JOURNAL OF ALGORITHMS 9, 181–204 (1988)

An Algorithm for Finding Hamilton Cycles in Random Directed Graphs

A. M. FRIEZE

Department of Computer Science and Statistics, Queen Mary College, Mile End Road, London E1 4NS, England

Received October 18, 1986

We describe a polynomial ($O(n^{1.5})$) time algorithm DHAM for finding hamilton cycles in digraphs. For digraphs chosen uniformly at random from the set of digraphs with vertex set $\{1,2,\ldots,n\}$ and $m = m(n)$ edges the limiting probability (as $n \to \infty$) that DHAM finds a hamilton cycle equals the limiting probability that the digraph is hamiltonian. Some applications to random "travelling salesman problems" are discussed.

1. INTRODUCTION

Some of the main problems in the study of hamilton cycles in random (undirected) graphs have been solved in recent years. For example Komlós and Szemerédi [13] showed that if $m = \frac{1}{2}n\log n + \frac{1}{2}n\log\log n + c_n n$ and $G_{n,m}$ denotes the random graph sampled uniformly from the set of graphs with vertex set $V_n = \{1,2,\ldots,n\}$ and m edges, then

$$\lim_{n\to\infty} \Pr(G_{n,m} \text{ is hamiltonian}) = \lim_{n\to\infty} \Pr(\delta(G_{n,m}) \geq 2) = \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-e^{-2c}} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to \infty, \end{cases} \tag{1.1}$$

where $\delta(G)$ is the minimum degree of graph G.

Bollobas [3] strengthened this result in the following way: let $e_1, e_2 \ldots e_N$, where $N = \binom{n}{2}$ be a random permutation of the edges of the complete graph K_n. Let $G_m = (V_n, \{e_1, e_2, \ldots, e_m\})$ and $m^* = \min\{m\colon \delta(G_m) \geq 2\}$.

0196-6774/88 $3.00

Bollobas showed that

$$\lim_{n\to\infty} \Pr(G_{m^*} \text{ is hamiltonian}) = 1. \tag{1.2}$$

Subsequently Bollobas, Fenner, and Frieze [5] described an algorithm HAM which can be implemented in $O(n^{3+o(1)})$ time and satisfies

$$\lim_{n\to\infty} \Pr(\text{HAM finds a hamilton cycle in } G_{n,m}) = \lim_{n\to\infty} \Pr(G_{n,m} \text{ is hamiltonian}). \tag{1.3}$$

There are now a large number of related results, many of which can be found in Bollobas [4, Chap. VIII].

When we come to digraphs we find that previous research has left considerable gaps in our knowledge. The main result of this paper provides the analogs of (1.1) and (1.3) for digraphs. Let $D_{n,m}$ denote the digraph sampled uniformly from the set of digraphs with vertex set V_n and m edges. Suppose now $m = n \log n + c_n n$.

Let $\delta^+(D), \delta^-(D)$ denote the minimum outdegree and indegree of a digraph D. We prove

THEOREM 1. *There is a (randomised) polynomial $(O(n^{1.5}))$ time algorithm DHAM which satisfies*

$$\lim_{n\to\infty} \Pr(\text{DHAM } \textit{finds a hamilton cycle in } D_{n,m}) = \lim_{n\to\infty} \Pr(\min\{\delta^+(D_{n,m}), \delta^-(D_{n,m})\} \geq 1) = \begin{cases} 0 & \text{if } c_n \to -\infty \\ e^{-2e^{-c}} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to \infty. \end{cases}$$

The previous best **existence** result is due to McDiarmid [14] who gave a nonconstructive proof that, in the notation of Theorem 1, if $c_n - \log\log n \to \infty$ then

$$\lim_{n\to\infty} \Pr(D_{n,m} \text{ is hamiltonian}) = 1.$$

He proved this by showing that for any $0 \leq p \leq 1$,

$$\Pr(D_{n,p} \text{ is hamiltonian}) \geq \Pr(G_{n,p} \text{ is hamiltonian}),$$

where $D_{n,p}$ (resp. $G_{n,p}$) is the random digraph (resp. graph) with vertex set V_n in which each of the $2N$ (resp. N) possible edges occur independently with probability p.

The best previous result concerning polynomial time algorithms is due to Angluin and Valiant [2] who described an $O(n \log n)$ time algorithm A and showed

$$\lim_{n \to \infty} \Pr(A \text{ finds a hamilton cycle in } D_{n,p}) = 1,$$

assuming $p = c \log n/n$ and c is a sufficiently large constant.

Instead of proving Theorem 1 directly we shall instead prove a stronger result, the directed constructive equivalent of (1.2), from which the theorem follows.

Let $e_1, e_2, \ldots e_{n(n-1)}$ be a random permutation of the edges of the complete digraph DK_n with vertex set V_n. Let $E_m = \{e_1, e_2, \ldots, e_m\}$ for $1 \leq m \leq n(n-1)$ and $D_m = (V_n, E_m)$.

THEOREM 2. *Let* $m^* = \min\{m\colon \delta^+(D_m) \geq 1, \delta^-(D_m) \geq 1\}$. *Then*

$$\lim_{n \to \infty} \Pr(\text{DHAM } \textit{finds a hamilton cycle in } D_{m^*}) = 1.$$

(Note that $\lim_{n \to \infty} \Pr(\text{HAM finds a hamilton cycle in } G_{m^*}) = 1$ was proved in [5].) The proof generalises easily to the case where we want k edge disjoint cycles.

THEOREM 3. *Let* $k \geq 1$ *be an integer constant and let* $m_k^* = \min\{m\colon \delta^+(D_m) \geq k, \delta^-(D_m) \geq k\}$. *Then there exists an* $O(n^{1.5})$ *time algorithm* DHAM_k *satisfying*

$$\lim_{n \to \infty} \Pr(\text{DHAM}_k \textit{ finds } k \textit{ edge disjoint hamilton cycles in } D_{m_k^*}) = 1.$$

We will also consider the use of DHAM in the exact solution of random travelling salesman problems. The first application is to the Bottleneck Travelling Salesman Problem (BTSP). An instance of BTSP is specified by the assignment of a numerical weight to the edges of DK_n. The objective is to find a hamilton circuit for which the maximum edge-weight is minimised.

Let us assume that edge-weights are drawn independently from the uniform $[0, 1]$ distribution. As a simple corollary of Theorem 2 we prove

THEOREM 4. *There is a polynomial* $(O(n^2))$ *time algorithm DBOT satisfying*

$$\lim_{n \to \infty} \Pr(\textit{DBOT solves BTSP exactly}) = 1.$$

We can use ideas from this paper and Frieze [9] to tackle the more usual travelling salesman problem (DTSP) in which we seek the hamilton circuit of minimum total weight. We change the distribution so that the edges of

DK_n are independently given integer weights from $\{0, 1, \ldots, B(n) - 1\}$. We can then prove

THEOREM 5. *If $B(n) = O(n/\log n)$ then there is a polynomial $(O(n^2))$ time algorithm DTSPSOLVE satisfying*

$$\lim_{n \to \infty} \Pr(\textit{DTSPSOLVE solves DTSP exactly}) = 1.$$

(In order not to make the paper too long we will only give an outline proof of this result, but we hope it will suffice to convince the reader.)

2. Some Notation and Preliminaries

For convenience we give here some definitions and basic results that are used throughout the paper.

$E(X)$ denotes the edge set of X, where X can be a graph, digraph, or cycle and $V(X)$ denotes its vertex set.

$B(a, p)$ denotes the binomial random variable with parameters a and p and

$$BS(b, c; a, p) = \Pr(b \le B(a, p) \le c).$$

The following inequalities for the tails of the binomial distribution are invaluable and can be derived, for example, from Theorem 1 of Hoeffding [10]:

$$BS(0, (1 - \varepsilon)ap; a, p) \le e^{-\varepsilon^2 ap/2}, \qquad 0 < \varepsilon < 1, \tag{2.1a}$$

$$BS((1 + \varepsilon)ap, \infty; a, p) \le e^{-\varepsilon^2 ap/3}, \qquad 0 < \varepsilon < 1. \tag{2.1b}$$

Let Q_n be a sequence of events. We say that Q_n occurs almost always (a.a.) if $\lim_{n \to \infty} \Pr(Q_n) = 1$. It is also useful to allow a.a. to stand for "almost all", with the obvious meaning.

If $m \approx \alpha n \log n$ for some constant $\alpha > 0$ and $p = m/n(n-1)$ then it is convenient to derive properties of $D_{n,m}$ from $D_{n,p}$. For a property Q

$$\Pr(D_{N,p} \text{ has } Q) = \sum_{m=0}^{n(n-1)} \Pr(D'_{n,m} \text{ has } Q) \Pr\left(\left|E(D_{n,p})\right| = m'\right) \tag{2.2}$$

as $D_{n,p} | m'$ edges is distributed as $D_{n,m'}$. We deduce from (2.2) that

$$\Pr(D_{n,m} \text{ has } Q) \le (1 + o(1))\sqrt{2\pi\alpha n \log n}\, \Pr(D_{n,p} \text{ has } Q) \tag{2.3}$$

and that if $\{Q_n\}$ are monotone properties, i.e., are preserved by adding edges or are preserved by deleting edges then

$$D_{n,p} \text{ has } Q_n \text{ a.a. implies } D_{n,m} \text{ has } Q_n \text{ a.a.} \tag{2.4}$$

The following quantities are defined later as well but are given here for quick reference:

$$m_0 = \lfloor n \log n - n \log\log\log n \rfloor, \qquad m_1 = \lfloor n \log n + n \log\log\log n \rfloor,$$
$$m_2 = \lceil 5n \log n/6 \rceil, \qquad \text{and} \qquad m_3 = \lceil 2n \log n/3 \rceil.$$
$$p_i = m_i/n(n-1) \qquad \text{for } i = 0,1,2,3.$$
$$l_o = \left\lceil \frac{3\log n}{\log\log n} \right\rceil, \ \mu_1 = \left\lceil \frac{(\log n)^2}{\log\log\log n} \right\rceil, \text{ and } \mu_2 = \left\lfloor \frac{\log n}{\log\log\log\log n} \right\rfloor.$$

3. Overview of the Algorithm

We assume that the input to DHAM consists of the edges ED_n in random order $e_1, e_2, \ldots, e_{n(n-1)}$. However, only the edges $e_1, e_2, \ldots, e_{m^*}$ are available to DHAM.

In the undirected case one tries to extend a current path and use rotations (see Fig. 1) to create extra paths. This approach fails in digraphs, for obvious reasons.

The ideas behind DHAM derive from the patching algorithms of Karp [11], Karp and Steele [12], and, more recently, Dyer and Frieze [6], for finding approximate solutions to travelling salesman problems.

The algorithm is split into 3 phases. In Phase 1 we construct a small set of edges $E^1 \subseteq E_{m^*}$ such that if $D^1 = (V_n, E^1)$ then $\delta^+(D^1) \geq 1$ and $\delta^-(D^1) \geq 1$ and D^1 a.a. contains a set of about $\log n$ vertex disjoint cycles covering V_n.

In Phase 2 we try to "patch" the cycles together in pairs by 2-edge exchanges (see Fig. 2). We will see that at the end of this phase there is a.a. a cycle $\tilde{C}$ of size $n - o(n)$ plus a few ($O(\log n)$) others.

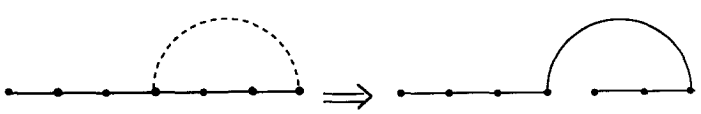

FIGURE 1

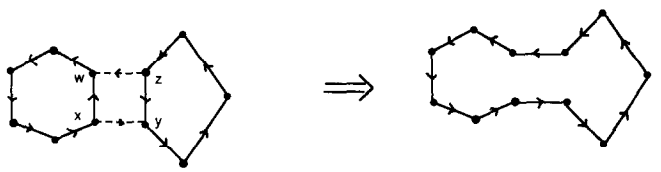

FIGURE 2

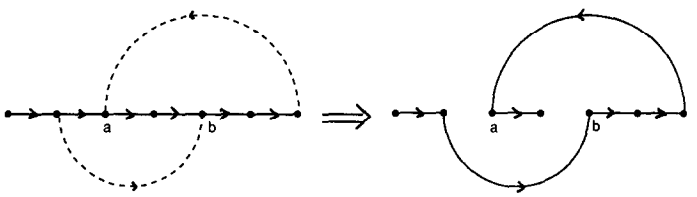

FIGURE 3

In the final phase we try to patch the small cycles one by one into $\tilde{C}$ by a more complex process of "double-rotations" (see Fig. 3).

4. Details of Phase 1

PHASE 1 OF DHAM.

begin
$\sigma :=$ a random permutation of V_n; {used to avoid bias in the set of cycles produced {by Phase 1
if $e = (v, w)$ let $\hat{e}$ denote $(v, \sigma(w))$;
$\hat{E}^{1+} := \hat{E}^{1-} := \varnothing$;
{construct a set $\hat{E}^{1+}$ of about $10n$ edges for which $\delta^+(\hat{E}^{1+}) \geq 1$ and the average out-degree is ≈ 10}
A:**for** $i = 1$ **to** m^* **do**
begin
if $d^+(v_i, \hat{E}^{1+}) \leq 9$ **then** $\hat{E}^{1+} := \hat{E}^{1+} \cup \{\hat{e}_i\}$;
{where $e_i = (v_i, w_i)$ and $d^+(v, E)$ (resp. $d^-(v, E)$) is the number of edges in E with tail (resp. head) v.}
end;

{construct a set $\hat{E}^{1-}$ of about $10n$ edges, disjoint from $\hat{E}^{1+}$, for which $\delta^-(\hat{E}^{1-}) \geq 1$ and the average indegree is ≈ 10}

B: **for** $i = 1$ **to** m^* **do**
begin
if $e_i \notin \hat{E}^{1+}$ **and** $d^-(\sigma(w_i), \hat{E}^{1-}) \leq 9$ **then** $\hat{E}^{1-} := \hat{E}^{1-} \cup \{\hat{e}_i\}$
end
$\hat{E} := \hat{E}^{1+} \cup \hat{E}^{1-}$;
BIP := bipartite graph (V_n, W_n, A) where W_n is a disjoint copy of V_n and $\{v, w\} \in A$ iff (v, w) or $(w, v) \in \hat{E}^1$;

{BIP is close to the model of Walkup [15] where each vertex randomly chooses 10 neighbours. 10 is sufficiently large so that it is easy to show it a.a. contains a perfect matching, 2 in fact should suffice.}

apply the algorithm of Even and Tarjan [8] to find a maximum matching M of BIP;
{This algorithm is not "aware" that the edges of E_{m^*} have been relabelled.}
if M is not a perfect matching **then** DHAM has failed **else**
begin
define the permutation ψ of V_n by $M = \{(v, \psi(v)): v \in V_n\}$;
$\phi := \sigma^{-1}\psi$;
output $F = \{(v, \phi(v)): v \in V_n\}$
{F defines a set of vertex disjoint cycles covering V_n}
end
end

Before proving that Phase 1 is a.a. successful we discuss some preliminary results. Let $m_0 = \lfloor n \log n - n \log\log\log n \rfloor$ and $m_1 = \lceil n \log n + n \log\log\log n \rceil$. It is straightforward to show

$$\delta^+(D_{m_0}) = \delta^-(D_{m_0}) = 0 \qquad \text{and} \qquad \delta^+(D_{m_1}) = \delta^-(D_{m_1}) = 1$$

a.a. (Erdos and Renyi [7])

and hence

$$m_0 \leq m^* \leq m_1 \qquad \text{a.a.}$$

Next let

$$\text{TINY} = \left\{ v \in V_n: d^+(v, E_{m_0}) \leq l_0 = \left\lceil \frac{3 \log n}{\log\log n} \right\rceil \text{or } d^-(v, E_{m_0}) \leq l_0 \right\}.$$

LEMMA 4.1.

$$v, w \in \textit{TINY implies } \operatorname{dist}(v, w; GD_{m^*}) \geq 5 \qquad \textit{a.a.} \qquad (4.1)$$

where if D is a digraph then GD is the graph obtained by ignoring edge orientation and parallel edges in D and $\operatorname{dist}(v, w; G)$ *is the distance from vertex v to vertex w in graph G.*

Proof.

$$\Delta\left(GD_{n,m_0}\right) \le 6\log n \qquad \text{a.a. (actually with probability } 1 - o\left(n^{-1}\right)\text{)}, \tag{4.2}$$

where Δ denotes maximum degree. (The expected number of vertices of degree exceeding $6\log n$ in GD_{n,p_0} is at most n BS $(6\log n, n-1; n-1, 2p_0) = o(n^{-2})$; now apply (2.3).)

Here and throughout the paper we do not aim for best possible interim results, only for ease of calculation:

$$|\text{TINY}| = n^{o(1)} \qquad \text{a.a.} \tag{4.3}$$

(The expected size of TINY in GD_{n,p_0} is n BS$(0, l_0; n-1, p_0) = n^{o(1)}$. The markov inequality ($\Pr(X > a) \le \mathrm{E}(X)/a$ for a nonnegative random variable X, E $\equiv$ Expectation here, naturally) implies that $|\text{TINY}| = n^{o(1)}$ a.a.; apply (2.4).)

We note next that

$$\Pr(\exists \text{ a path of length } \le 4 \text{ in } GD_{n,p_0} \text{ with distinct endpoints } v, w \in \text{TINY})$$

$$\le \binom{n}{2} \sum_{k=0}^{3} \binom{n-2}{k} k! p_0^{k+1} \text{BS}(0, l_0; n-5, 2p_0)^2 = O\left(n^{-(1-o(1))}\right)$$

and so, by (2.3), (4.1) holds in GD_{m_0}. Now GD_{m_1} is obtained from GD_{m_0} by adding $2n\log\log\log n$ (random) edges. Assuming (4.1) holds in GD_{n,m_0} and $\Delta(GD_{n,m_0}) \le 6\log n$ and $|\text{TINY}| = n^{o(1)}$ we find the probability that any of these extra edges is incident with two vertices, both at distance 3 or less from TINY is $O(n^{o(1)})(\log n)^6 n \log\log\log n/n^2) = o(1)$ and the result follows. □

LEMMA 4.2. *The following hold a.a.:*

(a) *Phase* 1 *succeeds.*

(b) ϕ *has at most* $2\log n$ *cycles.*

Proof. (a) Let us first change the problem slightly so that loops are allowed and return later to the loopless case. We now assume that $e_1, e_2, \ldots, e_{n^2}$ is a random permutation of $V_n \times V_n$. We can apply Phase 1 to this sequence as we did above.

Remark 1. The distribution of the sequence $(v_i, \sigma(w_i))$ is independent of σ. (This is not the case if no loops are allowed in the original sequence.)

We must show that BIP a.a. contains a perfect matching. Let us first consider the distribution of $\hat{E}^1$: In the construction of $\hat{E}^{1+}$ each $v \in V_n$ independently chooses $a(v) = \min\{10, d^+(v)\}$ out-neighbours at random

from V_n, where $d^+(v)$ (resp. $d^-(v)$) is the out-degree (resp. in-degree) of v in D_{m^*}. For $w \in V_n$ let $\Gamma^-(w) = \{v \in V_n : (v, w) \in \hat{E}^{1+}\}$. Given $\hat{E}^{1+}$, $\hat{E}^{1-}$ has the following distribution: each $w \in V_n$ independently chooses $b(v) = \min\{10, d^-(w) - |\Gamma^-(w)|\}$ in-neighbours from $V_n - \Gamma^-(w) - \{v: d^+(v) < 10\}$.

Suppose now that BIP does not contain a perfect matching. Then by Hall's theorem there exists $K \subseteq V_n$, $|K| = k$, and $L \subseteq W_n$, $|L| = k - 1$ such that $N(K, \text{BIP}) \subseteq L$ where $N(K, \text{BIP}) = \{y: \exists x \in K$ such that $\{x, y\} \in E(\text{BIP})\}$. Now $k \neq 1$ or n since δ (BIP) ≥ 1 by construction. Let $K' = K \cap \text{TINY}$ and $k' = |K'|$. Then (4.1) implies (i) $|L'| \geq k'$, where $L' = N(K', \text{BIP})$ and (ii) each $v \in K - K'$ chooses at least 9 out-neighbours in $L - L'$ in the construction of $\hat{E}^{1+}$. But the probability of the occurrence of (ii) with $k \leq n/2$ is at most

$$\sum_{t=9}^{\lfloor n/2 \rfloor} \binom{n}{t}\binom{n}{t-1}\left(10\left(\frac{t-1}{n-1}\right)^9\right)^t \qquad (t \equiv k - k')$$

$$\leq \sum_{t=9}^{\lfloor n/2 \rfloor} \left(\frac{n^2 e^2 10 t^9}{t^2 n^9}\right)^t$$

$$= O(n^{-63}). \tag{4.4}$$

The case $k > n/2$ is equivalent to the existence of a $K \subseteq W_n$, $|K| \leq n/2$ in which $|N(K, B)| < |K|$. We will show that

$$|\Gamma^-(w)| \leq \frac{2 \log n}{\log \log n} \qquad \text{for all } w \in W_n \text{ a.a.} \tag{4.5}$$

We can now argue as before, given (4.5), that this implies each $w \in K - \text{TINY}$ chooses at least 9 in-neighbours from $L - L'$ in the construction of $\hat{E}^{1-}$, where L, L' are defined analogously to the previous case. The probability of this is $O(n^{-63})$ as before.

To prove (4.5) we note that $|\Gamma^-(w)|$ is dominated by $B(10n, 1/n)$ and then a simple calculation yields (4.5).

(b) Let ϕ^* be some fixed permutation of V_n. Then

$$\Pr(\text{Phase 1 outputs } \phi^*) = \sum_{\sigma} \Pr(\psi = \sigma^*\phi^* | \sigma = \sigma^*)\Pr(\sigma = \sigma^*)$$

$$= \frac{1}{n!} \sum_{\sigma} \Pr(\psi = \sigma^*\phi^* | \sigma = \sigma^*).$$

But Remark 1 implies $\Pr(\psi = \psi^* | \sigma = \sigma_1) = \Pr(\psi = \psi^* | \sigma = \sigma_2)$ for any permutations $\psi^*, \sigma_1, \sigma_2$. Thus each ϕ^* is output with the same probability.

Now it is well known (for example, [3, p. 363]) that a random permutation of V_n a.a. has approximately $\log n$ cycles. We conclude that (when loops are present) Phase 1 ends a.a. with a set of $r_1 \leq 2\log n$ (say) vertex disjoint cycles covering V_n. Suppose now that we only consider loopless sequences of edges. This is the same as proceeding as above except that the edges e_i, $i \leq m^*$, cannot be used. We show that

$$\Pr(\text{no loop in } e_1, e_2, \ldots, e_{m^*} \text{ is used in the construction of BIP}) \geq (1 - o(1))e^{-20} \tag{4.6}$$

and then we can infer the result that we actually want.

We have to estimate the probability that no edge $(i, \sigma(i))$ is chosen. But referring to the discussion of the distribution of $\hat{E}^1$ we see that, using (4.5), this probability is at least

$$\left(1 - \frac{10}{n}\right)^n \left(1 - 10\Big/\left(n - \frac{2\log n}{\log\log n}\right)\right)^n + o(1).$$

and (4.6) follows. □

5. Details of Phase 2

We inherit from Phase 1 the permutation ϕ plus the associated set of edges F, which we now denote by ϕ_1, F_1, respectively. Let $m_2 = \lfloor 5n \log n/6 \rfloor$.

Phase 2 of DHAM.

being
$\phi_2 := \phi_1$; $F_2 := F_1$;
A: **for** $i = 1$ **to** m_2 **do**
begin
  if $e_i = (x, y)$ **and** x, y are in different cycles of
  F_2 **and** $e = (\phi_2^{-1}(y), \phi_2(x)) = (z, w) \in E_{m_2}$ **then**
    begin
      $F_2 := (F_2 \cup \{e_i, e\}) - \{(x, w), (z, y)\}$;
      $\phi_2(x)$; $= y$; $\phi_2(z)$; $= w$;
        go to A
    end
  end
    {if no patches are made we arrive here}
end.

Let $C_1, C_2, \dots, C_{r_2}$ denote the cycles defined by F_2, ϕ_2, at the end of Phase 2, where $|C_1| \geq |C_2| \geq \cdots \geq |C_{r_2}|$. We know that $r_2 \leq 2 \log n$ a.a. and the aim now is to show that $|C_1| = n - o(n)$ a.a.

Let now $m_3 = \lceil 2n \log n/3 \rceil$ and LARGE $= \{v \in V_n: d^+(v, E_{m_3}) \geq l_0$ and $d^-(v, E_{m_3}) \geq l_0\}$.

LEMMA 5.1. $|\text{LARGE}| \geq n - n^{.5}$ *a.a.*

Proof. Consider D_{n,p_3},

$$E\big(|V_n - \text{LARGE}| \text{in } D_{n,p_3}\big) \leq 2n\,\text{BS}(0, l_0; n-1, p_3) = O\big(n^{1/3+o(1)}\big).$$

Hence $|V_n - \text{LARGE}| \leq n^{1/2}$ a.a. in D_{n,p_3}. Now apply (2.4). □

The edges in the set $(E_{m_0} - E_{m_3}) \cap$ ELARGE (where ELARGE = LARGE × LARGE $- \{(v,v): v \in \text{LARGE}\}$) are only slightly conditioned by the construction of E^1, assuming that (4.5) holds. In this case we have, given E^1,

> if $m_2 < i \leq m_0$ and $e_i = (v, w)$ where $v, w \in$ LARGE then e_i is equally likely to be any edge in ELARGE which has not occurred previously. (5.1)

Now comes a "preparatory" lemma.

LEMMA 5.2. *The following holds a.a.:*

$$S \subseteq V_n, s_0 = \left\lceil n/\sqrt{\log n} \right\rceil \leq |S| \leq n - s_0 \text{ implies} \tag{5.2}$$
$$\left|\left\{(v,w) \in E_{m_3}: v \in S, w \notin S\right\}\right| \geq |S|(n - |S|)\log n/2n.$$

Proof. We consider the corresponding result for D_{n,p_3}. If S is fixed and $s = |S|$ then the number of edges from S to $V_n - S$ in D_{n,p_3} is distributed as $\text{B}(s(n-s), p_3)$. Now $\Pr(\text{B}(s(n-s), p_3) \leq 3s(n-s)p_3/4) \leq e^{-s(n-s)p_3/32}$ on using (2.1a). Applying (2.3) we find $\Pr((5.2) \text{ fails}) \leq 3\sqrt{n \log n}\, 2^n e^{-s_0(n-s_0)p_3/32} = o(1)$. □

We can now prove

LEMMA 5.3. $|C_1| \geq n_0 = \lceil n - \sqrt{n}/\log n \rceil$ *a.a.*

Proof. Throughout the lemma we condition on the value of $E^1 \cup E_{m_3}$. We assume this is such that $r_1 \leq 2 \log n$, $|\text{LARGE}| \geq n - n^{1/2}$, (4.5) and (5.2) hold, and where each $v \in V_n$ is incident with at most $6 \log n$ edges in $E^1 \cup E_{m_3}$. These conditions have been shown to hold a.a. We shall also assume that $|E^2| \geq n \log n/12$, where $E^2 = \{e_i \in \text{ELARGE}: m_3 < i \leq m_2\}$ (the probability that $e_i \in$ ELARGE for $m_3 < i \leq m_2$ is $1 - o(1)$

regardless of $e_{m_3+1}, \ldots, e_{i-1}$ and given our assumptions). None of our assumption affects (5.1), which we invoke to do our calculations. Partition E^2 into sets $E^2(k)$, $k = 1, 2, \ldots, r_2$ of roughly equal size, at least $n/24$ by assumption, so that if $e_i \in E^2(k)$ and $e_j \in E^2(k+1)$ then $i < j$. Let $a_k = \min\{i: e_i \in E^2(k)\}$.

Let A_k denote the event $\{|C_1| < n_0$ when i first reaches a_k in loop A of Phase 2 and no patch is made for $i = a_k, a_k + 1, \ldots, a_{k+1} - 1\}$—$C_1$ refers to the "current" largest cycle. We show

$$\Pr(A_k) \le e^{-\sqrt{\log n/50}}, \qquad k = 1, 2, \ldots, r_2. \tag{5.3}$$

Hence $\Pr(\cup_{k=1}^{r_2} A_k) = o(1)$ which implies the lemma as at most $r_2 - 1$ patches can be made in Phase 2.

Proof of (5.3). Fix k and suppose that when i first reaches a_k, $|C_1| < n_0$. Then, where $\{C_j\}$ denote the current cycles, there exists a smallest t such that

$$n - n_0 \le \sum_{i=0}^{t} |C_{r_2 - i}| \le n_0.$$

Let $S = \cup_{i=0}^{t} C_{r_2-i}$, $T_0 = \{(v, w) \in E_{m_3}: v \in S, w \notin S\}$, $T_1 = \{(\phi_2^{-1}(w), \phi_2(v)): (v, w) \in T_0\}$ and $T_2 = \{e = (v, w) \in T_1: v, w \in \text{LARGE}\}$. Then (5.2) implies

$$|T_1| = |T_0| \ge n_0(n - n_0)\log n/2n \tag{5.4a}$$

and then

$$|T_2| \ge |T_1| - 12n^{1/2}\log n, \tag{5.4b}$$

as for a given v there are at most $12 \log n$ neighbours of $\phi_2(v)$, $\phi_2^{-1}(v)$. But if A_k occurs then $E^2(k) \cap T_2 = \varnothing$. But then (5.1) implies that

$$\Pr\big(E^2(k) \cap T_2 = \varnothing\big) \le \big(1 - |T_2|/n(n-1)\big)^{|E^2(k)|}.$$

Using (5.4) and $|\mathrm{E}^2(k)| \ge n/24$ gives (5.3). □

6. Details of Phase 3

We a.a. start Phase 3 with cycles $C_1, C_2, \ldots, C_{r_2}$, where $|C_1| = n - o(n)$ and $r_2 \le 2 \log n$.

PHASE 3 OF DHAM.

```
begin
 for i = 2 to r_2 do {merge C_i into (the current) C_1}
 begin
    suppose C_i = (x_1, x_2, ..., x_k, x_{k+1} = x_1) as a sequence of vertices
    for j = 1 to k do
    begin
       FINDCYCLE (C_i, x_i, x_{i+1}, C_1, outcome);
       if outcome = success then goto A
    end {j − loop}
    terminate − failure
A: end {i − loop}
 terminate − success
end;

procedure FINDCYCLE (C_i, x_j, x_{j+1}, C_1, outcome);
begin
  outcome: = failure;
     suppose C_1 = (y_1, y_2, ..., y_p) and the out-neighbours of x_j are
     y_{i_1}, y_{i_2}, ..., y_{i_l}:
     if l = 0 then return else
     begin
      ℘_0: = {(x_{j+1}, x_{j+2}, ..., x_1, ..., x_j, y_{i_t}, y_{i_t+1}, ..., y_{i_t-1}):
      t = 1, 2, ..., l}
      = initial set of paths (see Fig. 4)
      END_0 := {y_{i_t-1}: t = 1, 2, ..., l};
      for t = 1 to T = ⌈2 log n / 3 log log n⌉ do
      begin
        suppose ℘_{t-1} = {P_1, P_2, ..., P_s};
        ℘_t := ∅; END_t := ∅;
        for r = 1 to s do
        begin
```

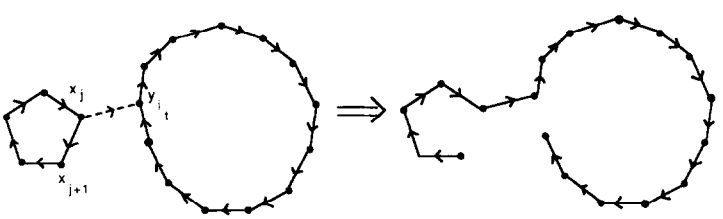

FIGURE 4

suppose $P_r = (u_1, u_2, \ldots, u_q)$;
if $(u_q, u_1) \in E_{m^*}$ **then**
begin
terminate FINDCYCLE successfully with
$C_1 = (u_1, u_2, \ldots, u_q, u_1)$
end else
begin
construct all possible paths obtainable from P_r in the following way (see Fig. 3);
if $(u_q, u_a) \in E_{m^*}$ and $(u_{a-1}, u_b) \in E_{m^*}$ where $b > a$ **then**
$\wp_t := \wp_t \cup \{\text{ROTATE}(P_r, a, b)\}$; $\text{END}_t := \text{END}_t \cup \{u_{b-1}\}$
where $\text{ROTATE}(P_r, a, b) = (u_1, u_2, \ldots, u_{a-1}, u_b, u_{b+1}, \ldots, u_q, u_a, \ldots, u_{b-1})$
end
end {r-loop}
end {t-loop} **end.**

We must show that Phase 3 a.a. succeeds. Let m_2, LARGE be as in Section 5. Consider the subgraph $\hat{H}$ of D_{m_0} induced by LARGE. Now the edges of $\hat{H}$ can be partitioned into $E^3 \cup E^4$, where E^3 consists of those edges that are in ELARGE but do not belong to E_{m_2}. Now, assuming (4.5), whatever E^4 is, E^3 is a random $|E^3|$ – subset of the remaining possible edges—see (5.1). But, given E^4, we obtain the same distribution for $\hat{H}$ by choosing a random $|E^3|$ – subset E^5 of ELARGE and then adding in $E^4 - E^3$ plus further randomly chosen edges.

Furthermore we can obtain E^5 by independently including each edge of ELARGE with probability $p = \log n/7n$ and excluding it with probability $1 - p$, then randomly adding or deleting edges to get a set of the required size. Since $|E^3| \approx n \log n/6$ a.a. we will a.a. be adding edges.

To summarise, we can a.a. consider $\hat{H}$ to be a supergraph of a graph H distributed as $D_{h,p}$ where $h = |\text{LARGE}| = n - o(n)$ and the edges of H occur independently of the outcome of Phase 1 or 2.

We need to be able to show that $|\text{END}_t|$ grows sufficiently rapidly with t inside FINDCYCLE. Unfortunately this is quite technical and requires several preliminary lemmas and constructions.

Let $C_1(i)$ denote the cycle C_1 **after** C_i has been merged in Phase 3 for $i \geq 2$. Let $C_{\text{Initial}} = C_1 = (u_1, u_2, \ldots, u_{n_1}, u_1)$ at the start of Phase 3 and let $n_2 = |\text{LARGE} \cap V(C_{\text{Initial}})| = n - o(n)$.

Let $\mu_1 = \lceil (\log n)^2/\log\log\log n \rceil$ and partition the path $P_{\text{Initial}} = (u_1, u_2, \ldots, u_{n_1})$ into consecutive blocks $A_1, A_2, \ldots, A_{\mu_1}$ each containing $\lfloor n_2/\mu_1 \rfloor$ or $\lceil n_2/\mu_1 \rceil$ large vertices. For $I \subseteq M = \{1, 2, \ldots, \mu_1\}$ let $A_I = \bigcup_{i \in I} A_i$. For each $I \subseteq M$, $|I| = \lfloor \mu_1/10 \rfloor$ we let $B_I = \{v \in \text{LARGE}:$

$|\{w \in \text{LARGE} \cap A_I : (v, w) \in E(H)\}| \leq \log n/75\}$ (= those v with substantially fewer than the expected number of out-neighbours in A_I).

LEMMA 6.1. *The following holds with probability* $1 - o(n^{-1})$:

$$I \subseteq M, |I| = \lfloor \mu_1/10 \rfloor, \textit{ implies } |B_I| \leq n^{0.99999}.$$

Proof. For a fixed I and positive integer k,

$$\begin{aligned}\Pr(|B_I| \geq k) &\leq \binom{h}{k} \mathrm{BS}(0, \lfloor \log n/75 \rfloor; \lfloor n_2/\mu_1 \rfloor \lceil \mu_1/10 \rceil, p)^k \\ &\leq \left(\frac{he}{k}\right)^k \exp\{-\tfrac{1}{2} \cdot \tfrac{1}{225} \cdot \tfrac{1}{71} \cdot k \log n\} \\ &\leq \left(n^{0.999985} e/k\right)^k.\end{aligned}$$

Hence

$$\Pr(\exists I : |B_I| \geq n^{0.99999}) \leq 2^{\mu_1} (n^{-0.000005} e) n^{0.99999} = o(n^{-1}).$$

Next let $\mu_2 = \lfloor \log n / \log\log\log\log n \rfloor$.

LEMMA 6.2. *The following holds with probability* $1 - o(n^{-1})$: $\nexists I \subseteq M$, $|I| = \lfloor \mu_1/10 \rfloor$, $v \in$ *LARGE and disjoint sets*

$$A = \{a_1, a_2, \ldots, a_{\mu_2}\} \subseteq \text{LARGE}, \qquad B = \{b_1, b_2, \ldots, b_{\mu_2}\} \subseteq B_I$$

such that

$$(v, a_i) \in E(H) \quad \textit{and} \quad ((b_i, a_i) \in E(H) \quad \textit{or} \quad a_i = \phi_2(b_i)), \quad i = 1, 2, \ldots, \mu_2.$$

Proof. For a fixed $I \subseteq M$, $v \in$ LARGE, $A, B \subseteq$ LARGE, where $a_i = \phi_2(b_i)$ for $i = 1, 2, \ldots, k$, the probability that the remaining conditions are satisfied is at most

$$p^{2\mu_2 - k} \mathrm{BS}(0, \lfloor \log n/75 \rfloor; \lfloor n_2/\mu_1 \rfloor \lceil \mu_1/10 \rceil, p)^{\mu_2} \leq p^{2\mu_2 - k} n^{-0.00001\mu_2}.$$

Thus

$$\begin{aligned}\Pr(\exists I, v, A, B \text{ as in the lemma}) &\leq 2^{\mu_1} n \binom{n}{\mu_2} \sum_{k=0}^{\mu_2} \binom{n}{\mu_2 - k} p^{2\mu_2 - k} n^{-0.00001\mu_2} \\ &\leq 2^{\mu_1} \sum_{k=0}^{\mu_2} (\log n)^{2\mu_2 - k} n^{-0.00001\mu_2} \\ &= o(n^{-1}).\end{aligned}$$

■

LEMMA 6.3. *Let* $0 < \alpha < 1$ *be fixed. Then* D_{m_1} *satisfies the following with probability* $1 - o(n^{-1})$: $\nexists A, B \subseteq V_n$ *such that* (i) $|A| \le a_0 = \alpha e^{-3} n/\log n$, (ii) $|B| \le \alpha|A|\log n/2$, *and* (iii) $|\{(v, w) \in E_{m_1}: v \in A, w \in B\}| \ge \alpha|A|\log n$.

Proof. We prove the result for D_{n,p_1} and apply (2.3):

$$\begin{aligned}
&\Pr(\exists A, B \text{ in } D_{n,p_1}) \\
&\le \sum_{k=1}^{a_0} \binom{n}{k}\binom{n}{\lfloor \alpha k \log n/2\rfloor}\binom{k\lfloor \alpha k \log n/2\rfloor}{\lceil \alpha k \log n\rceil} p_1^{\lceil \alpha k \log n\rceil} \\
&\le \sum_{k=1}^{a_0} \left(\frac{ne}{k}\left((1+o(1))\frac{2ne}{\alpha k \log n}\cdot\frac{e^2k^2}{4}\cdot\frac{(\log n)^2}{n^2}\right)^{\alpha \log n/2}\right)^k \\
&= \sum_{k=1}^{a_0} \left(\frac{ne}{k}\left(\frac{(1+o(1))e^3 k \log n}{2\alpha n}\right)^{\alpha \log n/2}\right)^k \\
&= o(n^{-2}).
\end{aligned}$$

■

Finally, a calculation similar to that for Lemma 4.1 yields

LEMMA 6.4. *The following holds with probability* $1 - o(n^{-3})$:

$$v, w \in \text{SMALL} = V_n - \text{LARGE} \text{ implies } \operatorname{dist}(v, w; GD_{m_1}) \ge 10. \quad (6.1)$$

We can now get down to proving that Phase 3 a.a. succeeds. We implicitly assume that (4.2) and the conditions of Lemmas 6.1–6.4 hold, but see Remark 2 below.

Consider any C_i, $i \ge 2$ and the first j such that both $x_j, x_{j+1} \in \text{LARGE}$. By Lemma 6.4 such a j a.a. exists. We show that for such an i, j

$$\Pr(\text{FINDCYCLE fails}) = O(n^{-\varepsilon_1}) \qquad \text{for some constant } \varepsilon_1 > 0. \quad (6.2)$$

It follows then that

$$\Pr(\text{Phase 3 fails}) = O(n^{-\varepsilon_1}\log n) + o(1) = o(1)$$

and Theorem 2 will follow.

We now consider such an execution of FINDCYCLE.

Proof of (6.2). The following observation is crucial.

At any stage the paths constructed are of the following form: we take P_{initial} and delete $O((\log n)^2/\log\log n)$ edges, permute the subpaths around and rejoin them with new edges (ROTATIONs) or paths (from the cycles

$C_2, C_3, \dots, C_{r_2}$).

Thus if n is large,

> If we partition the vertices of a constructed path P into consecutive sets $S_1, S_2, \dots, S_9$ each containing at least $\lfloor n_2/9 \rfloor$ vertices of LARGE $\cup$ $V(C_{\text{initial}})$ then for each $i = 1, 2, \dots, 9$, S_i contains a complete block A_{I_i}, where $|I_i| = \lfloor \mu_1/10 \rfloor$ for some I_i, each A_t, $t \in I_i$ appears on P as it did on P_{initial}. (6.3)

Let a constructed path $P = (z_1, z_2, \dots, z_p)$ be **good** if $z_p \in$ LARGE and $\exists\ I \subseteq M$, $|I| = \lfloor \mu_1/10 \rfloor$, and $r < s \le p/2$ with $s - r \le p/9$ such that $A_I \subseteq \{z_r, z_{r+1}, \dots, z_s\}$ and $z_p \notin B_I$.

Let now $\hat{\wp}_t$ denote those paths of $\wp_t$ which are good. Let $\hat{\text{END}}_t = \{v \in \text{END}_t : \hat{\wp}_t$ contains a path from x_{j+1} to $v\}$. By deleting paths, if necessary, we assume that for each $v \in \hat{\text{END}}_t$ there is a unique such $P_v \in \hat{\wp}_t$.

We first show

$$\Pr\left(|\hat{\wp}_0| \le \log n/8\right) = O(n^{-\varepsilon_1}). \tag{6.4}$$

Consider $C_1(i-1) = (w_1, w_2, \dots, w_q)$. If we partition it into subpaths $S_1, S_2, \dots, S_9$ then as in (6.3) we find a complete A_{I_k}, $|I_k| = \lfloor \mu_1/10 \rfloor$ contained in each S_k. For each w_l, at least 4 of these A_{I_k}'s will be in the first half of the path $(x_{j+1}, \dots, x_j, w_{l+1}, \dots, w_q, \dots, w_l)$ if $(x_j, w_{l+1}) \in E_{m^*}$. Now we can assume that $|B_{I_1} \cup B_{I_2} \cup \cdots \cup B_{I_9}| \le 9n^{0.99999}$ by Lemma 6.1. Thus $C_1(i-1)$ contains a set L of $n - o(n)$ large vertices w_l such that if (x_j, w_{l+1}) exists then the path produced is good. Since the H-edges x_i to $V(C_1(i-1)) \cap$ LARGE exist independently with probability $\log n/7$, (6.4) follows easily from (2.1a).

Remark 2. The definition of LARGE depends only on the first m_3 edges and so is independent of H. On the other hand, assuming the conditions of Lemmas 6.1–6.4 does (mildly) condition H. However,

$$\begin{aligned}
&\Pr\left(|\hat{\wp}_0| \le \log n/8\right) \\
&\quad \le \Pr\Big(x_j \text{ has } \le \log n/8 \text{ "good" neighbours } | \\
&\qquad\qquad |L| \ge n_1 = n_0 - 9n^{0.99999} - \sqrt{n}\Big) + \Pr(|L| < n_1) \\
&\quad \le (1-p)^{n_1} + o(n^{-3}),
\end{aligned}$$

where we can use $(1-p)^{n_1}$ as the H-edges incident with x_j are independent of L, and L will be large assuming the conditions of Lemmas 6.1–6.4.

The (almost) final piece in the jigsaw is, assuming the conditions of the previous lemmas,

$$1 \le |\hat{\mathrm{END}}_t| \le n/\left(456e^3(\log n)^2\right) \text{ implies}$$
$$|\hat{\mathrm{END}}_{t+1}| \ge (\log n/152)^2 |\hat{\mathrm{END}}_t|. \tag{6.5}$$

To prove this let $|\hat{\mathrm{END}}_t| \le n/(456e^3(\log n)^2)$. Let $v \in \hat{\mathrm{END}}_t$, $P_v = (z_1, z_2, \ldots, z_p)$ and r, s, I be as in the definition of a good path.

Case 1. $r > p/9$. Partition P_v into subpaths $S_1, S_2, \ldots, S_9$ containing complete blocks $I_1, I_2, \ldots, I_9$ as in (6.3). Let

$$F_1(v) = \{z_k : (v, z_k) \in E(H),\ z_k \in A_I,$$
$$z_{k-1} \in \mathrm{LARGE} - B_{I_9} \text{ and } (z_{k-1}, z_k) \in C_1(i-1)\}$$

and

$$F_2(v) = \{z_{k-1} : z_k \in F_1(v)\}.$$

Note that Lemmas 6.1, 6.2, and 6.4 imply

$$|F_1(v)| \ge \log n/75 - T - \mu_2 - 1 \ge \log n/76 \qquad \text{for } n \text{ large.} \tag{6.6}$$

For $z_{k-1} \in F_2(v)$ let

$$F_3(z_{k-1}) = \{z_l : l > k, (z_{k-1}, z_l) \in E(H),\ z_l \in A_{I_9},\ z_{l-1}$$
$$\in \mathrm{LARGE} - B_{I_1} \text{ and } (z_{l-1}, z_l) \in C_1(i-1)\}$$

and

$$F_4(z_{k-1}) = \{z_{l-1} : z_l \in F_3(Z_{k-1})\}.$$

Similarly to (6.6) we have

$$|F_3(z_{k-1})| \ge \log n/76 \qquad \text{for } n \text{ large.} \tag{6.7}$$

The important point to note is that

$$\text{if } z_k \in F_1(v),\ z_l \in F_3(z_{k-1}) \text{ then ROTATE } (P_v, k, l)$$
$$\text{is a good path with } I = I_1. \tag{6.8}$$

(This would not necessarily be true if we had used 8, say, in place of 9 in (6.3).)

Case 2. $r \le p/9$. We replace I_1 by I_3 throughout the above argument.

Let $\Phi_k = \bigcup_{v \in \hat{\mathrm{END}}_t} F_i(v)$ for $k = 1, 2$. Now by (6.6) there are at least $(\log n/76)|\hat{\mathrm{END}}_t|$ edges from $\hat{\mathrm{END}}_t$ to Φ_1. Hence by Lemma 6.3 with $\alpha = \frac{1}{76}$ and by (4.2),

$$(\log n/152)|\hat{\mathrm{END}}_t| \le |\Phi_1| \le (6 \log n)|\hat{\mathrm{END}}_t| \le n/(76e^3 \log n).$$

But $|\Phi_2| = |\Phi_1|$ since $z \in \Phi_2$ iff $z' \in \Phi_1$, where z' is the successor of z on $C_1(i-1)$. Putting $\Phi_k = \bigcup_{v \in \hat{\mathrm{END}}_t} \bigcup_{z \in F_2(v)} F_i(z)$, for $k = 3, 4$, we see by (6.7) that there are at least $(\log n/76)|\Phi_2|$ edges from Φ_2 to Φ_3.

Hence, by Lemma 6.3,

$$|\Phi_3| \ge (\log n/76)|\Phi_2|.$$

Also $|\Phi_4| = |\Phi_3|$ by the argument for $|\Phi_2| = |\Phi_1|$. Finally $\Phi_4 \subseteq \hat{\mathrm{END}}_{t+1}$ by (6.8) and so (6.5) follows.

Thus there exists $\tau \le T - 1$ such that $|\hat{\mathrm{END}}_\tau| \ge \lfloor n/(456e^3(\log n)^2 \rfloor$. Choose a subset END of $\hat{\mathrm{END}}_\tau$ of exactly this size and apply the preceding argument using END in place of $\hat{\mathrm{END}}_\tau$ to show that $|\hat{\mathrm{END}}_{\tau+1}| \ge \beta n$ for some constant $\beta > 0$. Now the existence of an edge in H from $\hat{\mathrm{END}}_t$ to x_{j+1} is independent of the previous history, at the time of checking (see Remark 2). Hence

$$\Pr(\text{FINDCYCLE fails}) \le (1-p)^{\beta n} \le n^{-\beta/7}.$$

and Theorem 2 has been proved, except for the claimed running time.

Running time of DHAM. We analyse this assuming that those events proved to occur a.a. actually occur. We can always put a time limit on DHAM to achieve the $O(n^{1.5})$ running time with certainty. We assume that D_{m^*} is processed in adjacency list form.

Phase 1. It takes $O(n \log n)$ time to construct BIP assuming $O(1)$ time to obtain a random bit. (To construct σ we need to be able to find the kth largest from a set of integers where k is randomly chosen.) This takes $O(\log n)$ time, using a height balanced tree—see, for example, Aho, Hopcroft, and Ullman [1]. The running time of the algorithm in [8] is $O(n^{1.5})$ since BIP has at most $20n$ edges.

Phase 2. We make $O(\log n)$ executions of the loop beginning at A. It takes $O(1)$ time to see if an edge belongs to F_2 and $O(\log n)$ time to see if it belongs to E_{m_2}. The vertex sets of the cycles of ϕ_2 partition V_n and the operations required on these sets is UNION and FIND—see, for example, [1]. It follows that $O(n(\log n)^3)$ time is sufficient for this phase.

Phase 3. Our assumptions imply that it takes at most 2 executions of FINDCYCLE to merge a C_i into C_1. Each execution of FINDCYCLE

generates $O((6 \log n)^T) = O(n^{4/3+o(1)})$ paths altogether. There is a natural tree structure to the set of paths produced where a path is the father of all the paths produced from it by one double-rotation. The description of the algorithm suggests a breadth-first search of this tree. It is, however, more efficient to explore this tree by depth-first search. We can then keep a single path, as a doubly linked list and create the next son in $O(1)$ time. Backtracking takes $O(1)$ time per level and we obtain a running time of $O(n^{4/3+o(1)})$ for the whole of Phase 3. □

It is easy to see that Theorem 1 follows from Theorem 2. Let $e_1, e_2, \ldots, e_{n(n-1)}$ be a random permutation of $E(DK_n)$ and $m = n \log n + c_n n$. Theorem 1 states the limiting value of Pr(DHAM finds a hamilton cycle in D_m). Now given $m < m^*$ then this is zero and given $m \geq m^*$ then the preceding proof shows that this probability tends to one (or only allow DHAM to use the first m^* edges). The value of $\lim_{n \to \infty} \Pr(m \geq m^*)$ is easy to compute and is in fact (implicitly) done in [7].

Outline Proof of Theorem 3. The idea is to (i) construct k disjoint sets $F_1, F_2, \ldots, F_k$ where each F_i is the set of edges of a set of vertex disjoint cycles covering V_n, (ii) split the edges of $(E_{m_0} - E_{m_3}) \cap \text{ELARGE}$ into k sets of roughly equal size $E_1, E_2, \ldots, E_k$ and use E_i to patch up the cycles in F_i as above, for $i = 1, 2, \ldots, k$.

To construct $F_1, F_2, \ldots, F_k$ we apply Phase 1 k times. It is straightforward to show that we a.a. succeed in constructing k permutations $\phi_1, \phi_2, \ldots, \phi_k$. If we consider a fixed i, then our randomising method (i.e., use of σ) ensures that ϕ_i is sampled uniformly and so a.a. has at most $2 \log n$ cycles. As k is constant $\phi_1, \phi_2, \ldots, \phi_k$ will a.a. all have at most $2 \log n$ cycles. The remaining proof goes through much as before with only trivial changes to constants.

7. Travelling Salesman Problems

Proof of Theorem 4. Given an instance of this problem let the edges of DK_n be ordered $e_1, e_2, \ldots, e_{n(n-1)}$, where $w(e_i) \leq w(e_{i+1})$ for $i = 1, 2, \ldots, n(n-1)$ and $w(e)$ is the weight of edge e. Note that our assumption implies that each such ordering is equally likely. If E_m, D_m, m^* are as before then we have

ALGORITHM DBOT.

begin
  apply DHAM to D_{m^*}
end

It is clear that if DHAM succeeds then it solves BTSP and Theorem 2 states

that it a.a. does. To obtain $O(n^2)$ running time we do not sort all the edges but instead build a heap out of them and generate them in increasing order of weight until m^* is reached. This does require $O(n^2)$ time a.a. □

Outline Proof of Theorem 5. We shall assume throughout the $B(n) = \lceil \alpha n/\log n \rceil$ for some constant $\alpha > 0$. We can assume $\alpha \geq 1$ for if not then it is easy to see that we can use a.a. DHAM to construct a Hamilton cycle using edges of length zero only.

We shall describe DTSPSOLVE part formally and part informally. DTSPSOLVE first, independently, randomly colours each element of $V_n \times V_n$ red or blue with probability $\frac{1}{2}$. E^r is the set of red edges and E^b the set of blue edges produced. Also let $E^{rk} = \{e \in E^r: w(e) \leq k\}$ for $k = 0, 1, \ldots, B(n) - 1$. Define E^{bk} similarly and let $E^k = E^{rk} \cup E^{bk}$. Note that $D^k = (V_n, E^k)$ is distributed as $D_{n,(k+1)p}$, where $p = 1/B$.

The strategy now is similar to that of [9]. We let

$$X = \{v \in V_n: d^+(v, E_{r0}) \leq l_0 \text{ or } d^-(v, E_{r0}) \leq l_0\}.$$

Now let

$$\mathcal{F} = \{F \subseteq E(DK_n): \text{(i) } F \text{ induces a set of vertex disjoint paths,}$$
$$\text{(ii) each } x \in X \text{ is an interior vertex of one such path,}$$
$$\text{(iii) each edge of } F \text{ is incident with a vertex of } X.\}$$

We first try to compute a minimum total weight member of $\mathcal{F}$ and then extend it, using zero-length red edges, to a set of $O(\log n)$ vertex disjoint cycles covering V_n. More formally,

PHASE 1.

begin
$\sigma :=$ a random permutation of V_n;
for $e \in V_n \times V_n$ **do** $\hat{w}(e) := w(\hat{e})$ where $(v, \hat{w}) = (v, \sigma(w))$ as before;
let $e_1, e_2, \ldots, e_{n^2}$ be random ordering of $V_n \times V_n$;
$E^{1+} := E^{1-} := \emptyset$;
A: **for** $i = 1$ **to** n^2 **do**
begin
if $d^+(v_i, E^{1+}) \leq \log\log n$ **and** $e_i \in E^{r0}$ **then** $E^{1+} := E^{1+} \cup \{\hat{e}_i\}$;
{where $e_i = (v_i, w_i)$}
end;
B: **for** $i = 1$ **to** n^2 **do**
begin
if $e_i \notin E^{1+}$ **and** $d^-(\sigma(w_i), E^{1-}) \leq \log\log n$ **and** $e_i \in E^{r0}$ **then** $E^{1-} := E^{1-} \cup \{\hat{e}_i\}$
end

$E^1 := E^{1+} \cup E^{1-}$;
compute F^* where $F^* = \min\{\hat{w}(F): F \in \mathscr{F}\}$ and $\hat{w}(F) = \Sigma_{e \in F}\hat{w}(e)$;
$Y_0 := \{v \in V_n: \exists e = (v, w) \in F^*\}$ and $Y_1 := \{w \in V_n: \exists e = (v, w) \in F^*\}$;
BIP := bipartite graph $(V_n - Y_0, V_n - Y_1, A)$ where $(v, w) \in A$ iff $(v, w) \in E^1$;
apply the algorithm of Even and Tarjan [8] to find a maximum matching M of BIP;
if M is not a perfect matching **then** DTSPSOLVE has failed **else**
$F' := M$; (as a set of edges of a graph with vertex set V_n.)
output $F = \{e: \hat{e} \in F' \cup F^*\} = \{(v, \phi(v)): v \in V_n\}$, say.
end

Assuming Phase 1 is successful we then (Phase 2) randomly choose half the edges of E^{b0} and use them to try to patch together cycles as in Phase 2 of DHAM. We then (Phase 3) use the remaining edges of E^{b0} to try to finish the patching as in Phase 3 of DHAM.

We now state as a lemma, some properties which can be proved by straightforward calculation.

LEMMA 7.1. *The following hold a.a.:*

$$|X| \leq n^{1-1/2\alpha}. \tag{7.1a}$$

If $\mu = \lceil 2\alpha \rceil$ *then*

$$\delta^+(D^\mu), \delta^-(D^\mu) \geq \log n/3, \tag{7.1b}$$

$$\Delta^+(D^{2\mu}), \Delta^-(D^{2\mu}) \leq 12 \log n, \tag{7.1c}$$

$GD^{2\mu}$ *does not contain any cycle* C *with* $|V(C) \cap X| \geq \lfloor |C|/10 \rfloor$, (7.1d)

$GD^{2\mu}$ *does not contain a tree with more than* 6α *vertices at least one third of which are in* X. (7.1e)

Now (7.1a) implies that $\mathscr{F} \neq \emptyset$. If we succeed in constructing a Hamilton cycle C^* then $w(C^*) = w(F^*)$ and for any Hamilton cycle C, $w(C) \geq w(F^*)$—just delete all edges of C not incident with a vertex in X to obtain a member of $\mathscr{F}$.

(7.1b) and (7.1e) imply that $F^* \subseteq E^{2\mu}$, for any $e \notin F^* \cap E^{2\mu}$ could be replaced by 2 edges of E^μ to yield an F of smaller weight. Thus to compute F^* we need only look at the edges E^* of $E^{2\mu}$ which are incident with X.

(7.1d) implies that E^* induces a forest in $GD^{2\mu}$. Let T be a tree in this forest. All its leaves are in $V_n - X$, by (7.1b). If $x \in X$ is adjacent to $t \geq 3$ leaves in T then there are either ≥ 2 edges of T directed into x from a leaf or ≥ 2 edges of T directed from x to a leaf. But we can then remove the larger weight edges to leave at most one edge directed from a leaf to a given

$x \in X$ or from a given $x \in X$ to a leaf. Assume this done. Then (7.1e) implies that the remaining tree has at most 12α vertices—if it has k nonleaf vertices then at least $\lfloor k/2 \rfloor + 1$ are in X. We can deal with each such tree in $O(1)$ time and hence compute F^* in $o(n)$ time.

We henceforth assume that the search for F^* is restricted to the subgraph induced by the vertices $Y = X \cup N(X, GD^{2\mu})$.

$$\text{This leaves the lengths of blue edges joining 2 vertices in } V_n - Y \text{ unconditioned by Phase 1.} \tag{7.2}$$

We must now show that Phase 1 a.a. succeeds. Now if BIP does not have a perfect matching then by Hall's theorem there exists $K \subseteq V_n - Y_0$, $L \subseteq V_n - Y_1$, $|L| = |K| - 1$, and $N(K, \text{BIP}) \subseteq L$. Suppose first as in the proof of Lemma 4.2 that $|K| \leq n/2$. Then either (7.1e) fails or, since $K \cap X = \varnothing$ each $v \in K$ makes at least $\log\log n - 3\alpha$ choices in the L in the construction of E^1. The probability of the latter event $\to 0$ by a similar calculation to that in Lemma 4.2(a). For $|K| > n/2$ we note that the equivalent of (4.5) holds and we can finish as we did in Lemma 4.2(a).

We can argue as in Lemma 4.2(b) that the permutation ϕ output is equally likely to be any permutation of V_n—it does not matter to the theorem if we allow loops in the construction of E^1.

The arguments that Phase 2 and Phase 3 work a.a. are as for DHAM except for trivial changes in constants, which now depend on α, the use of (7.1d) in place of Lemma 6.4, and $V_n - Y$ in place of LARGE (see 7.2).

ACKNOWLEDGMENT

I thank Martin Dyer for helpful comments and a particularly thorough reading of the paper.

REFERENCES

1. A. V. AHO, J. E. HOPCROFT, AND J. D. ULLMAN, "The Design and Analysis of Computer Algorithms," Addison-Wesley, Reading, MA, 1974.
2. D. ANGLUIN AND L. VALIANT, Fast probalistic algorithms for Hamilton circuits and matchings, *J. Comput. System Sci.* **18** (1979), 155–193.
3. B. BOLLOBAS, The evolution of sparse graphs, *in* "Graph Theory and Combinatorics Proceedings, Cambridge Combinatorial Conference in Honour of Paul Erdos, 1984" (B. Bollobas, Ed.), pp. 335–357.
4. B. BOLLOBAS, "Random Graphs," Academic Press, New York/London, 1985.
5. B. BOLLOBAS, T. I. FENNER, AND A. M. FRIEZE, An algorithm for finding hamilton cycles in random graphs, *in* "Proccedings, 17th Annual ACM Symposium on Theory of Computing, 1985," pp. 430–439.
6. M. E. DYER AND A. M. FRIEZE, On patching algorithms for random asymmetric travelling salesman problems, to appear.

7. P. ERDOS AND A. RENYI, On random matrices, *Publ. Math. Inst. Hungar. Acad. Sci.* (1964), 455–461.
8. S. EVEN AND R. E. TARJAN, Network flow and testing graph connectivity, *SIAM J. Comput.* **4** (1975), 507–518.
9. A. M. FRIEZE, On the exact solution of random travelling salesman problems with medium sized integer costs, to appear.
10. W. HOEFFDING, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.
11. R. M. KARP, A patching algorithm for the non-symmetric travelling salesman problem, *SIAM J. Comput.* **8** (1979), 561–573.
12. R. M. KARP AND J. M. STEELE, Probabilistic analysis of heuristics, *in* "The Travelling Salesman Problem: A Guided Tour" (E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy-Kan, and D. B. Shmoys, Eds.), Wiley, New York, 1985.
13. J. KOMLOS AND E. SZEMEREDI, Limit distribution for the existence of Hamilton cycles in random graphs, *Discrete Math.* **43** (1983), 55–63.
14. C. J. H. MCDIARMID, Clutter percolation and random graphs, *Math. Programming Stud.* **13** (1980), 17–25.
15. D. W. WALKUP, Matchings in randon regular bipartite graphs, *Discrete Math.* **31** (1980), 59–64.