HAMILTON CYCLES IN A CLASS OF RANDOM DIRECTED GRAPHS

by

Colin Cooper
School of Mathematical Sciences
Polytechnic of North London
London, U.K.

Alan Frieze
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213

Research Report No. 91–129

August 1991

The remaining case where $k = 1$ was discused in Cooper and Frieze [1], see also McDiarmid and Reed [6].

If $\ell = 0$ we write D_{k-in} and if $k = 0$ then we write $D_{\ell-out}$. If we drop the orientation in D_{k-out} then we obtain the underlying *undirected* graph G_{k-out}. This has been the object of considerable study and the main outstanding question is how large should k be in order that almost every (a.e.) G_{k-out} has a Hamilton cycle. It is currently known that $k \geq 5$ is sufficient, Frieze and Luczak [4] and it is conjectured that the correct lower bound for k is 3. This paper considers the directed version of this problem. We prove a slightly stronger result than is claimed in the abstract:

Theorem 1 *a.e.* $D_{4-in,5-out}$ *is Hamiltonian.*

□

To prove the theorem we will regard $D_{4-in,5-out}$ as the union of independent random digraphs $D_1 \cup D_2 \cup D_3 \cup D_4$. Here $D_1 \in \mathcal{D}_{2-in,2-out}$, $D_2 \in \mathcal{D}_{2-out}$, $D_3 \in \mathcal{D}_{2-in}$ and $D_4 \in \mathcal{D}_{1-out}$.

This result is unlikely to be best possible and we conjecture that a.e. $D_{2-in,2-out}$ is Hamiltonian.

We will use a *three phase* method as outlined below: a *cycle decomposition* is a set of vertex disjoint directed cycles that cover all n vertices. The *size* of the decomposition is the number of cycles.

Phase 1. We show that a.e. D_1 contains a directed cycle decomposition of size at most $2 \log n$.

Phase 2. Using $D_2 \cup D_3 \cup D_4$ we increase the minimum cycle size in the cycle decomposition to $\frac{20n}{\log n}$. This is done by growing each small cycle as a path in a way which does not allow the formation of any new cycle of size less than $\frac{20n}{\log n}$. We then close the path to a cycle of the required size.

Phase 3. Using $D_2 \cup D_4$ we convert the *Phase 2* cycle decomposition to a Hamilton cycle. To do this we break the cycles of the cycle decomposition and rearrange the path sections so formed into a Hamilton cycle. The manner of breaking and rearrangement is restricted in a way which allows us to use the second moment method to count the number of Hamilton cycles.

2 Phase 1. Making a cycle decomposition with at most $2\log n$ cycles

With any digraph D on n vertices there is an associated bipartite graph G with $n+n$ vertices which contains an edge (u,v) iff D contains the directed edge (u,v). It is well known that perfect matchings in G are in 1-1 correspondence with cycle decompositions of D.

We start with the random digraph D_1.

Lemma 2 *a.e. D_1 contains a cycle decomposition with at most $2\log n$ cycles.*

Proof. Walkup [7] has shown that the bipartite graph associated with D_1 a.s. contains a perfect matching $\{(i,\phi(i)), i=1,2,\ldots,n\}$. We can argue by symmetry that we can take ϕ to be a random permutation. It is well known e.g. Feller [2] that a.e. permutation contains at most $2\log n$ cycles. and thus the cycle decomposition has size at most $2\log n$. □

Thus at the end of Phase 1 we can assume we have a cycle decomposition of size at most $2\log n$.

3 Phase 2. Removing cycles of size $\frac{20n}{\log n}$ or less from the cycle decomposition

We partition the cycle decomposition into sets SMALL and LARGE, containing cycles C of size $|C| < \frac{20n}{\log n}$ and $|C| \geq \frac{20n}{\log n}$ respectively. We now describe an algorithm to replace all small cycles by large ones. Essentially the algorithm works as follows. We break an arbitrary edge (v_0,u_0) on a given $C \in$ SMALL. Using the 2-out digraph D_2 we grow paths from v_0 using the edges of D_2. These edges either attach the path to another cycle of the cycle decomposition, or the path intersects itself producing a cycle plus a new path. In our growth process, we only allow intersections which produce a cycle and a path both of size at least $\frac{20n}{\log n}$. This continues until we have $m=\sqrt{n\log n}$ paths u_0Pv_i $(i=1,\ldots,m)$ with distinct endpoints v_i. We now fix v_i and repeat the growth process for u_0 using the 2-in digraph D_3. The use of the independent 1-in digraph D_4 allows us to a.s. successfully close at least one of these paths to a cycle for each $C \in$ SMALL.

We now formally describe an algorithm (*OutPhase 2*) used to exclude the formation of small cycles when growing paths using the 2-out digraph as discussed above.

Algorithm OutPhase 2
For any $C \in$ SMALL we define as OUTPHASE(C) a layered tree of depth at most $1.5\log n$

whose set of nodes S_t at depth t consists of 5-tuples $\sigma(t) = (C, t, v, u_0Pv, \mathcal{D})$ defined inductively as follows.
C is the current cycle of SMALL we are enlarging, t is the depth of the layer, v the current active endpoint, $P = u_0Pv$ the path obtained by making the changes which produce v as an endpoint, $\mathcal{D}$ the current set of cycles which are a cycle decomposition for $V - V(P)$.
The root of the tree is $\sigma(0) = (C, 0, v_0, u_0Pv_0, \mathcal{D}_0 - C)$ where $\mathcal{D}_0$ is the cycle decomposition for the graph after the sucessful merging of the previous small cycle, and we have broken an arbitrary edge (v_0, u_0) of C to give u_0Pv_0.
If $\sigma \in S_t, \sigma = (C, t, v, uPv, \mathcal{D})$, the *potential* descendents σ' of σ, $\sigma' = (C, t+1, x, u_0P'x, \mathcal{D}')$ are formed at iteration $t+1$ as follows.

Let w be the terminal vertex of an out edge of v in the independent 2-out digraph D_2.
Case 1. w is a vertex of a cycle $C' \in \mathcal{D}$ with edge $(x, w) \in C'$. Let wQx be $C' - (x, w)$, and let $uP'x$ be $uPv \cup (v, w) \cup wQx$, and $\mathcal{D}' = \mathcal{D} - C'$.
Case 2. w is a vertex of uPv. Either $w = u$, or (x, w) is an edge of P, in which case $uP'x$ is made by breaking (x, w). Note that $\mathcal{D}' = \mathcal{D} + (wPv \cup (v, w))$ in either case.

In fact we only admit to S_{t+1} those σ' which satisfy the following conditions.
(i) In *Case 2* above, the cycle formed must have at least $\frac{20n}{\log n}$ vertices, and the path formed must either be empty or have at least $\frac{20n}{\log n}$ vertices.
(ii) $x \notin$ OUTEND, where OUTEND is the set of vertices whose out edges have been examined in this or some previous OUTPHASE. If σ' is admitted to level $t+1$, then OUTEND is updated to include x.
(iii) To simplify the discussion we do not allow w to be a vertex of a cycle currently in SMALL during the OUTPHASE.

An edge (v, w) which satisfies the above conditions is described as *sucessful*

In order to remove any ambiguity, we imagine the vertices v of $\sigma(t) \in S_t$ examined in ascending label order for the construction of S_{t+1}.

We note that the expected number of vertices on cycles of size at most $\frac{20n}{\log n}$ is $\frac{20n}{\log n}$ (see e.g. Kolchin [5]) and so we may use $\frac{n \log\log n}{2 \log n}$ as an upper bound on the size $|V(\text{SMALL})|$ of the vertex set of cycles in SMALL.

Lemma 3 *For given $C \in$ SMALL, and with probability of the converse $O\left(\left(\frac{\log\log n}{\log n}\right)^2\right)$, there exists a $t^* < \left\lceil \frac{\log\sqrt{n\log n}}{\log 3/2} \right\rceil$ such that $|S_{t^*}| \geq \sqrt{n \log n}$.*

Proof. We assume we stop OUTPHASE(C) when $|S_t| = \sqrt{n\log n}$, and show inductively that a.s. $(\frac{3}{2})^t \leq |S_t| \leq 2^t$, for $t \geq 3$. Thus |OUTEND| is at most |SMALL| $\times\, 2^{t^*} \leq n^{0.86}$. In general, let X_t be the number of unsucessful edges at iteration t, $(t = 1, 2, \ldots, t^*)$. The event of a particular out edge being unsucessful is stochastically dominated by a Bernouilli trial with

$$p = \frac{\frac{40n}{\log n} + |\text{OUTEND}| + |V(\text{SMALL})|}{n}$$

and thus $p < \dfrac{\log\log n}{\log n}$.
For $t \leq c$, constant, the probability of 2 or more unsucessful edges in levels $t \leq c$ is $O\left(\frac{2^{2c}(\log\log n)^2}{(\log n)^2}\right)$ and thus $|S_{t+1}| > 2|S_t| - 1 > (\frac{3}{2})^t, t \geq 3$.
In order to see this, note that in the case where there is only one sucessful edge at the first iteration, subsequent levels expand by a power of 2, and $|S_1| = 2$ otherwise.
For $t > c, c$ large, the expected number of unsucessful edges at iteration t is at most $\mu = 2p|S_t|$ and thus

$$Pr\left(X_t > \lfloor |S_t|/2 \rfloor\right) \leq \left(\frac{2e\log\log n}{\log n}\right)^{\lfloor |S_t|/2 \rfloor}.$$

□

After OUTPHASE(C) we have nodes $(C, t^*, v_i, u_0Pv_i, \mathcal{D}_i) \in S_{t^*}$, for $i = 1, ..., m$ ($m = \sqrt{n\log n}$), each with a path u_0Pv_i of length at least $\frac{20n}{\log n}$, (unless we have already sucessfully made a cycle) plus a cycle decompostion $\mathcal{D}$ of $V \setminus V(u_0Pv_i)$. We now carry out INPHASE(C, v_i) for each i. We start with u_0Pv_i and $\mathcal{D}_i$ and using the 2-in digraph D_3 we build a layered tree similar in description to one made by *Algorithm OutPhase 2.* Here all paths generated end with v_i.

Lemma 4 *With probability of the converse $O\left(\frac{(\log\log n)^3}{\log n}\right)$, a cycle decomposition with minimal cycle length $\frac{20n}{\log n}$ is produced in Phase 2.*

Proof. Describe a path u_0Pv_i as *bad,* if INPHASE(C, v_i) fails to generate $\sqrt{n\log n}$ paths wPv_i from a vertex w to v_i. By arguments similar to the previous lemma,

$$\Pr(u_0Pv_i \text{ is bad }) = O\left(\left(\frac{\log\log n}{\log n}\right)^2\right).$$

Thus

$$Pr\left(\text{the number of bad paths } \geq \frac{\sqrt{n\log n}}{\log\log n}\right) = O\left(\frac{(\log\log n)^3}{(\log n)^2}\right).$$

Hence

$$\Pr(\exists C \in \text{ SMALL with more than } \frac{\sqrt{n\log n}}{\log\log n} \text{ bad paths }) = O\left(\frac{(\log\log n)^3}{\log n}\right).$$

Now, let $m' = m\left(1 - \frac{1}{\log\log n}\right)$. Adding the independent copy D_4 of 1-out, we see that

$$\Pr(\nexists\ wPv_i \text{ s.t. } \mathrm{out}_4(v_i) = w) \leq \left(1 - \sqrt{\frac{\log n}{n}}\right)^{m'} = O\left(\frac{1}{n}\right).$$

$\square$

At this stage we have shown that a 4-in,5-out digraph almost always contains a cycle decomposition J in which the minimum cycle size is at least $\frac{20n}{\log n}$.

We shall refer to J as the *Phase 2* cycle decomposition.

Also let A denote the union of the sets OUTEND created as we removed each small cycle. Thus we know that $|A| \leq n^{9/10}$ a.s. . Furthermore, if $v \notin A$ then both the in-edges of D_2 and the out-edge of D_4 incident with v are unexamined and hence unconditioned.

4 Phase 3. Patching the Phase 2 cycle decomposition to a Hamilton cycle

Let $C_1, C_2, ..., C_k$ be the cycles of J, and let $c_i = |C_i \setminus A|$, $c_1 \leq c_2 \leq \cdots \leq c_k$, and $c_1 \geq \frac{20n}{\log n} - n^{-9/10}$. Let $a = \frac{n}{\log n}$. For each C_i we consider selecting a set of $m_i = 2\lfloor \frac{c_i}{a} \rfloor + 1$ vertices $v \notin A$, and deleting the edge (v, u) in J. Let $m = \sum_{i=1}^{k} m_i$ and relabel (temporarily) the broken edges as $(v_i, u_i), i \in [m]$ as follows: in cycle C_i identify the lowest numbered vertex x_i which loses a cycle edge directed out of it. Put $v_1 = x_1$ and then go round C_1 defining $v_2, v_3, \ldots v_{m_1}$ in order. Then let $v_{m_1+1} = x_2$ and so on. We thus have m path sections $u_i P v_j$ in J. If $P = u_i P v_j$ is such a section, define ϕ by $\phi(j) = i$, $(j = 1, ..., m)$. We see that ϕ is an even permutation as all the cycles of ϕ are of odd length.

We wish to try rejoin these path sections of J to make a Hamilton cycle using $D_2 \cup D_4$. Suppose we can. We define a permutation ρ where $\rho(i) = j$ if $u_{\phi(i)} P v_i$ is joined to $u_{\phi(j)} P v_j$ by $(v_i, u_{\phi(j)})$. This also defines a permutation γ where $\gamma(i) = \phi(j)$ and hence $\gamma(i) = \phi(\rho(i))$. Let H_m be the set of cyclic permutations on $[m]$. Let $R_\phi = \{\rho \in H_m : \phi\rho = \gamma, \gamma \in H_m\}$ be the *cyclic* solutions to $\gamma = \phi\rho$.
Thus we have not only constructed a Hamilton cycle in $J \cup D_2 \cup D_4$, but also in the *auxillary digraph* Γ, whose edges are $(i, \gamma(i))$.

Lemma 5 $(m-2)! \leq |R_\phi| \leq (m-1)!$

Proof. We grow a path $1, \gamma(1), \gamma^2(1), ..., \gamma^k(1)$ in Γ, maintaining feasibility in the way we join the path sections of J at the same time.

We note that at vertex i of Γ, an out edge corresponds to an edge from v_i in $u_{\phi(i)}Pv_i$; and an in edge to an edge to u_i in $u_iPv_{\phi^{-1}(i)}$. On adding the edge $(1,\gamma(1))$ we must avoid an edge to $\phi(1)$ (i.e. to $u_{\phi(1)}$ in J) and also an edge to 1 (i.e. joining v_1 to u_1). Thus there are $m-2$ choices for $\gamma(1)$ since $\phi(1) \neq 1$.

In general, at vertex $\gamma^k(1)$, $(k = 0, 1, ..., m-3)$, on adding the edge $(\gamma^k(1), \gamma^{k+1}(1))$, the subscripts $\gamma(1), ..., \gamma^k(1)$ of u are already used. We must also avoid the subscripts 1 and ℓ where u_ℓ is the initial vertex of the path terminating at $v_{\gamma^k(1)}$ made by joining path sections of J. Thus there are either $m-(k+1)$ or $m-(k+2)$ choices for $\gamma^{k+1}(1)$ depending on whether or not $\ell = 1$.

Hence, when $k = m-3$, there *may* be only one choice for $\gamma^{m-2}(1)$, the vertex h say. After adding this edge, let the remaining isolated vertex of Γ be w. We now need to show that we can complete γ, ρ so that $\gamma, \rho \in H_m$.

Which vertices are missing edges in Γ at this stage ? Vertices $1, w$ are missing in edges, and h, w out edges. Hence the path sections of J are joined so that either

$$u_1 \to v_h, \quad u_w \to v_w \quad \text{or} \quad u_1 \to v_w, \quad u_w \to v_h.$$

The first case can be (uniquely) feasibly completed in both Γ and J by setting $\gamma(h) = w, \gamma(w) = 1$. Completing the second case to a cycle in J forces

$$\gamma = (1, \gamma(1), ..., \gamma^{m-2}(1))(w) \tag{1}$$

and thus $\gamma \notin H_m$. We show this case cannot arise.

$\gamma = \phi\rho$ and ϕ even implies that γ and ρ have the same parity. On the other hand $\rho \in H_m$ has a different parity to γ in (1) - contradiction.

Thus there is a (unique) completion of the path in Γ. □

Let H stand for the cycle decomposition J to which $D_2 \cup D_4$ has beeen added.

Lemma 6 *Pr(H does not contain a Hamilton cycle)* $= O(n^{-0.04})$.

Proof. Let X be the number of Hamilton cycles in H resulting from rearranging the path sections generated by ϕ according to those $\rho \in R_\phi$. We will show that $E(X) \to \infty$ and

$$\text{Var}(X) \leq E(X) + E(X)^2 O(n^{-0.3})$$

and thus we may use the second moment method.

Let Ω denote the set of possible cycle re-arrangements.

$$E(X) \;=\; \sum_{\Omega} \left(1 - \left(1 - \frac{1}{n}\right)^3\right)^m$$

$$\geq \left(\frac{3}{n}\left(1-O\left(\frac{1}{n}\right)\right)\right)^m \prod_{i=1}^{k}\binom{c_i}{m_i}(m-2)!$$
$$\geq \frac{1}{m\sqrt{m}}\left(\frac{3m}{en}\right)^m \prod_{i=1}^{k}\left(\frac{c_i}{m_i}\right)^{m_i}.$$

However, $m_i = 2\lfloor \frac{c_i}{a} \rfloor + 1$ so $\frac{c_i}{m_i} \geq \left(\frac{20}{41} - o(1)\right)a$, and thus

$$\prod_{i=1}^{k}\left(\frac{c_i}{m_i}\right)^{m_i} \geq \left(\left(\frac{20}{41} - o(1)\right)a\right)^m.$$

Hence

$$E(X) \geq \frac{1}{m\sqrt{m}}\left(\left(\frac{20}{41} - o(1)\right)\frac{3ma}{en}\right)^m.$$

But $\left(\frac{39}{20} - o(1)\right)\log n \leq m \leq \frac{41}{20}\log n$ and so we have $E(X) \geq n^{0.046}$.

Let M, M' be two sets of selected edges which have been deleted in J and whose path sections have been rearranged into Hamilton cycles according to ρ, ρ' respectively. Let N, N' be the corresponding sets of edges which have been added to make the Hamilton cycles. What is the interaction between these two Hamilton cycles?

Let $s = |M \cap M'|$ and $t = |N \cap N'|$. Now $t \leq s$ since if $(v, u) \in N \cap N'$ then there must be a unique $(\tilde{v}, u) \in M \cap M'$ which is the unique J-edge into u. We claim that $t = s$ implies $t = s = m$ and $(M, \rho) = (M', \rho')$. (This is why we have restricted our attention to $\rho \in R_\phi$.) Suppose then that $t = s$ and $(v_i, u_i) \in M \cap M'$. Now the edge $(v_i, u_{\gamma(i)}) \in N$ and since $t = s$ this edge must also be in N'. But this implies that $(v_{\gamma(i)}, u_{\gamma(i)}) \in M'$ and hence in $M \cap M'$. Repeating the argument we see that $(v_{\gamma^k(i)}, u_{\gamma^k(i)}) \in M \cap M'$ for all $k \geq 0$. But γ is cyclic and so our claim follows.

We adopt the following notation. Let $t = 0$ denote the event that no common edges occur, and (s,t) denote $|M \cap M'| = s$ and $|N \cap N'| = t$.

$$\begin{aligned} E(X^2) &\leq E(X) + \sum_{\Omega}\left(\frac{3}{n}\right)^m \sum_{\substack{\Omega \\ t=0}}\left(\frac{3}{n}\right)^m \\ &\quad + \sum_{\Omega}\left(\frac{3}{n}\right)^m \sum_{s=2}^{m}\sum_{t=1}^{s-1}\sum_{\substack{\Omega \\ (s,t)}}\left(\frac{3}{n}\right)^{m-t} \\ &= E(X) + E_1 + E_2 \text{ say.} \end{aligned}$$

Clearly $E_1 \leq E(X)^2$. For given ρ, how many ρ' satisfy the condition (s,t)? Previously $|R_\phi| \geq (m-2)!$ and now $|R_\phi(s,t)| \leq (m-t-1)!$, (consider fixing t edges of Γ').
Thus

$$E_2 \leq E(X)^2 \sum_{s=2}^{m}\sum_{t=1}^{s-1}\left[\sum_{\substack{(\sigma_1,\ldots,\sigma_k) \\ s=\sigma_1+\cdots+\sigma_k}} \prod_{i=1}^{k}\frac{\binom{m_i}{\sigma_i}\binom{c_i-m_i}{m_i-\sigma_i}}{\binom{c_i}{m_i}}\right]\frac{(m-t-1)!}{(m-2)!}\left(\frac{n}{3}\right)^t.$$

Now

$$\frac{\binom{c_i-m_i}{m_i-\sigma_i}}{\binom{c_i}{m_i}} \leq \left(\frac{m_i}{c_i}\right)^{\sigma_i}\left(1+O\left(\frac{m^2}{c_1}\right)\right)$$

and

$$\frac{m_i}{c_i} \leq \frac{2}{a}+\frac{1}{c_1} \leq \frac{21}{10a}.$$

Also $\binom{m}{s}\frac{(m-t-1)!}{(m-2)!} \leq \frac{m^{s-(t-1)}}{s!}$ and so

$$\begin{aligned}
\frac{E_2}{E(X)^2} &\leq (1+o(1))\sum_{s=2}^{m}\frac{m^s}{s!}\left(\frac{21}{10a}\right)^s m\sum_{t=1}^{s-1}\left(\frac{n}{3m}\right)^t \\
&= (1+o(1))\frac{3m^2}{n}\sum_{s=2}^{m}\frac{1}{s!}\left(\frac{7n}{10a}\right)^s \\
&\leq \frac{3m^2}{n}n^{7/10} \\
&= O(n^{-1/4}).
\end{aligned}$$

The result follows by the Chebychev inequality.□

References

[1] C.Cooper and A.M.Frieze, *The limiting probability that* $\alpha - in, \beta - out$ *is strongly is connected*, JCT B 48 (1990) 117-134.

[2] W.Feller, *An introduction to probability theory and its applications* Vol. 1, John Wiley and Sons, New York, 1966.

[3] T.I.Fenner and A.M.Frieze, *On the connectivity of random m-orientable graphs and digraphs*, Combinatorica 2 (1982) 347-359.

[4] A.M.Frieze and T.Luczak, *Hamiltonian cycles in a class of random graphs: one step further* in Proceedings of Random Graphs '87, Edited by M.Karonski, J.Jaworski

[5] V.F.Kolchin, *Random mappings*, Optimization Software Inc., New York,1986.

[6] C.J.H.McDiarmid and B.Reed, to appear.

[7] D.D.W.Walkup, *Matchings in random regular bipartite digraphs*, Discrete mathematics 31 (1980) 59-64.