Packing and counting arbitrary Hamilton cycles in random digraphs

Asaf Ferber *  Eoin Long †

February 22, 2016

Abstract

We prove packing and counting theorems for arbitrarily oriented Hamilton cycles in $D(n,p)$ 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 $D \sim D(n,p)$ 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 $D \sim D(n,p)$ w.h.p. contains $(1 \pm o(1))np^n$ copies of $C$, provided $p \geq \log^{1+o(1)} n / n$.

1 Introduction

A Hamilton cycle in a graph or digraph is a cycle passing through every vertex of the graph exactly once, and a graph is Hamiltonian if it contains a Hamilton cycle. 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 most celebrated, sufficient conditions for Hamiltonicity in graphs was established by Dirac [5] in 1952. He proved that every graph on $n$ vertices, $n \geq 3$, with minimum degree at least $n/2$ is Hamiltonian. Ghouila-Houri [11] proved an analogue of Dirac’s theorem for digraphs by 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 Thomasson [19] showed that each regular tournament contains every orientation of a Hamilton cycle. Later on, Häggkvist and Thomason [13] showed an approximate analog of the result of Ghouila-Houri [11] while proving that $\delta^0(G) \geq n/2 + n^{5/6}$ is sufficient to guarantee every orientation of a Hamilton cycle to appear in $G$. Very recently, this problem has been settled completely by DeBiasio, Kühn, Molla, Osthus and Taylor [3]. 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, DeBiasio and Molla showed in [4] 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, Yale University, and Department of Mathematics, MIT. Emails: asaf.ferber@yale.edu, and ferbera@mit.edu.
†School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel. Email: eoinlong@post.tau.ac.il.
In this paper we restrict our attention to the sparse setting, that is, to random directed graphs. Let $D(n, p)$ 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 [17]. He showed (among other things) by using an elegant coupling argument that

$$\Pr[G \sim G(n, p) \text{ is Hamiltonian}] \leq \Pr[D \sim D(n, p) \text{ is Hamiltonian}].$$

Combined with the result of Bollobás [1] it follows that a typical $D \sim D(n, p)$ is Hamiltonian for $p \geq \frac{\ln n + \ln n + \omega(1)}{n}$. Frieze [10] later determined the correct threshold for the appearance of a Hamilton cycle in $D \sim D(n, p)$ is $p = \frac{\ln n + \omega(1)}{n}$.

While Frieze’s result gives a better bound than McDiarmid’s coupling argument, the latter is much more flexible (for some further applications, see [6]). For example, given an arbitrary oriented Hamilton cycle $C$, it follows immediately from McDiarmid’s proof that

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

In contrast, the result of Frieze is tailored to “consistently oriented” Hamilton cycles and gives no improvement for the obtained bound of $p = \frac{\ln n + \ln n + \omega(1)}{n}$. 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 $D \sim D(n, p)$ 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 [8]. Here it was proven using the “absorption method” that $D \sim D(n, p)$ 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 $p \geq \text{polylog}(n)/n$.

Here we deal with the problems of counting and packing arbitrary oriented Hamilton cycles in $D \sim D(n, p)$, for edge-densities $p \geq \text{polylog}(n)/n$. The analogous problems regarding the “consistently oriented” Hamilton cycles has been recently treated by Kronenberg and the authors in [7]. However, the proof method there is inapplicable to the arbitrary oriented case.

Enhancing a recent “online sprinkling” technique introduced by Ferber and Vu [9], we manage to tackle these two problems. Our first theorem gives an asymptotically optimal result for packing arbitrarily oriented Hamilton cycles in $D \sim D(n, p)$.

**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. $D \sim D(n, p)$ contains edge disjoint copies of $C_1, \ldots, C_t$, provided $p \gg \frac{\log^3 n}{n}$.

Our second result shows that given an arbitrarily oriented Hamilton cycle $C$, w.h.p. $D \sim D(n, p)$ 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 $D \sim D(n, p)$ contains $(1 \pm o(1))n!p^n$ distinct copies of $C$, provided $p \gg \frac{\log \log n \log n}{n}$.

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 [12, 14, 15, 16] and their references). The main difficulty when working in the directed case, is that the so called Posá rotation-extension technique (see [18]) does not work in its simplest form and therefore one should find more creative ways for generating Hamilton cycles.

**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 \(\delta^0(D) = \min_{u \in V(D), s \in \{+, -\}} d^s(u)\). 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 \(P := v_0v_1 \ldots v_k\) 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 \(\sigma : [0, k] \to \{+, -\}\), we say that \(P\) is a \(\sigma\)-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 \(\sigma : [1, k] \to \{+, -\}\) a \(\sigma\)-cycle \(C := v_1 \ldots v_k v_1\) 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 \(P^*\) denote the path induced by the edges of \(C\) which do not lie in \(P\), called the *compliment* 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 \(\mathcal{G}(n, p)\). When discussing events related to \(\mathcal{D}(D, p)\) or \(\mathcal{G}(G, p)\) we write with high probability (w.h.p.) to denote that the event occurs with probability \(1 - o(1)\). We say with very high probability (w.v.h.p.) to mean with probability \(1 - n^{-\omega(1)}\).

## 2 Tools

### 2.1 A concentration inequality

A sequence of random variables \(0 = X_0, X_1, \ldots, X_N\) is said a submartingale if

\[
\mathbb{E}(X_i | X_1, \ldots, X_{i-1}) \leq 0 \text{ for all } i \in [N].
\]

The next result gives a concentration bound for submartingales (see Theorem 27 in the survey of Chung and Lu [2], taking \(\phi_i = a_i = 0\)).

**Theorem 2.1.** Suppose that \(0 = X_0, X_1, \ldots, X_N\) is a submartingale which for all \(i \in [N]\) satisfies

\[
\text{Var}(X_i | X_1, \ldots, X_{i-1}) \leq \sigma; \quad \text{and} \quad X_i - \mathbb{E}(X_i | X_1, \ldots, X_{i-1}) \leq M.
\]

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 2.2.** 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 \(\mathbb{P}(A_i | A_1, \ldots, A_{i-1}) \leq q\). 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. Simply take \( Y_i \) to be the indicator random variable for the event \( A_i \), and set \( X_i = Y_i - q \) for all \( i \in [N] \), and \( X_0 \equiv 0 \). Then
\[
\mathbb{E}(X_i|X_1, \ldots, X_{i-1}) = \mathbb{P}(A_i|A_1, \ldots, A_{i-1}) - q \leq 0 \quad \text{for all } i \in [N],
\]
and so \( X_0, \ldots, X_N \) forms a submartingale. We also have
\[
\text{Var}(X_i|X_1, \ldots, X_{i-1}) \leq q \quad \text{for all } i.
\]
Taking \( M = 1 \) and \( \sigma = q \), Theorem 2.1 gives
\[
\mathbb{P}(E_m) = \mathbb{P}\left( \sum_{i=1}^{N} Y_i > nq + m \right) = \mathbb{P}\left( \sum_{i=1}^{N} X_i > m \right) \leq e^{-m^2/2(Nq+m/3)},
\]
as required. \( \square \)

2.2 Completing paths into a Hamilton cycle

We will also use the following result to complete paths into Hamilton cycles in \( D \sim D(n, p) \).

Lemma 2.3. Let \( \sigma \in \{+,-\}^n \). Suppose that \( G \) is an \( n \)-vertex digraph with \( \delta^0(G) = (1 - o(1))n \). Then a digraph \( D \sim D(G, p) \) contains a \( \sigma \)-path \( Q \) between any two distinct vertices in \( D \), provided \( p = \omega(\log n/n) \).

The analogue of Lemma 2.3 is well known for undirected graphs states that w.h.p. \( G \sim G(n, p) \) contains a Hamiltonian path between any two distinct vertices \( a, b \in V(G) \), provided \( p = \omega(\log n/n) \). (In fact, this result follows from many of the proofs of Hamiltonicity of \( G \sim G(n, p) \).) This undirected result can be ‘upgraded’ to the directed world to prove Lemma 2.3 using the nice coupling argument of McDiarmid [17] mentioned in the introduction.

3 Packing arbitrarily orientated Hamilton cycles in \( D(n, p) \)

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 \( D(n, p) \) 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 the 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 \( \sigma \)-path, for some arbitrary \( \sigma : [\ell - 1] \rightarrow \{+,-\} \). 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 \( \ell \) rounds. Throughout the algorithm, whenever we ‘expose an edge’ we mean to toss a biased coin with certain probability for this edge.

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 $x_i y_1$ with an orientation corresponding to $\sigma(i)$, with probability $p_{ex}$. If this pair is exposed as an edge and is an edge of $D$, set $x_{i+1} = y_1$ and $Q_{i+1} = Q_i x_{i+1}$. Otherwise expose $x_i y_2$ with an orientation corresponding to $\sigma(i)$, with probability $p_{ex}$. Again, if the exposed pair appears as an edge and is an edge of $D$, set $x_{i+1} = y_2$ and $Q_{i+1} = Q_i x_{i+1}$. 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:

$$F = \text{"the algorithm fails";}$$

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

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

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 $\ell$. Suppose $p_{ex}$ and $\ell$ satisfy $\log n / (n - \ell - \Delta) \ll p_{ex} \ll \min \{ \frac{(n-\ell)^2}{n^2 \Delta}, \frac{1}{(n \Delta)^{1/2}} \}$. Then running the path embedding algorithm with $p_{ex}$ to find a copy of $P$ in $D$, we have:

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

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

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

**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

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

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

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 $i = 1, 2$

$$\Pr[E_{u_1 v_2}] \leq \Pr[E_{u_2 v_1}] + (8 \Delta + 12)p_{ex}, \tag{1}$$
and that

$$\Pr[A_{u_1, v_1}] \leq \Pr[A_{u_2, v_2}] + (8\Delta + 12)p_{ex}. \tag{2}$$

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 $\Pr[E_{u_1v_1}] = \Pr[E_{u_2v_2}]$ and $\Pr[A_{u_1, v_1}] = \Pr[A_{u_2, v_2}]$. The key observation to proving (1) and (2) is that conditional on a high probability event, we can assume that this ‘same neighbourhood’ property holds.

Concretely, let

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

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

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

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 $u_1v_1$ and the edge $u_2v_2$, i.e. $\Pr[E_{u_1v_1} \mid B] = \Pr[E_{u_2v_2} \mid B]$. But this gives

$$\Pr[E_{u_1v_1}] = \Pr[E_{u_1v_1} \mid B] \Pr[B] + \Pr[E_{u_1v_1} \mid B^c] \Pr[B^c]$$

$$\leq \Pr[E_{u_2v_2}] \Pr[B] + \Pr[B^c]$$

$$\leq \Pr[E_{u_2v_2}] + \Pr[B^c]. \tag{3}$$

Similarly, conditional on $B$, by symmetry, the pairs $\{u_1, v_1\}$ and $\{u_2, v_2\}$ are equally likely to be disjoint from $V(Q) \setminus \{x_1, x_2\}$ and $\Pr[A_{u_1, v_1} \mid B] = \Pr[A_{u_2, v_2} \mid B]$. Therefore, an identical calculation to (3) gives

$$\Pr[A_{u_1, v_1}] \leq \Pr[A_{u_2v_2}] + \Pr[B^c].$$

But $\Pr[B^c] \leq (8\Delta + 12)p_{ex}$, as each $u \in V(D)$ has $\leq \Delta$ non in-neighbours, $\leq \Delta$ 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 (1) and (2).

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

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

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 (1), for any edge $\overrightarrow{uv} \in E(D)$ we have

$$\mathbb{E}(X) = \sum_{\overrightarrow{yz} \in E(D)} \Pr[C_{\overrightarrow{yz}}]$$

$$\geq \sum_{\overrightarrow{yz} \in E(D)} p_{ex} \times \left( \Pr[E_{\overrightarrow{yz}}] - (8\Delta + 12)p_{ex} \right)$$

$$\geq (n(n - 1) - \Delta n) p_{ex} \times \left( \Pr[E_{\overrightarrow{uv}}] - (8\Delta + 12)p_{ex} \right)$$

$$= (1 - o(1))n^2 p_{ex} \times \left( \Pr[E_{\overrightarrow{uv}}] - (8\Delta + 12)p_{ex} \right). \tag{4}$$
However we always have $X \leq \ell$, as each successfully exposed edge in $G \cap D$ completes a round of the algorithm and the algorithm consists of at most $\ell$ rounds. Combined with (4) this gives

$$Pr[E_{xy}] \leq (1 + o(1)) \frac{\ell}{n^2p_{ex}} + (8\Delta + 12)p_{ex} = (1 + o(1)) \frac{1}{np_{ex}},$$

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

Lastly, it is left to prove (iii). Let $Y$ denote the random variable which counts the number of pairs $\{u, v\}$ with $\{u, v\} \cap (V(Q) \setminus \{x_1, x_t\}) = \emptyset$. Note that we always have $Y \leq \binom{n}{2}$ and that if $F^c$ holds then $Y = \binom{n-\ell+2}{2}$. Since by (i) we have $Pr[F] = o(1/n)$, it therefore follows that

$$E(Y) \leq Pr[F]\left(\binom{n}{2}\right) + Pr[F^c]\left(\binom{n-\ell+2}{2}\right) \leq (1 + o(1))\binom{n-\ell}{2}.$$

But for distinct $u, v \in V(D)$, from (2) we have

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

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

### 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 $n^{1/6} \leq k \leq n^{5/6}$. Then, after generating $D \sim \mathcal{D}(n, p)$, we further partition $D$ into one of $k$ subdigraphs $D_1, \ldots, D_k$, 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 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 $D_1, \ldots, D_k$ above, we prove that 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 $\ell = n - n/\alpha \log n$, $\Delta = n^{1/3}$ and $p_{ex} = \alpha^2 \log^2 n/n$. Let $C_1, \ldots, C_t$ be cycles as given in the statement. Also set $p_1 = (1 - \ell/2)p$ and choose $p_2$ so that $(1 - p_1)(1 - p_2) = 1 - p$. Note that $p_2 = (1 + o_c(1))\epsilon p/2$. Also, from each $C_i$ we select an oriented subpath $P_i$ of length $\ell$, with orientation $\sigma_i$.

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 $P_1, \ldots, P_t$ in $D_1 \sim \mathcal{D}(n, p_1)$

In this stage we give a randomized algorithm which w.h.p. finds edge disjoint copies of $P_1, \ldots, P_t$ in $D_1 \sim \mathcal{D}(n, p_1)$. The algorithm exposes edges from $D_n$ in an online fashion with some small probability $p_{ex} = o(p_1))$. This will be carried out over a series of rounds, where in round $i$ we aim to create a copy of $P_i$.

Initialize to Round 1 and proceed as follows:
1. In round $i$ we have copies of $P_1, \ldots, P_{i-1}$, denoted $Q_1, \ldots, Q_{i-1}$. 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 $D^{(i)}$, denoted by $Q_i = x_{i,1} \ldots x_{i,t}$. If this subroutine fails, declare a failure and terminate the algorithm.

3. If $i < t$ then return to 1. in round $i+1$ to find $Q_{i+1}$. If $i = t$ then we have succeeded in finding edge disjoint copies of $P_1, \ldots, P_t$.

Given distinct vertices $u,v \in V(D_n)$, consider the following random variables:

$$X_{uv} = |\{i \in [t] : \text{edge } u \rightarrow v \text{ is exposed during round } i\}|; \quad Y_{u,v} = |\{i \in [t] : \{u,v\} \cap (V(Q_i) \setminus \{x_{i,1}, x_{i,t}\}) = \emptyset\}|.$$  

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

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

(b) $X_{uv} \leq \frac{p_{ex}}{p_{ex}}$ 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 of the algorithm. Indeed, by (b) w.h.p. we have exposed each edge with probability $p_{ex} \times X_{uv} \leq p_1$. The resulting digraph can therefore be coupled as a subgraph of $D_1 \sim \mathcal{D}(n, p_1)$ w.h.p.. By (a) this subdigraph contains edge disjoint copies of $P_1, \ldots, P_t$, as required. (Property (c) will be required 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

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

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

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

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

By Lemma 2.2, we have

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

This holds as $t/np_{ex} \geq np/np_{ex} \gg \log n$. Therefore, with probability $1 - o(n^{-1})$ we have $X_{uv} \leq (1 + \epsilon/2) \frac{t}{np_{ex}} = (1 + \epsilon/2) \frac{1 - \epsilon}{np_{ex}} \leq \frac{p_{ex}}{p_{ex}}$ for all $u \rightarrow v \in E(D_n)$. 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 \(\{u, v\} \cap (V(Q_i) \setminus \{x_i, x_i, x_i\}) = \emptyset\). By Lemma 3.1 \((iii)\), conditional on any choice of \(D^{(i)}\), we have

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

By Corollary 2.2 we find

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

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 \(D \sim D(n, p)\), and that property \((c)\) above holds. In this stage we will prove that w.h.p. it is possible to use edges of \(D_2 \sim D(n, p)\) to complete each oriented path \(Q_i\) to a copy of \(C_i\) which is edge-disjoint to other \(C_j\)-s.

To see this, for each \(i \in [t]\) let \(W_i = V(D) \setminus \{x_i, x_i, x_i\}\) (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 \(\delta^0(G_i) \geq |W_i| - 2t \geq |W_i| - 2\Delta = (1 - o(1))|W_i|\). Also, by property \((c)\) from Stage 1 for each \(u, v \in E(G_i)\) we have \(Y_{u,v} \leq (1 + \epsilon)t(n - \ell)^2/n^2\).

Now select \(D_2 \sim D(n, p)\), where \(p_2 = (1 + o(1))\epsilon p\). Given \(D_2\), we obtain a random subdigraph \(F_i\) of \(G_i\) by assigning \(\bar{uv} \in E(D_2)\) with probability \(1/Y_{u,v}\) to some \(F_i\) with \(\{u, v\} \subseteq V(W_i)\). By \((c)\), each edge of \(G_i\) appears independently in \(F_i\) with probability

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

Therefore we can apply Lemma 2.3 to obtain that with probability \(1 - o(1/n)\) for all \(i \in [t]\), the digraph \(H_i\) (and therefore also \(F_i\)) contains a copy of \(P_2\) from \(x_{i,1}\) to \(x_{i,\ell}\) in \(W_i\), denoted \(Q_2^i\). But combining \(Q_i\) with \(Q_2^i\) for each \(i \in [t]\) we obtain a copy of \(C_i\). Therefore w.h.p., 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 \(D_1 \sim D(n, p)\) and \(D_2 \sim D(n, p)\) then with probability at least \(1 - o(n^{-1})\) 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 \(D \sim D(n, p)\). This proves that the theorem holds with probability \(1 - o(n^{-1})\) for \(3n \ll p \ll n^{-2/3}\), 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. To prove the theorem, 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 $\sigma$-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 $\rho$-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 $D \sim D(n,p)$ contains many copies of $C$ in two stages. In the first stage we show that w.h.p. $D \sim D(n,p_1)$ contains $(1 - o(1))n!p^n$ copies of $P$. In the second stage, we expose a further random digraph $D_2 \sim D(n,p_2)$ 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: $D_1 \sim D(n,p_1)$ contains at least $(1 - 3\epsilon)n!p^n$ 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 $1 \leq i \leq \ell - 1$ 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. $R_i = N^\sigma(i)(x_i) \cap (V(D) \setminus V(Q_i))$.

3. Select a vertex uniformly at random from $R_i$ and set it equal to $x_{i+1}$ and $Q_{i+1} := Q_ix_{i+1}$. 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 $\rho$-paths $Q$ in $D$. We will write $\mathcal{F}(D)$ for this distribution.

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

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

Note that if the algorithm ended in failure, $E_i^c$ must occur. We claim that

$$\Pr_{D_1 \sim D(n,p_1), Q \sim \mathcal{F}(D_1)}[E_i] = 1 - o(1). \quad (5)$$

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 $D_1 \sim D(n,p_1)$. Clearly with this process, $|R_i| \sim \text{Bin}(n - i, p_1)$. Therefore, by Chernoff’s inequality, we have

$$\Pr_{D_1 \sim D(n,p_1), Q \sim \mathcal{F}(D_1)}[|R_i| < (1 - \epsilon)(n - i)p_1 \mid Q_i] \leq e^{-2\epsilon^2(n-i)p_1} \leq e^{-2\epsilon^2(n-\ell)p_1} = o(n^{-1}).$$

Here we have used that

$$\epsilon^2(n-\ell)p_1 \geq \epsilon^2(n - \ell)p/2 \geq \alpha \log n/2 \gg \log n.$$
However, this gives that
\[ \Pr_{D_1 \sim \mathcal{D}(n, p_1)} \left[ E_{i+1} \mid E_i \right] \geq 1 - o(n^{-1}). \]

In turn this gives (5), since \( \ell \leq n \) and
\[ \Pr_{D_1 \sim \mathcal{D}(n, p_1)} \left[ E_i \right] \geq \prod_{i \in \{1, \ldots, \ell - 1\}} \Pr_{D \sim \mathcal{F}(D_1)} \left[ E_{i+1} \mid E_i \right] \geq (1 - o(n^{-1}))^{\ell - 1} = 1 - o(1). \]

Now note that (5) shows that if we select \( D_1 \sim \mathcal{D}(n, p_1) \) then w.h.p.
\[ \Pr_{Q \sim \mathcal{F}(D_1)} \left[ E_\ell \right] = 1 - o(1). \]

However, for each \( \sigma \)-path \( \tilde{Q} = x_1 \cdots x_\ell \) in \( D_1 \) which satisfies \( E_\ell \) we have
\[ \Pr_{Q \sim \mathcal{F}(D_1)} (Q = \tilde{Q}) = \prod_{i \in \{1, \ldots, \ell - 1\}} \frac{1}{R_i} \leq \prod_{i \in \{1, \ldots, \ell - 1\}} \frac{1}{(1 - \epsilon)(n - i)p_1}. \]

Therefore, letting \( Q(D_1) \) denote the collection of all \( \sigma \)-paths in \( D_1 \), which satisfy \( E_\ell \), from (5) we have
\[ 1 - o(1) \leq \Pr_{Q \sim \mathcal{F}(D_1)} \left[ E_i \right] = \sum_{Q \in \mathcal{Q}} \Pr_{Q \sim \mathcal{F}(D_1)} \left[ Q = \tilde{Q} \right] \leq \frac{|Q(D_1)|}{(1 - \epsilon)^\ell (n)p_1^\ell}. \]
Rearranging, this gives
\[ |Q(D_1)| \geq (1 - \epsilon)^n (n)p_1^\ell \geq (1 - 2\epsilon)^n (n)p_1^\ell \geq (1 - 3\epsilon)^n n! p^n. \]

Here we used that
\[ (n - \ell)! p^{n-\ell} \leq (n - \ell)p)^{n-\ell} = (\alpha \log n)^{n/\alpha \log \log n} = (1 + o(1))^n. \]

**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 compliment path of \( P \) in \( C \) (see notation). Note from the bound in Stage 1, w.h.p. we have \(|\mathcal{P}| \geq (1 - 3\epsilon)^n p^n\). Let us fix \( Q \in \mathcal{P} \), which starts at \( x_1 \) and ends at \( x_\ell \). Select \( D_2 \sim \mathcal{D}(n, p_2) \). We will show that
\[ \Pr \left[ Q \text{ is contained in a copy of } C \text{ in } Q \cup D_2 \right] = 1 - o(1). \] (6)

To see this, set \( W_Q := (V(D_1) \setminus V(Q)) \cup \{x_1, x_\ell\} \), 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 \( D_1 \), but this only helps us). Using that
\[ p_2 = \epsilon p = \frac{\epsilon \alpha^2 (\log \log n) \log n}{n} \gg \frac{\log n}{n - \ell}, \]
by Lemma 2.3 we find that \( D_2|W_Q| \) w.h.p. contains a \( P^c \) path \( Q_2 \) from \( x_1 \) to \( x_\ell \). Combined with \( Q \), this gives a copy of \( C \) in \( Q \cup D_2 \). This gives (6).
We now complete the proof of the theorem. Let $P_{\text{bad}}$ denote the set of $Q \in P$ which are not contained in a copy of $C$ in $Q \cup D_2$. From (6) we have

$$E(|P_{\text{bad}}|) = o(|P|).$$

By Markov’s inequality this gives that w.h.p. $|P_{\text{bad}}| = o(|P|)$. Therefore, w.h.p. there are $|P \setminus P_{\text{bad}}| = (1 - o(1))|P| \geq (1 - 4\epsilon)n!p^n$ 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, this gives $|P \setminus P_{\text{bad}}|/2n \geq (1 - 5\epsilon)n!p^n$ copies of $C$ in $D_1 \cup D_2$. 

References


