrow{xy}}) \le (1+o(1))\frac{1}{np_{ex}}$ for all distinct $x, y \in V(D)$, completing *(ii)*.

Lastly, it is left to prove *(iii)*. Let Y denote the random variable which counts the number of pairs $\{u, v\}$ with $\{u, v\} \cap \left(V(Q) \setminus \{x_1, x_\ell\}\right) = \emptyset$. Note that we always have $Y \le \binom{n}{2}$ and that if F^c holds then $Y = \binom{n-\ell+2}{2}$. Since by *(i)* we have $\Pr(F) = o(n^{-2})$, it therefore follows that

$$\mathbb{E}(Y) \le \Pr(F)\binom{n}{2} + \Pr(F^c)\binom{n-\ell+2}{2} \le (1 + o(1))\frac{(n-\ell)^2}{2}.$$

But for distinct $u, v \in V(D)$, from (4) we have

$$\mathbb{E}(Y) = \sum_{\{x,y\} \in \binom{[n]}{2}} \Pr(A_{x,y}) \ge \binom{n}{2}\left(\Pr(A_{u,v}) - (8\Delta + 12)p_{ex}\right).$$

Rearranging, we obtain $\Pr(A_{u,v}) \le (1 + o(1))\left(\frac{n-\ell}{n}\right)^2 + (8\Delta + 12)p_{ex} = (1 + o(1))\left(\frac{n-\ell}{n}\right)^2$, since $p_{ex} \ll \frac{1}{\Delta}\left(\frac{n-\ell}{n}\right)^2$, as required. □

3.2 Finding edge disjoint Hamilton cycles in $\mathcal{D}(n,p)$

In this subsection we prove Theorem 1.2. In the proof, it will be useful to assume that p is not too large at certain points in the argument. We will assume that $\log^3 n/n \ll p \leq n^{-2/3}$. The general situation can be reduced to this as follows. If $p \geq n^{-2/3}$ let $p' = n^{-5/6}$ and $k = n^{5/6}p$, so that $n^{1/6} \ll k \leq n^{5/6}$. Then, after generating $D \sim \mathcal{D}(n,p)$, we further partition D into k subdigraphs $D_1, \dots, D_k$, where each edge e is assigned to D_i with probability $1/k$. It is clear that each D_i is distributed as $\mathcal{D}(n,p')$. Therefore, if we prove that the statement of the Theorem holds w.v.h.p (probability $1-o(1/n)$) when $\log^3 n/n \ll p \leq n^{-2/3}$, by taking a union bound over all the digraphs $D_1, \dots, D_k$ above, we prove it w.h.p. for all $p \gg \log^3 n/n$.

Suppose that $p = \alpha^3 \log^3 n/n$. Since $\log^3 n/n \ll p \leq n^{-2/3}$ we have $1 \ll \alpha \leq n^{1/9}$. Also let $\ell = n - n/\alpha \log n$ and $\Delta = n^{1/3}$. Let $0 < \epsilon$, let $t = (1-\epsilon)np = (1-\epsilon)\alpha^3 \log^3 n$ and let $C_1, \dots, C_t$ be cycles as given in the statement. Note that we may assume that ϵ is sufficiently small (i.e. $\epsilon \ll 1$). Set $p_1 = (1-\epsilon/2)p$ and choose p_2 so that $(1-p_1)(1-p_2) = 1-p$. Note that $p_2 = (1+o_\epsilon(1))\epsilon p/2$. Furthermore, let $M = \alpha \log n$ and take p_{ex} so that $(1-p_{ex})^M = 1-p_1$. This gives $p_{ex} = (1+o_\epsilon(1))\alpha^2 \log^2 n/n$. Also, from each C_i we select an oriented subpath P_i of length ℓ, with orientation σ_i.

Our general plan is to embed the paths $\{P_i\}_{i\in[t]}$ into $\mathcal{D}(n,p_1)$ by repeatedly applying the algorithm described in the previous section. We will then expose new edges with probability p_2, to complete each copy of P_i into a copy of the cycle C_i. Of course, we ensure that the obtained cycles are edge disjoint. The embedding scheme proceeds in two stages.

Stage 1: Finding edge disjoint copies of $P_1, \dots, P_t$ in $D_1 \sim \mathcal{D}(n,p_1)$

Following an idea introduced in [10], in this stage we give a randomized algorithm which w.v.h.p. finds edge disjoint copies of $P_1, \dots, P_t$ in $D_1 \sim \mathcal{D}(n,p_1)$. To formally describe this process, it will be helpful to generate $\mathcal{D}(n,p_1)$ in an alternative manner. To each directed edge $e \in E(D_n)$, associate e with t independent Bernoulli random variables $C_1^e, \dots, C_t^e$, each coin coming up heads with probability p_{ex}. Then let D_1 denote the random subdigraph of D_n where each edge e is included if some $\{C_j^e\}_{j\in[M]}$ appears heads up (NB: note the appearance of M rather than t here, and that as $M \leq t$ this is always well-defined). Each edge e appears independently in D_1 with probability $1-(1-p_{ex})^M = p_1$, and so D_1 is distributed according to $\mathcal{D}(n,p_1)$. We will gradually expose D_1 using the coins $\{C_i^e\}_{i\in[t]}$, always maintaining a 'fresh coin' for each unused edge. Provided we never examine more than the first M coins for any $e \in E(D_n)$, the above coupling shows that the exposed random digraph is generated according to $\mathcal{D}(n,p_1)$ (for more details about this idea, the reader is referred to [10]).

We now describe the algorithm. To begin, initialize to Round 1 and set counters $M_e = 1$ for all $e \in E(D_n)$. Proceed as follows:

1. In round i we have copies of $P_1, \dots, P_{i-1}$, denoted $Q_1, \dots, Q_{i-1}$. Set $D^{(i)} = D_n \setminus (\cup_{j<i} E(Q_j))$.

2. Apply the path embedding algorithm from the previous subsection to find a copy of P_i in $\mathcal{D}(D^{(i)}, p_{ex})$, which we denote by Q_i. If the edge $e \in E(D^{(i)})$ is exposed during this algorithm, it appears as an edge according to whether the fresh coin $C_{M_e}^e$ appears as heads. If this subroutine fails, declare a failure and terminate the algorithm.

3. If the edge e was exposed during the previous step, increment M_e by one. Otherwise leave M_e unchanged.

4. If $i < t$ then return to 1. in round $i+1$ to find Q_{i+1}.

5. If $i = t$ but some $M_e > M$ declare a failure and terminate the algorithm. Otherwise return $Q_1, \dots, Q_t$.

As mentioned above, in the key step 2. of the algorithm we always uses 'fresh coins' to expose edges $e \in E(D^{(i)})$. Furthermore, if the algorithm does not terminate in failure then we have found $Q_1, \dots, Q_t$ and $M_e \leq M$ for all e. By our coupling above, this guarantees that the resulting paths lie in $D_1 \sim \mathcal{D}(n, p_1)$.

We now analyse the failure probability. Given distinct vertices $u, v \in V(D_n)$, consider the following random variables:

$$X_{\overrightarrow{uv}} := M_{\overrightarrow{uv}}; \quad Y_{u,v} := \left|\left\{i \in [t] : \{u,v\} \cap \left(V(Q_i) \setminus \{x_{i,1}, x_{i,\ell}\}\right) = \emptyset\right\}\right|.$$

We claim that the following three properties hold with probability $1 - o(n^{-1})$:

(a) The algorithm succeeds in finding copies of $P_1, \dots, P_t$;

(b) $X_{\overrightarrow{uv}} \leq \frac{p_1}{p_{ex}}$ for all distinct $u, v \in V(D)$;

(c) $Y_{u,v} \leq (1+\epsilon) \times t \times \left(\frac{n-\ell}{n}\right)^2$ for all distinct $u, v \in V(D)$.

This will complete Stage 1, as if both (a) and (b) hold then our algorithm did not end in failure. Indeed, by (a) we have found copies of $P_1, \dots, P_t$ and by (b) we have $M_{\overrightarrow{uv}} = X_{\overrightarrow{uv}} \leq \frac{p_1}{p_{ex}} \leq M$, using that $1 - Mp_{ex} \leq (1-p_{ex})^M = 1 - p_1$. (Property (c) will be needed for Stage 2.)

To see (a) note that in round i, each vertex has at most 2 neighbours in each Q_j for $j < i$, and therefore $\delta^0(D^{(i)}) \geq n - 2(i-1) \geq n - 2t \geq n - 2\Delta$. Note that

$$\begin{aligned}\frac{\log n}{n-\ell-\Delta} = \frac{\log n}{n/\alpha \log n - n^{1/3}} \leq \frac{2\alpha \log^2 n}{n} \ll p_{ex} \ll p &\leq \min\left\{\frac{1}{\alpha^2 \log^2 n (n^{1/3})}, n^{-2/3}\right\} \\ &\leq \min\left\{\frac{(n-\ell)^2}{n^2 \Delta}, \frac{1}{(n\Delta)^{1/2}}\right\}.\end{aligned}$$

Therefore, by Lemma 3.1 *(i)* the path embedding algorithm succeeds in round i with probability at least $1 - o(n^{-2})$. Therefore, by a union bound, the algorithm succeeds in producing a copy of $P_1, \dots, P_t$ with probability $1 - o(n^{-1})$.

We now prove (b). Given distinct vertices $u, v \in V(D_n)$ we have $X_{\overrightarrow{uv}} = \sum_{i \in [t]} X_{\overrightarrow{uv}}(i)$, where $X_{\overrightarrow{uv}}(i)$ denotes the indicator random variable of the event that one of the coins $C_j^{\overrightarrow{uv}}$ is exposed during round i. Note that from Lemma 3.1 *(ii)*, conditional on any choice of $D^{(i)}$, we have

$$\Pr\left(X_{\overrightarrow{uv}}(i) = 1 | D^{(i)}\right) \leq (1 + \epsilon/4)\frac{1}{np_{ex}}.$$

By Corollary 2.4, we have

$$\Pr\left(X_{\overrightarrow{uv}} \geq (1+\epsilon/2)\frac{t}{np_{ex}}\right) \leq e^{-\epsilon^2 t/(64 np_{ex})} = o(1/n^3).$$

This holds as $t/np_{ex} \geq (1-\epsilon)np/np_{ex} \gg \log n$. Therefore, with probability $1 - o(n^{-1})$ we have $X_{\overrightarrow{uv}} \leq (1+\epsilon/2)\frac{t}{np_{ex}} = (1+\epsilon/2)\frac{(1-\epsilon)np}{np_{ex}} \leq \frac{p_1}{p_{ex}}$ for all $\overrightarrow{uv} \in E(D_n)$. This proves (b).

Lastly, (c) is similar to (b). Given distinct $u, v \in V(D_n)$ we have $Y_{u,v} = \sum_{i\in[t]} Y_{u,v}(i)$, where $Y_{u,v}(i)$ denotes the indicator random variable of the event $\{u, v\} \cap \big(V(Q_i) \setminus \{x_{i,1}, x_{i,\ell}\}\big) = \emptyset$. By Lemma 3.1 *(iii)*, conditional on any choice of $D^{(i)}$, we have

$$\Pr\left(Y_{u,v}(i) = 1|D^{(i)}\right) \leq (1 + \frac{\epsilon}{2})(\frac{n-\ell}{n})^2.$$

By Corollary 2.4 we find

$$\Pr\left(Y_{u,v} \geq (1+\epsilon)\frac{t(n-\ell)^2}{n^2}\right) \leq e^{-\epsilon^2 t(n-\ell)^2/16n^2} = o(1/n^4).$$

Here we used that $t(n-\ell)^2/n^2 \gg \log n$. By applying the union bound we obtain (c).

Stage 2: Completing the copies of $P_1, \ldots, P_t$ to copies of $C_1, \ldots, C_t$.

Let us suppose that in Stage 1 we found $Q_1, \ldots, Q_t$ in $D_1 \sim \mathcal{D}(n, p_1)$, and that property (c) above holds. In this stage we will prove that with probability $1 - o(1/n)$ it is possible to use edges of $D_2 \sim \mathcal{D}(n, p_2)$ to complete each oriented path Q_i to a copy of C_i which is edge disjoint from the other C_j's.

To see this, for each $i \in [t]$ let $W_i = V(D) \setminus \{x_{i,2}, \ldots, x_{i,\ell-1}\}$ (recall that $Q_i = x_{i,1} \ldots x_{i,\ell}$). Let G_i denote the digraph on vertex set W_i consisting of all directed edges which do not lie in the paths $P_1, \ldots, P_t$. Clearly we have $\delta^0(G_i) \geq |W_i| - 2t \geq |W_i| - 2\Delta$, which by the choice of the parameters is at least $(1 - 1/\log^2 n)|W_i|$. Also, by property (c) from Stage 1 for each $\overrightarrow{uv} \in E(G_i)$ we have $Y_{u,v} \leq (1+\epsilon)t(n-\ell)^2/n^2$.

Now select $D_2 \sim \mathcal{D}(n, p_2)$, where (recall) $p_2 = (1 + o_\epsilon(1))\epsilon p$. Given D_2, we obtain a random subdigraph F_i of G_i by assigning $\overrightarrow{uv} \in E(D_2)$ with probability $1/Y_{u,v}$ to some F_i with $\{u, v\} \subset V(W_i)$ (if $Y_{u,v} = 0$ then simply discard the edge $\overrightarrow{uv}$). By (c), each edge of G_i appears independently in F_i with probability

$$\frac{p_2}{Y_{u,v}} \geq \frac{(1 + o_\epsilon(1))\epsilon p n^2}{(1+\epsilon)t(n-\ell)^2} \geq \frac{\epsilon n}{2(n-\ell)^2} := p_{in}.$$

Therefore the distribution of F_i stochastically dominates that of $H_i \sim \mathcal{D}(G_i, p_{in})$.

Now to complete the proof, let P_i^c denote the complementary path to P_i in C_i. Using $n - \ell = n/\alpha \log n$, we find

$$p_{in} \geq \frac{\epsilon\alpha \log n}{2(n-\ell)} \gg \frac{\log |W_i|}{|W_i|}.$$

Therefore we can apply Lemma 2.5 to obtain that with probability $1 - o(1/n^2)$ for all $i \in [t]$, the digraph H_i (and therefore also F_i) contains a copy of P_i^c from $x_{i,1}$ to $x_{i,\ell}$ in W_i, denoted Q_i^c. But combining Q_i with Q_i^c for each $i \in [t]$ we obtain a copy of C_i. Therefore with probability $1 - o(1/n)$, for all $i \in [t]$, the digraph $Q_i \cup F_i$ contains a copy of C_i.

Stage 1 and 2 together prove that if $D_1 \sim \mathcal{D}(n, p_1)$ and $D_2 \sim \mathcal{D}(n, p_2)$ then with probability at least $1 - o(1/n)$ the digraph $D_1 \cup D_2$ contains edge disjoint copies of $C_1, \ldots, C_t$. As $D_1 \cup D_2$ can be coupled as a subgraph of $D \sim \mathcal{D}(n, p)$. This proves that the theorem holds with probability $1 - o(n^{-1})$ for $\log^3 n \ll p \ll n^{-2/3}$, and therefore by the reduction mentioned at the beginning, w.h.p. for all $p \gg \log^3 n/n$.

4 Counting

In this section we prove Theorem 1.3.

Proof of Theorem 1.3. First, let us prove the upper bound. Given p and any σ-cycle C, the expected number of copies of C in $D \sim \mathcal{D}(n,p)$ is at most $n!p^n$. Therefore, using Markov's inequality, we obtain that with probability at least $1-1/K$ there are at most $Kn!p^n$ many such copies. Therefore, by setting $K = \log n$ (say) we obtain that w.h.p. $\mathcal{D}(n,p)$ contains at most $n!p^n \log n = (1+o(1))^n n!p^n$ many such copies.

Next, we wish to prove the lower bound. In order to do so, suppose that $p = \alpha^2(\log\log n)\log n/n$ for some function $\alpha = \alpha(n)$ which tends to infinity with n. Let us also set $\ell = n - n/\alpha(\log\log n)$.

Let C be an n-vertex σ-cycle for some $\sigma \in \{+,-\}^n$. Let $\rho \in \{+,-\}^\ell$ denote the vector given by $\rho(i) = \sigma(i)$ for all $i \in [\ell]$ and let P denote the ρ-subpath of C. Let us set $p_1 = (1-\epsilon)p$ and $p_2 = \epsilon p$, for fixed small constant $\epsilon > 0$. We prove that $D \sim \mathcal{D}(n,p)$ contains many copies of C in two stages. In the first stage we show that w.h.p. $D_1 \sim \mathcal{D}(n,p_1)$ contains $(1-o_\epsilon(1))^n n!p^n$ copies of P. In the second stage, we expose a further random digraph $D_2 \sim \mathcal{D}(n,p_2)$ and show that w.h.p. 'most' of the copies Q of P in D_1 extend to a copy of C in $D_2 \cup Q$.

Stage 1: $D_1 \sim \mathcal{D}(n,p_1)$ contains at least $(1-3\epsilon)^n n!p^n$ copies of P w.h.p.

To begin, consider the following way to select a random copy of P, denoted $Q = x_1 \cdots x_\ell$, in some *fixed* digraph D on n vertices.

1. In the first round, select a vertex $x_1 \in V(D)$ uniformly at random and set $Q_1 := x_1$.

2. Suppose now that we are in round i, for some $1 \le i \le \ell - 1$ and so far we have found $Q_i = x_1 \cdots x_i$ and aim to extend it to Q_{i+1}, by selecting x_{i+1}. Let R_i denote the $\sigma(i)$-neighbourhood of x_i in $V(D) \setminus V(Q_i)$, i.e. $R_i = N^{\sigma(i)}(x_i) \cap \big(V(D) \setminus V(Q_i)\big)$.

3. Select a vertex uniformly at random from R_i and set it equal to x_{i+1} and $Q_{i+1} := Q_i x_{i+1}$. If no such vertex exists declare a failure and terminate the algorithm. If $i < \ell - 1$, return to 1. for round $i+1$.

4. If $i = \ell - 1$, output $Q := Q_\ell$.

Running this randomized algorithm results in a distribution on the set of all ρ-paths Q in D. We will write $\mathcal{F}(D)$ for this distribution.

We will now analyse the above algorithm while running on $\mathcal{D}(n,p_1)$. Select $D_1 \sim \mathcal{D}(n,p_1)$ and $Q \sim \mathcal{F}(D_1)$. For each $i \in [\ell]$, we will be interested in the following event:

$$E_i = \text{“}|R_j| \ge (1-\epsilon)(n-i)p_1 \text{ for all } j < i\text{”}.$$

Note that if the algorithm ends in failure, E_ℓ^c must occur. We claim that

$$\Pr_{\substack{D_1 \sim \mathcal{D}(n,p_1) \\ Q \sim \mathcal{F}(D_1)}} (E_\ell) = 1 - o(1). \tag{7}$$

To see this, we analyse the algorithm by generating D_1 in an 'online fashion', exposing edges as we go. Suppose now that we are in round i of the algorithm and have so far found $Q_i = x_1 \cdots x_i$. Expose all edges of D_1 in direction $\sigma(i)$ between x_i and $V(D_1) \setminus V(Q_i)$. Note that under this process, each edge is exposed at most once, and so can be coupled as a subgraph of $D_1 \sim \mathcal{D}(n, p_1)$. Clearly with this process, $|R_i| \sim \mathrm{Bin}(n-i, p_1)$. Therefore, by Chernoff's inequality (see Remark 2.5 in [17]), we have

$$\Pr_{\substack{D_1 \sim \mathcal{D}(n,p_1) \\ Q \sim \mathcal{F}(D_1)}} \big(|R_i| < (1-\epsilon)(n-i)p_1 \bigm| Q_i\big) \le e^{-2\epsilon^2 (n-i)p_1} \le e^{-2\epsilon^2 (n-\ell)p_1} = o(n^{-1}).$$

Here we have used that

$$\epsilon^2 (n-\ell) p_1 \ge \epsilon^2 (n-\ell)p/2 \ge \alpha \log n/2 \gg \log n.$$

However, this gives that

$$\Pr_{\substack{D_1 \sim \mathcal{D}(n,p_1) \\ Q \sim \mathcal{F}(D_1)}} (E_{i+1} |\ E_i) \ge 1 - o(n^{-1}).$$

In turn this gives (7), since $\ell \le n$ and

$$\Pr_{\substack{D_1 \sim \mathcal{D}(n,p_1) \\ Q \sim \mathcal{F}(D_1)}} (E_\ell) \ge \prod_{i \in [\ell-1]} \Pr_{\substack{D_1 \sim \mathcal{D}(n,p_1) \\ Q \sim \mathcal{F}(D_1)}} (E_{i+1} |\ E_i) \ge \big(1 - o(n^{-1})\big)^{\ell-1} = 1 - o(1).$$

Now note that (7) shows that if we select $D_1 \sim \mathcal{D}(n, p_1)$ then w.h.p.

$$\Pr_{Q \sim \mathcal{F}(D_1)} (E_\ell) = 1 - o(1).$$

However, for each σ-path $\widetilde{Q} = x_1 \cdots x_\ell$ in D_1 which satisfies E_ℓ we have

$$\Pr_{Q \sim \mathcal{F}(D_1)} (Q = \widetilde{Q}) \le \prod_{i \in [\ell-1]} \frac{1}{|R_i|} \le \prod_{i \in [\ell-1]} \frac{1}{(1-\epsilon)(n-i)p_1}.$$

Therefore, letting $\mathcal{Q}(D_1)$ denote the collection of all σ-paths in D_1, which satisfy E_ℓ, from (7) we have

$$1 - o(1) \le \Pr_{Q \sim \mathcal{F}(D_1)} (E_\ell) = \sum_{\widetilde{Q} \in \mathcal{Q}(D_1)} \Pr_{Q \sim \mathcal{F}(D_1)} \Big(Q = \widetilde{Q}\Big) \le \frac{|\mathcal{Q}(D_1)|}{(1-\epsilon)^{\ell-1} (n)_{\ell-1} p_1^{\ell-1}}.$$

Rearranging, this gives

$$|\mathcal{Q}(D_1)| \ge (1-\epsilon)^n (n)_{\ell-1} p_1^{\ell-1} \ge (1-2\epsilon)^n (n)_{\ell-1} p^{\ell-1} \ge (1-3\epsilon)^n n! p^n.$$

Here we used that

$$(n-\ell+1)! p^{n-\ell+1} \le ((n-\ell+1)p)^{n-\ell+1} \le (2\alpha \log n)^{n/\alpha \log\log n + 1} = (1 + o(1))^n.$$

Stage 2: Completing 'most' copies of P in D_1 to a copy of C.

Let $\mathcal{P}$ denote the collection of all copies of P in D_1. Let P^c denote the complement path of P in C (see notation). Note from the bound in Stage 1, w.h.p. we have $|\mathcal{P}| \ge (1-3\epsilon)^n n! p^n$. Let us fix $Q \in \mathcal{P}$, which starts at x_1 and ends at x_ℓ. Select $D_2 \sim \mathcal{D}(n, p_2)$. We will show that

$$\Pr\left(Q \text{ is contained in a copy of } C \text{ in } Q \cup D_2\right) = 1 - o(1). \tag{8}$$

To see this, set $W_Q := \big(V(D_1) \setminus V(Q)\big) \cup \{x_1, x_\ell\}$, so that $|W_Q| = n - \ell + 2$. But it is easy to see that $D_2[W_Q] \sim \mathcal{D}(n - \ell + 2, p_2)$ (perhaps some edges also appear in D_1, but this only helps us). Using that

$$p_2 = \epsilon p \geq \frac{\epsilon \alpha^2 (\log \log n) \log n}{n} \gg \frac{\log n}{n - \ell},$$

by Lemma 2.5 we find that $D_2[W_Q]$ w.h.p. contains a P^c path Q_2 from x_1 to x_ℓ. Combined with Q, this gives a copy of C in $Q \cup D_2$. This gives (8).

We now complete the proof of the theorem. Let $\mathcal{P}_{bad}$ denote the set of $Q \in \mathcal{P}$ which are *not* contained in a copy of C in $Q \cup D_2$. From (8) we have

$$\mathbb{E}(|\mathcal{P}_{bad}|) = o(|\mathcal{P}|).$$

By Markov's inequality this gives that w.h.p. $|\mathcal{P}_{bad}| = o(|\mathcal{P}|)$. Therefore, w.h.p. there are $|\mathcal{P} \setminus \mathcal{P}_{bad}| = (1 - o(1))|\mathcal{P}| \geq (1 - 4\epsilon)^n n! p^n$ paths Q which extend to a copy of C in $Q \cup D_2$. As each copy of C can be obtained from at most $2n$ such paths (rotations along the cycle or flipping all the orientations), this gives $|\mathcal{P} \setminus \mathcal{P}_{bad}|/2n \geq (1 - 5\epsilon)^n n! p^n$ copies of C in $D_1 \cup D_2$. □

Acknowledgment. The authors would like to thank the referees for their very careful reading of the manuscript and for providing us with many valuable comments.

References

[1] N. Alon, J. Spencer, **The Probabilistic Method**, 3rd ed., John Wiley and Sons (2008).

[2] B. Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1), 257–274, 1984.

[3] F. Chung and L. Lu. Concentration inequalities and martingale inequalities – a survey, *Internet Math.* **3**(1) (2006), 79-127.

[4] L. DeBiasio, D. Kühn, T. Molla, D. Osthus and A. Taylor. Arbitrary orientations of Hamilton cycles in digraphs, in *SIAM Journal Discrete Mathematics* 29 (2015), 1553–1584.

[5] L. DeBiasio and T. Molla. Semi-degree threshold for anti-directed Hamilton cycles, *arXiv preprint*, arXiv:1308.0269.

[6] G. A. Dirac. Some theorems on abstract graphs, in *Proceedings of the London Mathematical Society*, 2 (1952), 69–81.

[7] A. Ferber. Closing gaps in problems related to Hamilton cycles in random graphs and hypergraphs, *The Electronic Journal of Combinatorics*, 22(1):P1–61, 2015.

[8] A. Ferber, G. Kronenberg and E. Long. Packing, counting and covering Hamilton cycles in random directed graphs, *Israel Journal of Mathematics*, to appear.

[9] A. Ferber, R. Nenadov, U. Peter, A. Noever, and N. Škoric. Robust Hamiltonicity of random directed graphs. *SODA '14*.

[10] A. Ferber and V. Vu. Packing perfect matchings in random hypergraphs, preprint.

[11] A. M. Frieze. An algorithm for finding Hamilton cycles in random directed graphs. *Journal of Algorithms*, 9(2):181–204, 1988.

[12] A. Frieze, and M. Krivelevich. On two Hamilton cycle problems in random graphs. *Israel Journal of Mathematics*, 166, no. 1, (2008): 221–234.

[13] A. Ghouila-Houri. Une condition suffisante dexistence dun circuit Hamiltonien, in *C.R. Acad. Sci. Paris* 25 (1960), 495-497.

[14] R. Glebov and M. Krivelevich. On the number of Hamilton cycles in sparse random graphs. *SIAM Journal on Discrete Mathematics*, 27(1):27–42, 2013.

[15] R. Häggkvist and A. Thomason. Oriented Hamilton cycles in digraphs, in *Journal of Graph Theory* 19.4 (1995), 471–479.

[16] D. Hefetz, M. Krivelevich and T. Szábo, Hamilton cycles in highly connected and expanding graphs. Combinatorica 29 (2009), 547–568.

[17] S. Janson, T. Łuczak, and A. Ruciński, **Random graphs**. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York (2000).

[18] F. Knox, D. Kühn, and D. Osthus. Edge-disjoint Hamilton cycles in random graphs. *Random Structures & Algorithms*, 2013.

[19] M. Krivelevich and W. Samotij. Optimal packings of Hamilton cycles in sparse random graphs. *SIAM Journal on Discrete Mathematics*, 26(3):964–982, 2012.

[20] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: applications. *Journal of Combinatorial Theory, Series B*, 104:1–27, 2014.

[21] C. McDiarmid. Clutter percolation and random graphs, in *Combinatorial Optimization II*, 17–25. Springer, 1980.

[22] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[23] A. Thomason. Paths and cycles in tournaments, in *Transactions of the American Mathematical Society*, 296(1) (1986), 167–180.