SIAM J. COMPUT.
Vol. 16, No. 6, December 1987

007

# ON THE EXACT SOLUTION OF RANDOM TRAVELLING SALESMAN PROBLEMS WITH MEDIUM SIZE INTEGER COEFFICIENTS*

A. M. FRIEZE†

**Abstract.** Let edge weights for the complete graph on vertex set $\{1, 2, \cdots, n\}$ be chosen independently and uniformly from $\{0, 1, \cdots, B(n)-1\}$ where $B(n) = o(n/\log\log n)$. We show that there exists a polynomial time ($O(n^3 \log n)$) algorithm which solves the associated travelling salesman problem with probability tending to 1 as $n$ tends to $\infty$.

**Key words.** travelling salesman problem, probabilistic analysis

**AMS(MOS) subject classifications.** 68A20, 68F15

**1. Introduction.** The Travelling Salesman Problem (TSP) is one of the most studied problems in combinatorial optimisation and a recent book [16] is devoted entirely to its analysis.

We shall consider the *symmetric* problem in which one is given edge weights $c: E(K_n) \to Z$ where $E(K_n)$ is the edge set of the complete graph $K_n$ with vertex set $V_n = \{1, 2, \cdots, n\}$. The problem is to find a tour of hamilton cycle of $K_n$ of minimum total weight.

We shall look at the problem from the point of view of probabilistic analysis. Previous work e.g. Karp [14], [15] or Halton and Terada [13], has concentrated on solving the problem approximately with high probability. In this paper we make some progress on the problem of finding a polynomial time algorithm that can solve the problem *exactly* with probability tending to 1 as $n$ tends to infinity.

In order to do this, we have to be more restrictive about edge weights than previous authors. We assume that $\{c(e): e \in E(K_n)\}$ is a set of independent random variables where each $c(e)$ is chosen uniformly from $\{0, 1, \cdots, B-1\}$ where $B = B(n)$. We prove the following.

THEOREM 1.1. *If* $B(n) = o(n/\log\log n))$ *then there is a polynomial time* ($O(n^3 \log n)$) *algorithm ALGT such that*

$$\lim_{n\to\infty} \Pr(\textit{ALGT solves TSP } \underline{\textit{exactly}}) = 1. \qquad \square$$

ALGT can be considered as an extension of the algorithm HAM described in Bollobás, Fenner and Frieze [4] for finding hamilton cycles in random graphs, although we have had to make several changes in the description and proof of validity. To see the relationship, consider the set of edges of length zero. The subgraph $G_0$ of $K_n$ induced by these edges is distributed like $G_{n,p}$, $p = 1/B(n)$, where each edge of $K_n$ is independently included with probability $p$ and excluded with probability $1-p$. If $p$ is large enough then $G_0$ is hamiltonian with probability tending to 1 and HAM can be used to find such a cycle with probability tending to 1. Here $p$ large enough is equivalent to $B(n) \approx n/\log n$. If $B(n) > n/\log n$ then $G_0$ is almost always nonhamiltonian and this paper is concerned with dealing with this problem for $B(n)$ growing faster but not too fast.

The paper is in 2 parts. In the first part, §§ 2–4, we describe an algorithm which we can prove correct for $B(n) = o(n/\sqrt{\log n})$ and which runs in $O(n^{2+o(1)} \log n^{(\log n/\log d)})$ time, where $d = n/B(n)$. If $d$ grows proportionally to $\log n$

* Received by the editors October 23, 1985; accepted for publication (in revised form) February 13, 1987.

† Department of Mathematics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, and Department of Computer Science and Statistics, Queen Mary College, Mile End Road, London E1 4NS, England.

then this is nearly as good as is claimed in the abstract. But if $d$ grows like $\sqrt{\log n}$ we can only claim $O(n^{4+o(1)})$. The second part is devoted to actually proving Theorem 1.1. The reason for describing 2 methods is that the first method is deterministic. The only randomisation needed is in the input, i.e., we have a deterministic algorithm that works on almost all problems. Our second algorithm is randomised: it requires the selection of a random subset of the zero length edges. The methods are similar enough so that only a few extra lemmas are needed to prove the validity of the second approach.

**Solution strategy.** This is quite simple at the top level. We identify a set of "troublesome" vertices $X_0$ which we "cover" as cheaply as possible by a set of vertex disjoint paths $F^*$ (Phase 1). Having done this we construct a tour through $V_n$ which contains $F^*$ and zero length edges otherwise. It is not then difficult to see that this will be a minimum length tour. Unfortunately there are quite a lot of details.

**Notation.** For ease of reference we gather together in one place the main notation common to the rest of the paper.

Let $p = 1/B(n) = \omega\sqrt{\log n/n}$ initially, where $\omega = \omega(n) \to \infty$, and let $d = np$. We assume that $d \leqq 1.1 \log n$. (For $d > 1.1 \log n$ we look for a hamilton cycle in $G_0$.)

Let $E_k = \{e \in E(K_n): c(e) \leqq k\}$, $0 \leqq k < B$ and let $G_k = (V_n, E_k)$. Note that $G_k$ is distributed like $G_{n,(k+1)p}$.

For $v \in V_n$ let $d_k(v)$ denote the degree of $v$ in $G_k$ and

$$\delta_k = \min\{d_k(v): v \in V_n\}.$$

For a graph $G$, $W \subseteq V(G)$ let $N(W, G) = \{v \in V(G) - W: \exists w \in W$ such that $\{v, w\} \in E(G)\}$

$$S(m, p, k) = \sum_{t=0}^{k} \binom{m}{t} p^t (1-p)^{m-t},$$

$\lambda = \lceil\sqrt{\log n}\rceil$ and $\mu = \lceil 2 \log n/d\rceil$.

An event $\mathscr{E}_n$ dependent on $n$ will be said to occur *almost always* (a.a.) if $\lim_{n\to\infty} \Pr(\mathscr{E}_n) = 1$.

In what follows we will sometimes use real values where we should round up or down. Where we do so the reader can be assured that by rounding up or down the desired objective can be achieved. Furthermore when we claim certain inequalities we will implicitly only claim them for $n$ large.

**2. Structural properties of the $G_k$.** Our main problem is in dealing with the following set of vertices:

$$X_0 = \{v \in V_n: d_0(v) \leqq d/2\}$$

that are equal to the set of vertices incident with "few" edges of weight zero. The following lemma summarises part of what we usually expect of $X_0$ and $G_{2\mu}$: $X_0$ is small, $G_{2\mu}$ is reasonably sparse and contains few small cycles. $X_0$ is (almost) randomly distributed over $V_n$ and so is spread out and a long way from any small cycle.

LEMMA 2.1. *The following properties hold a.a.*

(a) $|X_0| \leqq ne^{-d/10}$.

(b) $G_{2\mu}$ *is connected.*

(c) $Z \subseteq V_n$, $|Z| \leqq 3\lambda$ *and* $|Z \cap X_0| \geqq \lambda$ *implies* $G_{2\mu}[Z]$ *is not connected* {*i.e.*, $X_0$ *is "spaced out" in* $G_{2\mu}$}.

(d) $C$ *is a cycle in* $G_{2\mu}$, $|C| \leqq 2\lambda$ *implies* $|C \cap X_0| \leqq |C|/10$.

(e) $G_{2\mu}$ *contains fewer than* $(\log n)^9$ *cycles of length* 8 *or less.*

(f) *The maximum degree in* $G_{2\mu}$ *is* $\leqq 30 \log n$.

*Proof.* (a)

$$E(|X_0|) = n \sum_{k=0}^{d/2} \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

$$\leqq c_1 \binom{n-1}{\lfloor d/2 \rfloor} p^{\lfloor d/2 \rfloor} (1-p)^{n-1-\lfloor d/2 \rfloor}$$

(for some constant $c_1$. We use this notation for unspecified constants)

$$\leqq c_2 n \left( \frac{de}{\lfloor d/2 \rfloor} \right)^{\lfloor d/2 \rfloor} e^{-d} \qquad \text{using } \binom{n}{k} \leqq \left( \frac{ne}{k} \right)^k \text{ and } 1+x \leqq e^x$$

$$\leqq c_3 n (e/2)^{-d/2}$$

$$\leqq c_3 n e^{-.15d}.$$

Now use the Markov inequality $-\Pr(|X_0| \geqq x) \leqq E(|X_0|)/x$ for $x > 0$.

(b) This follows from Erdös and Rényi [5] as $G_{2\mu}$ "has more edges" than $G_{n,4\log n/n}$.

(c) If (b) holds and (c) fails, then there exists $Z \subseteq V_n$, $E \subseteq E_{2\mu}$ such that

(i) $(Z, E)$ is a tree,

(ii) $|Z \cap X_0| \geqq \lambda$, and

(iii) $|Z| = k = 3\lambda$.

Let $\alpha =$ the number of triples $(Z, Z \cap X_0, E)$ in $G_{2\mu}$ satisfying (i), (ii) and (iii). Then

$$\Pr(\alpha \neq 0) \leqq E(\alpha) \leqq \binom{n}{k} \binom{k}{\lambda} k^{k-2} ((2\mu+1)p)^{k-1} S(n-k, p, d/2)^{\lambda}$$

$$\leqq c_4 \left( \frac{ne}{k} \right)^k \left( \frac{ke}{\lambda} \right)^{\lambda} k^{k-1} \left( \frac{4 \log n}{n} \right)^{k-1} \left( 1 + \frac{3d}{4 \log n} \right)^{k-1} e^{-.15 d\lambda}$$

$$\leqq c_5 \, e^{\log n + 3\lambda \log\log n + 9\lambda - .15\lambda d}$$

$$= o(1).$$

(d) Let $\beta =$ the number of pairs $(C, C \cap X_0)$ where $C$ is a cycle of length $\leqq 2\lambda$ in $G_{2\mu}$ and $|C \cap X_0| \geqq |C|/10$. Then

$$\Pr(\beta \neq 0) \leqq E(\beta) \leqq \sum_{k=3}^{2\lambda} \binom{n}{k} ((k-1)!/2) \binom{k}{\lceil k/10 \rceil}$$

$$\times ((2\mu+1)p)^k S(n-k, p, d/2)^{\lceil k/10 \rceil}$$

$$\leqq c_6 \sum_{k=3}^{2\lambda} \left( \frac{ne}{k} \right)^k ((k-1)!/2)(10e)^{k/10} \left( \frac{8 \log n}{n} \right)^k e^{-.015dk}$$

$$\leqq \sum_{k=3}^{2\lambda} (c_7 \log n \, e^{-.015d})^k$$

$$= o(1).$$

(e) The expected number of cycles of length 8 or less in $G_{2\mu}$ is $O((\log n)^8)$. Now use the Markov inequality.

(f) Let $\gamma$ equal the number of vertices of degree exceeding $30 \log n$ in $G_{2\mu}$. Then

$$\begin{aligned}\Pr(\gamma \neq 0) \leqq E(\gamma) &\leqq n\binom{n-1}{\lceil 30 \log n \rceil}((2\mu+1)p)^{\lceil 30 \log n \rceil}\\ &\leqq n\left(\left(\frac{ne}{30 \log n}\right)\left(\frac{8 \log n}{n}\right)\right)^{30 \log n}\\ &= o(1)\end{aligned}$$

(we assumed that $d \leqq 1.1 \log n$). □

The next lemma shows that a.a. all vertices have some small incident edges.

LEMMA 2.2. *$\delta_\mu \geqq (\log n)/3$ a.a.*

*Proof.*

$$\begin{aligned}\Pr(\exists v \in V_n\colon d_\mu(v) < (\log n/3) &\leqq E(\text{number of vertices of degree} < (\log n)/3)\\ &\leqq n \sum_{k=0}^{(\log n)/3} \binom{n-1}{k}\\ &\quad \times ((\mu+1)p)^k(1-(\mu+1)p)^{n-1-k}\\ &\leqq 2n \sum_{k=0}^{(\log n)/3}\left(\frac{ne}{k}\right)^k\\ &\quad \times \left(\left(\frac{2 \log n}{n}\right)\left(1+\frac{np}{\log n}\right)\right)^k e^{-(np+2\log n)}\\ &\leqq 2n \sum_{k=0}^{(\log n)/3}\left(\frac{ne}{k}\right)^k\left(\frac{2 \log n}{n}\right)^k n^{-2}\\ &= o(1)\end{aligned}$$

as the largest term in the above summation $= O(n^{-1.06})$. □

**3. Phase 1 of ALGT.** With these preliminaries taken care of we describe Phase 1 of ALGT in detail.

We first describe a simple lower bound on the length of any tour. Let

$\mathcal{F} = \{F \subseteq E(K_n)\colon F$ induces a set of vertex disjoint paths such that:
(1) the endpoints of each path of $F$ are in $V_n - X_0$,
(2) every vertex in $X_0$ is in the interior of some path of $F$,
(3) $e \cap X_0 \neq \varnothing$ for all $e \in F\}$.

If $|X_0| < n/2$ (as it a.a. is by a long way—by Lemma 2.1) then, for any tour $T$

$$c(T) \geqq LB = \min\{c(F)\colon F \in \mathcal{F}\} \tag{3.1}$$

where for $S \subseteq E(K_n)$, $c(S) = \sum_{e \in S} c(e)$.

To show (3.1) simply delete all edges of $T$ joining 2 vertices not in $X_0$ in order to obtain a member of $\mathcal{F}$. This also shows that $\mathcal{F} \neq \varnothing$ in this case. The main result of this paper is that in polynomial time we can a.a. find an $F^*$ and extend it to a tour $T^*$ for which

$$c(T^*) = LB \tag{3.2}$$

which implies $T^*$ is optimal.

Phase 1 is concerned with finding $F^*$.

Gurevich and Shelah [12] also use the idea of covering "troublesome" vertices with paths.

Let L2.1 and L2.2 denote the events defined in Lemmas 2.1 and 2.2. (We shall use this notation to describe several events.)

Let $\mathcal{F}^* = \{F \in \mathcal{F}: c(F) = LB\}$.

LEMMA 3.1. *If* L2.1 $\cap$ L2.2 *occurs and* $F^* \in \mathcal{F}^*$, *then* $F^* \subseteq E_{2\mu}$, *for large* $n$.

*Proof.* Assume L2.1 $\cap$ L2.2 occurs. Suppose then that $e = \{a, b\} \in F^*$ and $c(e) > 2\mu$. Assume first that $e \subseteq X_0$. Let $A = \{a_1, a_2, \cdots, a_r\}$, $B = \{b_1, b_2, \cdots, b_s\}$, $r, s \geqq (\log n)/3 - 1$ be such that $\{a, a_i\} \in E_\mu$ and $\{b, b_j\} \in E_\mu$ for $i = 1, 2, \cdots, r$, $j = 1, 2, \cdots, s$ — $A$, $B$ exist because of L2.2., L2.1(d) implies that $|A \cap B| \leqq 1$ and L2.1(c) implies that at most $3\lambda$ members of $A \cup B$ are vertices of paths of $F^*$. Thus we can pick $a_k \neq b_l$ such that $a_k, b_l \in V_n - X_0$ and $F' = (F^* - \{e\}) \cup \{\{a, a_k\}, \{b, b_l\}\} \in \mathcal{F}$. But then $c(F') < c(F)$—contradiction. The case where $a \in X_0$, $b \notin X_0$ is similar. We replace $\{a, b\}$ by an edge $\{a, a_k\} \in E_\mu$ chosen as above. □

**Phase 1 of ALGT.**

*begin*

check for the occurrence of L2.1 $\cap$ L2.2—this can easily be done in $O(n^3)$ time;

*if* either event fails to occur *then* terminate unsuccessfully

*else* find $F^* \in \mathcal{F}^*$

*end*

LEMMA 3.2. *If* L2.1 $\cap$ L2.2 *occurs and n is large, then* $F^*$ *can be found in* $O(|E_{2\mu}|)$ *time.*

*Proof.* Assume that L2.1 $\cap$ L2.2 occurs. Let $E^* = \{e \in E_{2\mu}: e \cap X_0 \neq \varnothing\}$ and $V^* = \{v \in V_n: \exists e \in E^* \text{ such that } v \in e\}$. Let $G^* = (V^*, E^*)$. Lemma 3.1 implies that to find $F^*$ we must cover $X_0$ as cheaply as possible by vertex disjoint paths in $G^*$ all of whose endpoints are in $V^* - X_0$. We note that L2.1(f) implies

$$|V^*| \leqq 30n\, e^{-d/10} \log n = o(n) \tag{3.3}$$

and that L2.1 implies that $G^*$ is a forest. (Every edge of $G^*$ contains a member of $X_0$ and so by L2.1(d) there are no cycles of size $2\lambda$ or less. But a larger cycle will contradict L2.1(c).) We are thus left with a set of path covering problems on trees. We can solve these in linear time using the dynamic programming approach of Bern, Lawler and Wong [2].

Let $T$ be a component of $G^*$ and let $r \in T$ be chosen arbitrarily as a root. Let

$$\Phi = \{X \subseteq E(T): X \text{ induces a set of vertex disjoint paths}\}.$$

For $X \in \Phi$ let $I(X) = \{\text{internal nodes of the paths of } X\}$ and $O(X) = \{\text{endpoints of the paths of } X\}$ and $V(X) = I(X) \cup O(X)$. Let

$$\Phi(T, r) = \{X \in \Phi: \text{(i) } X_0 - \{r\} \subseteq I(X), \text{(ii) } V(T) - X_0 \supseteq O(X) - \{r\}\}$$

and

$$\Phi_i(T, r) = \{X \in \Phi(T, r): r \in A_i(X)\}, \qquad i = 1, 2, 3$$

where $A_1(X) = V(T) - V(X)$, $A_2(X) = I(X)$ and $A_3(X) = O(X)$. Assume now that $\Phi(T, r)$ and the $\Phi_i(T, r)$ are ordered by total weight, i.e., if $X, X' \in \Phi(T, r)$, then $X \leqq X' \leftrightarrow c(X) \leqq c(X')$.

Then

$$\Psi(T, r) = \begin{cases} \min \Phi(T, r) & \text{if } r \notin X_0, \\ \min \Phi_2(T, r) & \text{if } r \in X_0, \end{cases}$$

is the answer to our problem—the min operator breaks ties arbitrarily. Now if $(T_1, r_1)$, $(T_2, r_2)$ are vertex disjoint rooted subtrees of $(T, r)$ then we define a rooted tree $(T_1 \circ T_2, r_1)$ where $V(T_1 \circ T_2) = V(T_1) \cup V(T_2) \cup \{r\}$ and $E(T_1 \circ T_2) = E(T_1) \cup E(T_2) \cup \{e\}$ where $e = \{r_1, r_2\}$. $(T, r)$ can be constructed from a set of single vertices by $|V(T)| - 1$ applications of $\circ$.

All we have to demonstrate is that if we know $\min \Phi_i(T_j, r_j)$ for $i = 1, 2, 3$ and $j = 1, 2$ then we can compute $\min \Phi_i(T_1 \circ T_2, r_1)$ in $O(1)$ time. Let

$$\Psi_i = \min \{Z \in \Phi_i(T_1 \circ T_2, r_1): e \notin Z\}$$

and

$$\Psi_i' = \min \{Z \in \Phi_i(T_1 \circ T_2, r_1): e \in Z\}.$$

Then $\min \Phi_i(T_1 \circ T_2, r_1) = \min \{\Psi_i, \Psi_i'\}$ and

$$\Psi_i = (\min \Phi_i(T_1, r_1)) \cup \Psi(T_2, r_2), \qquad i = 1, 2, 3,$$

$$\Psi_1' = \varnothing,$$

$$\Psi_2' = (\min \Phi_3(T_1, r_1)) \cup \{e\} \cup Y$$

where

$$Y = \begin{cases} \min \{\min \Phi_1(T_2, r_2), \min \Phi_3(T_2, r_2)\} & \text{if } r_2 \notin X_0, \\ \min \Phi_3(T_2, r_2) & \text{if } r_2 \in X_0, \end{cases}$$

and

$$\Psi_3' = (\min \Phi_1(T_1, r_1)) \cup \{e\} \cup Y.$$

To make the implied operations obviously executable in constant time we keep the minimal weights $c(\Phi_i(T_j, r_j))$ plus flags indicating how they were constructed and then work backwards to construct the optimal solution. □

Let $X_1 = \{$interior vertices of paths of $F^*\}$.

LEMMA 3.3. *If* L2.1 ∩ L2.2 *occurs, then* $|X_1| = O(n\, e^{-d/10})$. □

**4. Phase 2 of ALGT.** Assuming that L2.1 ∩ L2.2 occurs and $F^*$ has been constructed we apply a modification of HAM of [4] in order to find a tour containing $F^*$ which only uses additional edges from $E_0$ and hence satisfies (3.2). We remove edges in $E_0 - F^*$ which are incident with $X_1$. Thus let

$$H = (V_n, (E_0 - \{e \in E_0: e \cap X_1 \neq \varnothing\}) \cup F^*).$$

Any hamilton cycle of $H$ will be a tour of length LB and optimal, if one exists.

**Phase 2 of ALGT.** We first check (see Lemma 4.3) that $H$ is connected and $\delta(H) \geqq 2$. If either condition fails we terminate unsuccessfully. Let now $F^*$ consist of paths $R_1, R_2, \cdots, R_s$ and let $R_i$ have endpoints $a_{2i-1}, a_{2i}$ and be of length $l_i$ for $i = 1, 2, \cdots, s$ and let $\mathrm{ENDF}^* = \{a_1, a_2, \cdots a_{2s}\}$. The following idea has been used by many authors (e.g. [1], [4], [6], [7], [9], [10], [11]): given a path $P = (v_1, v_2, \cdots, v_k)$ plus an edge $e = (v_k, v_i)$ where $1 \leqq i \leqq k-2$, we can create another path of length $k-1$

by deleting edge $\{v_i, v_{i+1}\}$ and adding $e$. Thus let

$$\text{ROTATE}(P, e) = (v_1, v_2, \cdots, v_i, v_k, v_{k-1}, \cdots, v_{i+1}), \text{ and}$$

$$\text{NEW}(P, e) = v_{i+1}.$$

$v_1$ is called the *fixed endpoint*, $v_k$ is called the *rotated endpoint* and $e$ is called the *rotation edge* of the rotation.

The algorithm we describe proceeds by a sequence of stages. At the beginning of the $k$th stage we have a path $P_k$ of length $k$, with endpoints $w_0 \leqq w_1$. We try to extend $P_k$ from $w_1$. If we fail but $\{w_0, w_1\} \in E(H)$ then connectivity tells us that we can find a longer path. Failing this, we do a sequence of rotations with $w_0$ as fixed endpoint which creates new paths that we can try to extend or close. We apply the same construction to all these paths and so on until either we have succeeded in obtaining a path of length $k+1$ or we have run out of paths to rotate. We then take this set of paths and treat each of them like $P_k$ but using $w_0$ as the first rotated endpoint.

We construct our sequence of paths in a "depth-first" manner. Suppose the "current" path is $Q$. One end $u$ will be kept fixed. Suppose its other end $v$ has neighbours $x_1, x_2, \cdots, x_p$ where $x_1 \in Q$. We replace $Q$ by ROTATE$(Q, \{v, x_1\})$ and continue with this "new" $Q$ *before* considering $x_2$ and and "old" $Q$. This will be done after backtracking. We also place a limit on the depth of the rotation sequence before backtracking. This is one way of keeping the algorithm polynomial, but its most important role is in the proof that HAM a.a. succeeds. We now give a formal description:

```
ALGORITHM HAM.
    begin
      let P_0 be the degenerate path consisting of v_1 = min (V_n − X_1) alone;
      k := 0;
L0    begin [stage k begins here]
        longerpathfound := false;
        let P_k have endpoints w_0, w_1 where w_0 ≦ w_1;
        storepaths := true;        [When this variable is true the new paths
                                   [generated by rotations during SEARCH
                                   [are stored for later.
        END_k := {w_1}             [We keep track of the endpoints of some
                                   [paths.
        SEARCH(P_k, w_0, P);       [Do rotations with w_0 as fixed endpoint
        if longerpathfound then    [i.e. if SEARCH has found a path
                                   [longer than P_k
        begin
L1        k := |P| = the number of edges in P;
          P_k := P;
          goto L0
        end else
        storepaths := false;       [No need to store paths now
        for w ∈ END_k do           [SEARCH(P_k, w_0, P) constructs a set of
                                   [paths {P(w_0, w): w ∈ END_k} where
                                   [P(w_0, w) has endpoints w_0 and w.
        begin
          SEARCH(P(w_0, w), w, P);
          if longerpathfound then goto L1
        end;
```

terminate unsuccessfully [successful termination with a hamilton [cycle containing $F^*$ occurs in SEARCH
*end*;
*procedure* SEARCH($Q, u, P$);
*begin*
let $v$ be the endpoint of $Q$ other than $u$;
*if* $v = a_j \in$ ENDF* *and* $Q \not\supseteq R_{\lceil j/2 \rceil}$ *then*
*begin*
$Q := Q + R_{\lceil j/2 \rceil}$; [catenation of paths
longerpathfound := true
*end else*
DFS($u, -, 0$)
*end*;
*procedure* DFS($v, y, t$); [vertex $y$ is such that [ROTATE($Q, \{v, y\}$) reverses the [rotation made immediately prior to [this call of DFS.

S1:
*if* $t > T = [\log n/(\log d - \log\log d)] + 1$ *then* return *else*
*begin*
Suppose that the edges incident with $v$, not contained in $Q$, and not incident with ENDF* are $\{v, x_1\}, \cdots \{v, x_p\}$ where $x_1 < x_2 < \cdots < x_{p-1}$ and $x_p = y$;
*if* $\exists i$ such that $x_i \notin Q$ *then*
*begin*
$j := \min\{i: x_i \notin Q\}$
$P := Q + \{v, x_j\}$; {*extension*}
longerpathfound := true
*end else*
*if* $u \in \{x_1, \cdots, x_{p-1}\}$ *then*
*begin* [cycle extension
let $C$ be the cycle $Q + \{u, v\}$;
*if* $C$ is a hamiltonian cycle *then* terminate HAM successfully
*else*
*begin*
starting from $u$, let $a$ be the first vertex along $Q$ which is adjacent to some vertex not in $C$; let $b$ be the lowest numbered neighbour of a not in $C$ and let $a_1$ and $a_2$ be the neighbours of $a$ on $C$ where $a_1 < a_2$, then
*if* $\{a, a_1\} \in F^*$ *then* $a_1 := a_2$;
$P := C + \{a, b\} - \{a, a_1\}$; longerpathfound := true
*end*
*end else*
*for* $i = 1$ *to* $p - 1$ *do* [$x_i \in Q - u$

S2:
*if not* longerpathfound *then*
*begin*
$Q :=$ ROTATE($Q, \{v, x_i\}$); $v' :=$ NEW($Q, \{v, x_i\}$);
*if* storepaths *and* $v' \notin \text{END}_k$ *then*
*begin*

$\mathrm{END}_k := \mathrm{END}_k \cup \{v'\}$;
$P(w_0, v') := Q$
*end*;
$\mathrm{DFS}(v', x_i, t+1)$
*end*;
*end*
*if* $t > 0$ *then* $Q := \mathrm{ROTATE}(Q, \{v, y\})$[backtrack to the 'parent path'
*end*;

The running time of HAM will be estimated later. We now introduce some notation used in the analysis of HAM. Suppose that HAM terminates unsuccessfully in stage $k$. Let

$$\mathrm{END}(H) = \{v \in \mathrm{END}_k \colon P(w_0, v) \text{ is obtained from } P_k \text{ by } t \text{ rotations, } 0 \leqq t \leqq T-1\}.$$

For $v \in \mathrm{END}(H)$ let

$$\mathrm{END}(H, v) = \{x \colon \exists \text{ a path } Q \text{ with } v, x \text{ as endpoints which is obtained from } P(w_0, v) \text{ by } t \text{ rotations with } v \text{ as fixed endpoint, where } 0 \leqq t \leqq T-1\}.$$

We note that,

(4.1) $H$ cannot contain an edge $\{x, y\}$ where $x \in \mathrm{END}(H)$ and $y \in \mathrm{END}(H, x)$.

We must of course show that HAM a.a. succeeds and (4.1) is a simple yet key observation. We see (Lemmas 4.1–4.5) that if HAM fails then $|\mathrm{END}(H)| \geqq n/16$ and $|\mathrm{END}(H, v)| \geqq n/16$ for $v \in \mathrm{END}(H)$ a.a. Given (4.1) we can see that this is unlikely. However the edges referred to in (4.1) are heavily conditioned and we use an argument of Fenner and Frieze [6] to deal with this problem.

The next lemma yields some important structural properties of $H$. It states that sets of vertices disjoint from $X_0$ have "large" neighbourhoods. This enables us to deduce that the number of path endpoints grows quickly from iteration to iteration.

LEMMA 4.1. *Assuming only that $d \to \infty$, the following properties hold a.a.*

(a) $S \subseteq V_n - X_0$, $|S| \leqq n/d$ *implies* $|N(S, G_0)| \geqq d|S|/6$.

(b) $S \subseteq V_n - X_0$, $n/d \leqq |S| \leqq n/2$ *implies* $|N(S, G_0)| \geqq |S|/2$.

*Proof.* (a) A similar result has been proved in several other papers but with different parameters, so we only give an outline here. (See [4], [9].) First consider the case where $s = |S| \leqq s_0 = n/d^5$. By considering $N(S, G_0)$ we see that if (a) fails then there exists a set $T$, $\lceil d/2 \rceil = t_0 \leqq t = |T| \leqq t_1 = \lfloor n/d^4 \rfloor$ which contains at least $3t/2$ edges. The probability of this event is no more than

$$\sum_{t=t_0}^{t_1} \binom{n}{t} \binom{\binom{t}{2}}{\lceil 3t/2 \rceil} p^{\lceil 3t/2 \rceil} \leqq \sum_{t=t_0}^{t_1} (te^5 d^3/27n)^{t/2} = o(1).$$

For $|S| \geqq s_0$ we ignore the condition $s \subseteq V_n - X_0$. The probability that $S$ exists is no more than

$$\begin{aligned} &\sum_{s=s_0}^{n/d} \binom{n}{s} \sum_{t=0}^{\lfloor sd/6 \rfloor} \binom{n-s}{t} (1-(1-p)^s)^t (1-p)^{s(n-s-t)} \\ &\leqq c_8 \sum_{s=s_0}^{n/d} \left(\frac{ne}{s}\right)^s \left(\frac{ne}{sd/6}\right)^{sd/6} (sp)^{sd/6} e^{-s(n-s-sd/6)p} \\ &\leqq c_8 \sum_{s=s_0}^{n/d} \left(\left(\frac{ne}{s}\right) \cdot (6e)^{d/6} \cdot e^{-d} \cdot e \cdot e^{d/6}\right)^s \\ &= o(1). \end{aligned}$$

(b) The probability that an $S$ exists contradicting (b) is no more than

$$\sum_{s=n/d}^{n/2} \binom{n}{s}\binom{n}{s/2}(1-p)^{s(n-3s/2)} = o(1)$$

as the largest term is $O((8d^{3/2}e^{-d/4})^{n/d})$. □

We can now "translate" these results to $H$:

LEMMA 4.2. *If* L2.1 ∩ L2.2 ∩ L4.1 *occurs, then for n large*

(a) $\delta(H) \geqq 2$,

(b) $S \subseteq V_n - X_1, |S| \leqq n/d$ *implies* $|N(S, H)| \geqq d|S|/7$ ($X_1$ *is defined prior to Lemma* 3.3),

(c) $S \subseteq V_n - X_1, |S| \geqq n/d$ *implies* $|N(S, H)| \geqq n/7 - |S|$,

(d) $H$ *is connected.*

*Proof.* Assume L2.1 ∩ L2.2 ∩ L4.1 occurs. Each vertex in $X_1$ has degree 2 in $H$. Each vertex in $V_n - X_1$ has degree at least $d/2$ in $G_0$. But from L2.1(c)

(4.2) no vertex is adjacent to $\lambda$ or more vertices in $X_1$.

(a) This follows immediately.

(b) This follows from L4.1(a) and (4.2).

$$(X_1 \supseteq X_0 \text{ and so } |N(S, H)| \geqq |N(S, G_0)| - |N(S, G_0) \cap X_1| \geqq d|S|/6 - \lambda|S|).$$

(c) This follows from (b) by taking $S' \subseteq S$ of size $\lfloor n/d \rfloor$ and using

$$|N(S, H)| \geqq |N(S', H)| - |S - S'|.$$

(d) Suppose there exists a partition of $V_n$ into 2 sets $A, B, |A| \leqq |B|$ such that $N(A, H) = \emptyset$. Now clearly $A \not\subseteq X_1$. If $|A - X_1| \leqq n/d$ then (b) and (4.2) with $S = A - X_1$ yield the contradiction, $d|A - X_1|/7 \leqq |A \cap X_1| \leqq \lambda|A - X_1|$. If $|A - X_1| > n/d$ we can use Lemma 4.1(b) and Lemma 3.3. □

We now consider what happens if HAM terminates unsuccessfully in stage $k$ of Phase 2.

Consider $P_k$, the initial path in stage $k$. It is the final path in a sequence $P^{(0)}, P^{(1)}, \cdots, P^{(M)} = P_k$ where $P^{(i+1)}$ is obtained from $P^{(i)}$ by a single extension, cycle extension or rotation.

Let $W(H) = \{\text{edges in } P^{(i)}, \ i = 1, \cdots, M\} \cup \{\{u, v\}$: HAM executes a cycle extension on a path with endpoints $u$ and $v\}$.

Let $X \subseteq E_0$ be *deletable* if:

(4.3a) No edge of $X$ is incident with a vertex of degree $\leqq d/2 + 1$ in $G_0$;

(4.3b) No pair of edges of $X$ are incident with a common vertex;

(4.3c) $X \cap W(H) = \emptyset$.

The colouring argument of [6] essentially shows that there is a.a. a deletable set $X$ of size $\approx d/2$. Being deletable means that HAM can ignore it up to the last stage and follow the same execution path. We can thus use (4.1) because $X$ is now "unconditioned"—this is only an intuitive interpretation of what follows.

For $X \subseteq E_0$ we define $H_X$ to be the graph obtained from $(V_n, E(K_n) - X)$ by applying Phase 1 but using $E_0 - X$ everywhere $E_0$ has been used previously.

*Remark* 4.1. For later purposes it is convenient to assume that the check for L2.1 ∩ L2.2 is not carried out in this case and that if $X = \emptyset$ then $H_X = G_0$. Thus $H_X$ is always defined but sometimes $H$ itself may not be. We note

LEMMA 4.3. *If* L2.1 ∩ L2.2 *occurs and X is deletable, then*

(a) $E(H_X) = E(H) - X$.

(b) *If HAM terminates unsuccessfully in stage k on input H, then HAM will also terminate unsuccessfully in stage k on $H_X$.*

*Proof.* (a) is clear and for (b) note that HAM will actually generate $P_k$ at the start of stage $k$ via the same sequence $P^{(i)}$, $i=1,\cdots,M$, as for $H$. □

We also note that similar calculations to those used for Lemma 2.1(a) yield

LEMMA 4.4. $|\{v\in V_n\colon d_0(v)\leqq d/2+1\}|\leqq ne^{-d/10}$*a.a.* □

We now show that failure a.a coincides with a large number of endpoints:

LEMMA 4.5. *Suppose* L2.1 ∩ L2.2 ∩ L4.2 ∩ L4.4 *occurs and HAM terminates unsuccessfully in stage k on H. Suppose $X\subseteq E_0$ is deletable. then for n large*

$$|\mathrm{END}(H_X)|\geqq n/16, \tag{4.4}$$

$$|\mathrm{END}(H_X,x)|\geqq n/16 \quad \textit{for } x\in \mathrm{END}(H_X). \tag{4.5}$$

*Proof.* Assume L2.1 ∩ L2.2 ∩ L4.2 ∩ L4.4 occurs. As noted in Lemma 4.3, HAM will generate the same $P_k$ at the start of stage $k$ as for $H$. HAM will fail to find a path of length $k+1$ in $H_X$ because HAM failed on $H$. Suppose $P_k$ has endpoints $w_0$ and $w_1$. Note that $w_0, w_1\notin X_1$. Let $S_t=\{v\colon\exists$ a path $Q_s$, $s=s(v)$, with endpoints $w_0$ and $v$, such that SEARCH $(P_k, w_0, P)$ constructs $Q_s$ using $t$ rotations$\}$. The rotations above are, of course, concerned with the execution of HAM on $H_X$.

We fix one value $s(v)$ for each $v$. We prove (4.4) by showing that

$$\left|\bigcup_{t=0}^{T-1} S_t\right|\geqq\frac{n}{16}. \tag{4.6}$$

To do this we show first that

$$|S_t|\leqq\frac{n}{d}\quad\text{implies } |S_{t+1}|\geqq\frac{d|S_t|}{15}. \tag{4.7}$$

Note that $|S_0|=1$ and let $t\geqq 0$ be such that $|S_t|\leqq n/d$. We use separate arguments for $t\leqq 20$ and $t>20$. Suppose first that $t\leqq 20$.

Consider now pairs $(v,w)$ where $v\in S_t$ and $w\in W(v)=N(\{v\},H_X)$. If $\{v,w\}$ is not an edge of $Q_{s(v)}$ let $x=x(v,w)=\mathrm{NEW}(Q_{s(v)},\{v,w\})$ and let $x=v$ otherwise. Let $Z=S_0\cup S_1\cup\cdots\cup S_t$. If $w\notin\mathrm{ENDF}^*$ and $\{v,w\}\notin Q_{s(v)}$ then $x\in S_{t+1}$. Thus

$$\begin{aligned}|S_{t+1}|&=|\{x=x(v,w)\colon v\in S_t,\{v,w\}\notin Q_{s(v)}\text{ and } w\in W(v)-\mathrm{ENDF}^*\}|\\&\geqq|\{x=x(v,w)\colon v\in S_t, w\in W(v)\}|-\lambda|S_t|\end{aligned} \tag{4.8}$$

using the fact that for each $v\in S_t$ there are at most $\lambda-1$ edges incident with $v$ and a vertex of ENDF* and there is the final edge of $Q_{s(v)}$. Now each $v\in S_t$ is adjacent to at most $t$ vertices $w$ for which $x=x(v,w)$ satisfies $\{w,x\}\notin P_k$ since $Q_{s(v)}$ contains at most $t$ edges not in $P_k$. For each $w\in N_t=N(S_t,H_X)$ choose $v=v(w)\in S_t$ such that $\{v,w\}\in E(H_x)$ and then let $x(w)=x(v,w)$.

It follows from the above that

$$|\{w\in N_t\colon\{w,x(w)\}\notin P_k\}|\leqq t|S_t|. \tag{4.9a}$$

Furthermore

(4.9b) There can be at most 2 distinct vertices $w, w'\in N_t$ such that $x(w)=x(w')$ and $\{x,w\},\{x,w'\}\in P_k$.

Thus, using (4.9)

$$|N_t| \leqq 2|\{x(w): w \in N_t \text{ and } \{x, x(w)\} \in P_k\}| + t|S_t|.$$

Hence, using (4.8) we have

$$\begin{aligned} |S_{t+1}| &\geqq |N_t|/2 - (\lambda + t/2)|S_t| \\ &\geqq (d/14 - (\lambda + t/2))|S_t| \end{aligned} \tag{4.10}$$

and (4.7) follows for $t \leqq 20$, in fact for $t \leqq \min\{d/105 - 2\lambda, T-1\}$. Suppose now that $T > t \geqq 21$. Let $X_W = \{x(w): w \in N_t\}$ and for $x \in X_W$ let $W^{-1}(x) = \{w \in N_t: x = x(w)\}$ and $S(x) = \{v(w): w \in W^{-1}(x)\}$. Let

$$Y = \bigcup_{\substack{x \in X_W \\ |W^{-1}(x)| \geqq 2}} W^{-1}(x)$$

and let

$$A = \{v \in V_n: v \text{ lies on some cycle of length 8 or less of } G_{2\mu}\}.$$

Then

$$|N_t| \leqq |X_W| + |A| + |Y - A|. \tag{4.11}$$

Now we can write

$$|Y - A| \leqq |Y_1| + |Y_2| + \cdots + |Y_p| \tag{4.12}$$

where

$$Y_i = W^{-1}(x_i) - A, \quad |W^{-1}(x_i)| \geqq 2 \quad \text{for } i = 1, 2, \cdots, p.$$

Let now $Z_i = \{v(w): w \in Y_i\}$. We note that

$$|Z_i| = |Y_i|, \qquad i = 1, 2, \cdots, p, \tag{4.13a}$$

$$|Z_i \cap Z_j| \leqq 1, \qquad 1 \leqq i < j \leqq p. \tag{4.13b}$$

It follows from (4.12) and (4.13) that

$$|Y - A| \leqq \frac{3|S_t|}{2}$$

and so from (4.11) and L2.1(e)

$$|X_W| \geqq |N_t| - 8(\log n)^9 - \frac{3|S_t|}{2}.$$

Using (4.8) we obtain

$$\begin{aligned} |S_{t+1}| &\geqq |X_W| - \lambda|S_t| \\ &\geqq |N_t| - (\lambda + 2)|S_t| \end{aligned}$$

as if $t > 20$, $|S_t| \geqq (\log n/169)^{10}$ by (4.7) for $t \leqq 20$ and the proof of (4.7) is complete.

We deduce that for some $\tau \leqq T-1$, $|S_\tau| \geqq n/d$. Choose a subset $S \subseteq S_\tau$ of size $\lfloor n/d \rfloor$ and apply the argument that proved (4.7) using $S$ in place of $S_\tau$ and $S_0 \cup S_1 \cup \cdots \cup S_{\tau-1} \cup S$ in place of $Z$. We deduce that $|S_{\tau+1}| \geqq d|S|/15$ and (4.4) follows.

To prove (4.5) consider $v \in \mathrm{END}(H_X)$ and redefine $S_t = \{x$: $\exists$ a path $Q$ with endpoints $v, x$ such that SEARCH$(P(w_0, v), v, P)$ constructs $Q$ using $t$ rotations$\}$.

Now apply the argument used to prove (4.7), using $Q$ in place of $P_k$, to prove (4.5). □

We now consider the size of $W$ and the running time of HAM:

LEMMA 4.6. (a) $|W| \leqq 3n \log n/(d \log d)$ *a.a.*

(b) *HAM runs in* $O(n^2(30 \log n)^T \log n/d) = O(n^{3+o(1)})$ *time a.a.*

*Proof.* Assume L2.1 ∩ L2.2 ∩ L4.2 ∩ L4.4 occurs.

(a) Let $Z = \{v \in V_n - X_0\colon c(v, w) > 2\mu$ for all $w \in X_0\}$. These vertices are not used at all in the construction of $F^*$ and

$$|Z| \geqq n - 30n \log n\, e^{-d/10}. \tag{4.14}$$

At the start of stage $k$, let $Z_k = Z - P_k$ and $P_k - P_{k-1} = \{v_k\}$. Note that $v_k$ is an endpoint of $P_k$. Now let

$$\xi_k = \begin{cases} 1 & \text{if } v_k \in Z \text{ and there is an edge of } E_0 \text{ incident with } v_k \text{ and } Z_k, \\ 0 & \text{otherwise} \end{cases}$$

(if there is no stage $k$, $\xi_k = 0$).

Then

$$\Pr(\xi_k = 1 | v_k \in Z) \geqq 1 - (1-p)^{|Z_k|} \tag{4.15}$$

using the FKG inequality [8], as the conditioning introduced only tells us that the degrees of $v_k$ and the members of $Z_k$ in $G_0$ are at least $d/2$. Furthermore

$$|W| \leqq \sum_{k=0}^{n-1} \xi_k + 2T \sum_{k=0}^{n-1} (1 - \xi_k). \tag{4.16}$$

Now, using (4.14) and (4.15) we have

$$\begin{aligned} E\left(\sum_{k=0}^{n-1} \xi_k\right) &\geqq \sum_{r=0}^{|Z|-1} (1 - (1-p)^r) \\ &\geqq |Z| - \frac{1}{p} \\ &\geqq n - 30n \log n\, e^{-d/10} - n/d \\ &\geqq n - \frac{5n}{4d} \quad \text{for } n \text{ large} \end{aligned}$$

and so

$$E\left(\sum_{k=0}^{n-1} (1 - \xi_k)\right) \leqq \frac{5n}{4d}.$$

As the lower bound in (4.15) holds whatever the algorithm's history up to stage $k$, we have

$$\sum_{k=0}^{n-1} (1 - \xi_k) \leqq \frac{3n}{2d} \quad \text{a.a.} \tag{4.17}$$

as it would if the $\xi_k$ were independent. (a) now follows from (4.16).

(b) Let us count the number of rotations in each stage. If $\xi_k = 1$ there are none, otherwise, using L2.1(f), each execution of SEARCH requires $O((30 \log n)^T)$ rotations. Now Angluin and Valiant [1] describe a data structure which enables one to do a rotation in $O(\log n)$ time. Thus, using (4.17), the time spent doing rotations is $O(n^2(30 \log n)^T \log n/d)$. This dominates the time for the other operations. □

We complete the proof of Theorem 1.1 by using the colouring argument of Fenner and Frieze [6]. Having generated our costs we colour the edges of $E_0$ blue and green. $e \in E$ is coloured blue with probability $1-1/n$ and green with probability $1/n$. These colourings are done independently. Let $E_{0b}$ denote the blue edges and $E_{0g}$ the green edges.

We define 2 events:

$$\mathscr{E}_1 = L \cap \{\text{HAM terminates unsuccessfully in Phase 2}\}$$

where $L = \text{L2.1} \cap \text{L2.2} \cap \text{L4.2} \cap \text{L4.4} \cap \text{L4.6}$,

$$\mathscr{E}_2 = \mathscr{E}_1 \cap \{X = E_{0g} \text{ is deletable}\}.$$

We note that if $\mathscr{E}_2$ occurs, then

(4.18) HAM terminates unsuccessfully on $H_X$, and

(4.19) $|\text{END}(H_X)| \geqq n/16$ and $|\text{END}(H_X, x)| \geqq n/16$ for all $x \in \text{END}(H_X)$.

(4.20) There does not exist $e = \{x, y\} \in X$ such that $x \in \text{END}(H_X)$ and $y \in \text{END}(H_X, x)$.

We prove that

$$\Pr(\mathscr{E}_2|\mathscr{E}_1) \geqq (1-o(1))(1-1/n)^{(3n \log n/d \log d)+n}, \tag{4.21}$$

$$\Pr(\mathscr{E}_2) \leqq (1-p/n)^{n^2/512}. \tag{4.22}$$

(The restricted version of) Theorem 1.1 follows immediately from (4.21), (4.22) and the fact that $\Pr(L) = 1-o(1)$.

*Proof of* (4.21). This follows from

$$\Pr(\mathscr{E}_2|E_i, i = 0, 1, \cdots B(n)-1) \geqq (1-o(1))(1-1/n)^{(3n \log n/d \log d)+n} \tag{4.23}$$

for all $E_0, E_1, \cdots$ compatible with the occurrence of $\mathscr{E}_1$. Once $E_0, E_1, \cdots$ are fixed then (4.23) follows easily from L4.4, L4.6, L2.1(f) and (4.2).

*Proof of* (4.22). Let $X_{0b} = \{v \in V_n\text{: degree of } v \text{ in } (V_n, E_{0b}) \leqq D/2\}$ and for $i \geqq 1$, $E_{ib} = \{e \in E_i;\ e \cap X_{0b} \neq \varnothing\}$. We prove (4.22) by showing that

$$\Pr(\mathscr{E}_2|E_{ib}, i = 0, 1, \cdots B(n)-1) \leqq (1-p/n)^{n^2/512}. \tag{4.24}$$

We note that if $E_{0b}, E_{1b}, \cdots$ are given then so is $H_X$. If either of (4.18) or (4.19) do not hold then the probability in (4.24) is zero. So assume that $E_{0b}, E_{1b}, \cdots$ are such that (4.18) and (4.19) hold. Consider next the set $Y = \{\{x, y\}: x \in \text{END}(H_X)$ and $y \in \text{END}(H_X, x)\}$. Note that $Y \cap E_{ib} = \varnothing$ for $i \geqq 0$. For each $e \in Y$ we have, given $E_{0b}, E_{1b}, \cdots$, $\Pr(e \in E_{0g}) = p/(n-d+1)$ and the events $e \in E_{0g}$ are independent. Clearly

$$\Pr(\mathscr{E}_2|E_{ib}, i = 0, 1, \cdots B(n)-1) \leqq \Pr((4.20)|E_{ib}, i = 0, 1, \cdots B(n)-1) \leqq (1-p/n)^{|Y|}$$

and (4.24) and the theorem follows. □

**5. The general case.**

**Change of assumptions.**

(1) $d = \omega \log \log n \leqq \log n/2$.

(2) $\lambda = \lceil \log n/\log \log n \rceil$.

Before proving Theorem 1.1 for $B(n) = o(n/\log \log n)$ we will indicate the places where the proof we have given so far breaks down.

*Problem* A. We have no problems with Phase 1 but because now $d < \lambda$ it is possible that we may find some vertices not in $X_1$ which lose all of their incident zero edges when we construct $H$. Our solution, essentially, is to absorb these vertices into $X_0$ beforehand.

*Problem* B. The estimate for the execution time of HAM has a term $(\log n)^{\log n/\log d}$ in it, which is not always polynomial for $d = \omega \log \log n$.

*Problem* C. Our final colouring argument relies on $|W|/n = o(d)$. The upper bound for $|W|$ in Lemma 4.6 is not strong enough to prove this.

Having given the problems, we proceed to their solution.

*Overview.* We define a new set $X_0$ which takes care of Problem A and we cover it with paths as before. We then make a small change to HAM which will take care of the other 2 problems.

We start with Problem C. Our solution is a little bizarre: given $E_0$ we randomly partition it into 2 sets $E_{0+} = \{\text{positive edges}\}$ and $E_{0-} = \{\text{negative edges}\}$. If $e \in E_0$ we place it in $E_{0+}$ with probability $\frac{1}{2}$ and in $E_{0-}$ with probability $\frac{1}{2}$. Naturally these placings are done independently of each other and so $(V_n, E_{0+}) = G_{n,p/2}$. We then alter algorithm HAM as follows:

(5.1a) The depth of rotation check $S_1$ is abandoned,

(5.1b) $S_2$ is changed to
*if* $e = \{v, x_i\} \in E_{0+}$ *and not* longerpathfound *and* $e$ has not been used previously as a rotation edge in the *current* execution of SEARCH *then*

Thus, foolishly it seems, we only use edges in $E_{0-}$ to extend paths and not to rotate them. On the other hand, in the final edge-colouring argument, we only colour edges in $E_{0-}$. Since we make $\leqq n$ cycle extensions we have $|E_0 - \cap W| \leqq 2n$ which gets around the $|W| = o(nd)$ problem as we will see later. One consequence of (5.1b) is that each execution of SEARCH requires no more than $|E_{0+}|$ rotations—each $e \in E_{0+}$ is used at most once to rotate. Now $|E_{0+}|$ is a binomial random variable with expectation $(n-1)d/4$ and so is a.a. less than $nd$. By Lemma 4.6 the number of stages that require rotations is a.a. $O(n/d)$. Putting this together we obtain the following lemma.

LEMMA 5.1. *HAM runs in* $O(n^3 \log n)$ *time a.a.* □

Thus Problem B is definitely solved, we have intimated that Problem C is solved and we now turn to Problem A: In this part of the paper $X_0$ will be redefined. So let

$$Y_0 = \{v \in V_n \colon \text{degree of } v \text{ in } (V_n, E_{0+}) \text{ is no more than } d/4\}.$$

It is convenient to define a set $X_{0,m}$ for positive integer $m \geqq 2$ and use $X_0 = X_{0,7}$ in the algorithm.

**Construction of $X_0 = X_{0,m}$.** Let $Y_0$ be as above and $Y_k = Y_0 \cup \{w_1, \cdots, w_k\}$ be defined inductively by $w_k = \min\{w \in V_n - Y_{k-1} \colon \exists$ independent edges $e_1, \cdots, e_m \in E_{2\mu+}$ (where $E_{k+} = E_k - E_{0-}$ for $k \geqq 0$) and edges $f_1, \cdots, f_m \in E_{2\mu+}$ such that

$$\begin{aligned} &e_i \cap Y_{k-1} \neq \emptyset, && i = 1, 2, \cdots, m, \\ &w_k \in f_i, && i = 1, 2, \cdots, m, \\ &|e_i \cap f_i| = 1, && i = 1, 2, \cdots, m\}. \end{aligned}$$

The process continues until $k = s+1$ is such that no $w_k$ exists. We take $X_0 = Y_s$. It is important to note that

(5.2) The set $X_0$ constructed is actually independent of the numbering of the vertices of $G_{2\mu+}$.

Before proving properties of $X_0$ we motivate its definition. Having constructed $X_0$ we shall carry out Phase 1 as previously and cover $X_0$ with a set of vertex disjoint paths $F^*$—we use the edges in $E_{0-}$ in this phase as normal. Consider now a vertex $v$ that is not an interior vertex of one of these paths and consider the edges of $E_{0+}$ which are incident with $v$ and with vertices on $F^*$. Let these edges be $\{v, w_i\}$, $i = 1, 2, \cdots, k$ and let $w_i$ be incident with edges $a_i, b_i$ where $a_i = b_i$ if $w_i$ is an endpoint. As $v \notin X_0$ the set $\{a_1, b_1, \cdots, a_k, b_k\}$ does not contain a subset of seven independent edges. Furthermore it is easy to see that $G_{2\mu+}$ a.a. contains no pair of triangles or quadrilaterals with a common vertex. (The expected number of such pairs is $O((\log n)^8/n)$.)

Defining $X_1$ as prior to Lemma 3.3 we have

(5.3) $\quad v \notin X_1$ is adjacent in $G_{2\mu+}$ to at most 7 vertices of $F^*$.

Thus our definition of $X_0$ will solve Problem A provided the equivalent of Lemma 2.1 can be shown, so that we can compute $F^*$ quickly. For (5.3) tells us that an endpoint not in $X_1$ has many ($\geqq d/4 - 7$) edges in $E_{0+}$ that can be used for rotations, if necessary.

We ought to be sure that $X_0$ can be computed quickly. But this is no problem as given $Y_k$, to test whether $w \notin Y_k$ can be added we have to look at the edges of $G_{2\mu+}$ which are incident with neighbours of $w$ and contain vertices of $Y_k$ and see whether they contain seven independent edges. We can clearly do all this within the $O(n^3 \log n)$ time bound.

We now give the equivalent of Lemma 2.1. Its proof is rather long and technical and we advise the reader to omit it at first reading.

LEMMA 5.2.

(a) $\Pr(|X_0| \geqq 16ne^{-d/10}) = o(e^{-d \log n})$.

(b) *$Z \subseteq V_n$, $|Z| \leqq 3\lambda$ and $|Z \cap X_0| \geqq \lambda$ implies $G_{2\mu+}[Z]$ is not connected a.a.*

(c) *$C$ is a cycle in $G_{2\mu+}$ implies $|C \cap X_0| \leqq |C|/10$ a.a.*

*Proof.* It is convenient to start by proving:

(5.4a) $\quad \Pr(|Y_0| \geqq ne^{-d/10}) = o(e^{-d \log n})$,

(5.4b) $\quad S \subseteq V_n$, $|S| \leqq N/(\log n)^6$ implies that $S$ contains less than $5|S|/4$ edges of $G_{2\mu+}$ a.a.,

(5.4c) $\quad \Pr(\exists S \subseteq V_n\colon ne^{-d} \leqq |S| \leqq n/(\log n)^6$ and $S$ contains at least $5|S|/4$ edges of $G_{2\mu+}) = o(e^{-d \log n})$,

(5.4d) $\quad S \subseteq V_n$, $|S| \leqq 4\lambda$ implies that $S$ contains at most $|S| + 4$ edges of $G_{2\mu+}$ a.a.

*Proof of* (5.4a). We use the Markov inequality

$$\Pr(X \geqq a) \leqq E(f(X))/f(a) \tag{5.5}$$

for any nonnegative random variable $X$, any $a \geqq 0$ and any nonnegative function $f$ (see, for example, Grimmett and Stirzaker [11, pp. 185–186]). We take $f(X) = \max\{0, X(X-1)(X-2)\cdots(X-t+1)\}$ where $t = \lceil(\log n)^2\rceil$ and $a = ne^{-d/10}$. Note that $f(|Y_0|)$ is the number of $t$-tuples of elements of $Y_0$. Hence

$$E(f(|Y_0|) \leqq t!\binom{n}{t} S(n-t, (2\mu+1)p, d/2)^t$$

$$\leqq n^t e^{-.15dt}.$$

The result goes through with a lot to spare on using (5.5).

*Proof of* (5.4b). The probability that $S$ exists contradicting (5.4b) is no more than

$$\sum_{k=4}^{n/(\log n)^6} \binom{n}{k} \binom{\binom{k}{2}}{\lceil 5k/4 \rceil} ((2\mu+1)p)^{(5k/4)}$$

$$\leqq \sum_{k=4}^{n/(\log n)^6} ((k/n)^{1/4}(\log n)^{5/4}(6/5)^{5/4}e^{9/4})^k$$

$$= O((\log n)^5/n).$$

*Proof of* (5.4c). The calculations that we have done support this.

*Proof of* (5.4d). The probability that $S$ exists contradicting (5.4c) is no more than

$$\sum_{k=5}^{4\lambda} \binom{n}{k} \binom{\binom{k}{2}}{k+5} ((2\mu+1)p)^{\lceil k+5 \rceil} = o(1)$$

after using $(\log n)^\lambda \leqq n \log n$.

*Proof proper.* (a) We shall prove that $\Pr(|X_{0,m}| \geqq 16ne^{-d/10}) = o(e^{-d \log n})$ for all constants $m \geqq 2$. Define $Y_0' = Y_0$ and $Y_k'$ inductively by $Y_k' = Y_{k-1}' \cup e_1 \cup f_1 \cup \cdots \cup e_m \cup f_m$ where the $e_i, f_i$ are those associated with $w_k$. We let $E_k' = \{e \in E_{2\mu}: e \subseteq Y_k'\}$. Then

(5.6a) $$|Y_0| + k \leqq |Y_k'| \leqq |Y_0| + (m+1)k, \text{ and}$$

(5.6b) $$|E_k'| \geqq (2m/(m+1))(|Y_k'| - |Y_0|)$$

follow easily by induction.

Suppose then that $|X_0| \geqq 16ne^{-d/10}$. Then since, by (5.4a), $|Y_0| \leqq ne^{-d/10}$ with probability $1 - o(e^{-d \log n})$ we can assume that there exists $k \geqq 15ne^{-d/10}$ such that $|Y_k| = 16ne^{-d/10}$. But then (5.6) implies

$$\frac{|Y_k'|}{|E_k'|} \leqq ((m+1)/2m)(|Y_k'|/(|Y_k'| - |Y_0|)$$

$$\leqq ((m+1)/2m)(16/15)$$

$$\leqq 4/5 \quad \text{as } m \geqq 2.$$

But this happens with probability $o(e^{-d \log n})$—(5.4c).

(b) If $G_{2\mu+}$ is connected, and [5] implies that it a.a. is, then the failure of (b) implies the existence of a set $Z \subseteq V_n$, $E \subseteq E_{2\mu+}$ satisfying

(i) $(Z, E)$ is a tree,

(ii) $|Z \cap X_0| \geqq \lambda$, and

(iii) $|Z| = 3\lambda$.

Let us first note that

(5.7) There a.a. exist no $Z$, $E$ satisfying (i), (ii)′ and (iii)

where

(ii)′ $|Z \cap Y_0| \geqq \lambda/14$.

The proof of this follows the lines of the proof of Lemma 2.1(c). Now the probability $\pi_0$ that $Z$ exists contradicting (b) is no more than

$$\binom{n}{3\lambda}\binom{3\lambda}{\lambda}(3\lambda)^{3\lambda-2}\pi_1 + \Pr(\bar{L})$$

where $\pi_1 = \max\{\Pr(A_T)\}$ and $L = \bigcap$(the events that have been proved so far to a.a. occur) and $A_T$ is the event $\{\{1, 2, \cdots, \lambda\} \subseteq X_0$, some fixed tree $T$ with vertices $\{1, 2, \cdots, 3\lambda\}$ exists in $G_{2\mu}\} \cap L$. We show that

$$\pi_1 \leqq e^{-\alpha\lambda d}((2\mu+1)p)^{3\lambda-1} \quad \text{for some constant } \alpha > 0. \tag{5.8}$$

Hence

$$\begin{aligned}\pi_0 &\leqq (ne/3\lambda)^{3\lambda} 3^{3\lambda} (3\lambda)^{3\lambda-2} (3 \log n/n)^{3\lambda-1} e^{-\alpha\lambda d}\\ &= O(e^{-\alpha\omega\lambda \log\log n/2})\\ &= o(1).\end{aligned}$$

It remains to prove (5.8). We do this by demonstrating that if $A_T$ occurs then so does some other event that we can show is unlikely. Assume $A_T$ occurs.

Now $T$ contains at most $6\lambda/7$ vertices of degree $\geqq 7$ in $T$ and hence at least $\lambda/7$ vertices of $X_0$ which have degree 6 or less in $T$. Now $|T \cap Y_0| \leqq \lambda/14$ by (5.7) and so there are at least $\lambda/14$ vertices in $X_0 - Y_0$ which have degree 6 or less in $T$. Let these vertices be denoted by $T_0$.

Now for $v \in T_0$ there are edges $e_i(v), f_i(v)$, $i = 1, 2, \cdots, 7$ such that the $e_i(v)$'s are independent and meet $X_0$, the $f_i(v)$'s meet $v$ and $|e_i(v) \cap f_i(v)| = 1$ for $i = 1, 2, \cdots, 7$. But at most 6 out of $f_1(v), \cdots, f_7(v)$ can be edges of $T$. Assume then that $f_1(v)$ is not an edge of $T$. At most 4 edges of the set $\{f_1(v)\colon v \in T_0\}$ are incident with 2 vertices of $T$ (5.4d) and so at least $\lambda/14 - 4$ vertices in $T_0$ are such that $f_1(v)$ is incident with a vertex outside $T$. Let $T_1$ denote these vertices of $T$.

Let $K = \{e_1(v)\colon v \in T_1\}$. (5.4d) implies that at most 8 of these edges intersect another edge in $K$. Thus by removing at most 8 edges we are left with an independent set of edges $K_1 \subseteq K$, $|K_1| \geqq \lambda/14 - 12$. Now (5.4d) implies that there are at most 4 vertices not in $T$ which are adjacent to 2 or more vertices of $T$ and no vertex not in $T$ is adjacent to 6 or more members of $T$. Also at most 4 edges not incident with $T$ can have both endpoints adjacent to members of $T$. Thus by removing at most 24 edges from $K_1$ we obtain a set $K_2 \subseteq K_1$, $|K_2| \geqq \lambda/15$ such that $e \in K_2$ implies (i) $e \cap X_0 \neq \varnothing$, (ii) $e \cap T = \varnothing$, and (iii) exactly one endpoint of $e$ is adjacent to $T$.

Let now $Z_1 = \{v \not\in X_0\colon \exists w \in T \text{ and } e \in K_2 \text{ such that } \{v, w\} \in E_{2\mu+} \text{ and } v \in e\}$ and $Z_2 = (\bigcup_{e \in K_2} e) - Z_1$. (5.4d) implies that

(5.9) no vertex in $V' = V_n - (T \cup Z_1)$ is within distance 2 in $G_{2\mu+}$ of 5 or more members of $T \cup Z_1$.

Let $G' = G_{2\mu+}[V']$. We claim that if we apply the process used to construct $X_0$ to $G'$, taking $m = 3$, then $X_{0,7}(G) - (T \cup Z_1) \subseteq X_{0,3}(G')$, where we have given the $X_{0,m}$ an extra parameter to indicate from which graph they are constructed.

If not let $X_{0,7}(G) = Y_0(G) \cup \{w_1, w_2, \cdots, w_s\}$ where the $w_i$ are in the order in which they were selected and let $t = \min\{j\colon w_j \not\in T \cup Z_1 \cup X_{0,3}(G')\}$. Let $e_i, f_i$, $i = 1, 2, \cdots, 7$ be the edges used to place $w_t$ in $X_0(G)$. (5.9) implies that at least 3 of these can be used to place $w_t$ in $X_0(G')$, a contradiction.

We have thus proved that if $T$ exists, $L$ occurs and $|X_0 \cap T| \geqq \lambda$ then there exist sets $Z_1, Z_2$ with

(5.10a) $|Z_1| \leqq |Z_2| \leqq 6\lambda$,

(5.10b) $\lambda/15 \leqq |Z_2|$,

(5.10c) $Z_2 \subseteq X_{0,3}(G')$,

(5.10d) Each vertex of $Z_1$ is adjacent to a vertex of $T$,

(5.10e) Each vertex of $Z_2$ is adjacent to a vertex of $T \cup Z_1$.

Note that once $T, Z_1, Z_2$ are given the event (5.10c) is independent of (5.10d) and (5.10e). Thus

$$\pi_1 \leqq \Pr(T \text{ exists and } \exists\, Z_1, Z_2 \text{ satisfying (5.10)})$$
$$\leqq \sum_{z_1=\lceil\lambda/15\rceil}^{6\lambda} \sum_{z_2=z_1}^{6\lambda} \binom{n}{z_1}\binom{n}{z_2}((2\mu+1)p)^{3\lambda-1}(3\lambda(2\mu+1)p)^{z_1}(9\lambda(2\mu+1)p)^{z_2}\pi_2$$

where $\pi_2 = \Pr(\mathrm{LOW} \subseteq X_{0,3}(G'))$ and LOW is the $\lceil\lambda/15\rceil$ lowest numbered vertices of $V(G')$ and so

$$\pi_1 \leqq ((2\mu+1)p)^{3\lambda-1}(\log n)^{24\lambda}\pi_2.$$

Bearing in mind (5.3), we see that by symmetry

$$\Pr(\mathrm{LOW} \subseteq X_{0,3}(G') \mid |X_{0,3}(G')| = n_1) = \binom{n_1}{\lceil\lambda/15\rceil} \Big/ \binom{|V(G')|}{\lceil\lambda/15\rceil} \tag{5.11}$$
$$\leqq (n_1/|V(G')|)^{\lceil\lambda/15\rceil}.$$

Now

$$n - 9\lambda \leqq |V(G')| \leqq n - 3\lambda$$

and so by (a) (with $n_1$ in place of $n$) we see that

$$\Pr(|X_{0,3}(G')| \geqq ne^{-d/10}) = o(e^{-d\log n}).$$

This combined with (5.11) yields (5.8) and (b).

The proof of (c) is similar and slightly easier as we have to deal with a cycle rather than an arbitrary tree. □

**6. Modifications for the second phase.** We must now prove that HAM a.a. succeeds. Let $H$ be defined as before. Assume that HAM fails on $H$ in stage $k$. Let $W(H)$ be defined as prior to (4.3). Because the edges of $E_{0-}$ are only used for extensions we have

$$|E_{0-} \cap W(H)| \leqq 2n. \tag{6.1}$$

Let $X \subseteq E_{0-}$ be *deletable* if

(6.2a) $X \cap W(H) = \varnothing$.

(6.2b) No edge of $X$ is incident with a vertex which is distance 2 or less from $X_0$ in $G_{2\mu+}$.

Note that if L4.1 ∩ L5.2 occurs then

(6.3) the number of vertices at distance 2 or less from $X_0$ in $G_{2\mu+}$ is $O(n(\log n)^2 e^{-d/20}) = o(n)$.

If $H_X$ is defined as in § 4 then Lemmas 4.1 and 4.3 continue to hold, with $d$ replaced by $d/2$.

LEMMA 6.1. *Assume* L4.1 ∩ L5.2 *occurs and HAM terminates in stage k on H. Suppose $X \subseteq E_{0-}$ is deletable. Then for n large*

$$|\mathrm{END}(H_X)| \geqq n/67. \tag{6.4}$$

$$|\mathrm{END}(H_X, x)| \geqq n/67 \textit{ for } x \in \mathrm{END}(H_X). \tag{6.5}$$

*Proof.* Assume L4.1 ∩ L5.2 occurs. Let $H_{X+} = (V_n, E(H_X) \cap E_{0+})$. We modify the argument of Posa [17] to show that $Z = \mathrm{END}(H_X)$ has relatively few neighbours in $H_{X+}$. In fact we show that

$$|N(Z, H_{X+}) - \mathrm{ENDF}^*| \leqq 2|Z| \tag{6.6}$$

(recall that ENDF* is the set of endpoints of the paths in $F^*$). This follows once we have shown that if $x \notin (Z \cup \text{ENDF}^*)$ is adjacent to some $z \in Z$ then $x$ is adjacent on $P_k$ to some $z' \in Z$. Suppose $x$, $z$ contradict this statement. As $x \notin Z$ there will not be an execution of DFS($x$, ?, ?) during SEARCH($P_k$, $w_0$, $P$). As $z \in Z$ there will be an execution of DFS($z$, ?, ?) and the edge $\{z, x\}$ will be a rotation edge. But at this time the neighbours of $x$ on $Q$ will be its neighbours on $P_k$. For if a $P_k$ edge gets deleted by a rotation then one of its endpoints is placed in $Z$. We have thus proved (6.6).

It follows from (5.3) and $Z \cap X_1 = \emptyset$ that

$$|N(Z, G_{0+})| \leqq 9|Z|. \tag{6.7}$$

Applying Lemma 4.1 (with $d$ replaced by $d/2$) shows that $|Z| > 2n/d$. Let $Z' \subseteq Z$ be of size $\lfloor 2n/d \rfloor$. Now

$$\begin{aligned} |N(Z, G_{0+})| &\geqq |N(Z', G_{0+})| - |Z| \\ &\geqq \lfloor 2n/d \rfloor (d/12) - |Z|. \end{aligned}$$

Formula (6.4) follows on using (6.7). Formula (6.5) is proved similarly. □

We can finish the proof using our colouring argument. Having generated our costs and chosen our partition of $E_0$, we randomly colour the edges in $E_{0-}$ blue with probability $1 - 1/n$ and green with probability $1/n$. Our 2 events are

$$\mathscr{E}_1 = L \cap \{\text{HAM terminates unsuccessfully in Phase 2}\}$$

where $L = \text{L4.1} \cap \text{L5.2}$ and

$$\mathscr{E}_2 = \mathscr{E}_1 \cap \{X = E_{0+g} \text{ is deletable}\}.$$

Proceeding as before we can prove

$$\Pr(\mathscr{E}_2 | \mathscr{E}_1) \geqq (1 - o(1))(1 - 1/n)^{2n} \tag{6.8}$$

and

$$\Pr(\mathscr{E}_2) \leqq (1 - p/2n)^{n^2/9000}. \tag{6.9}$$

(To fix $H_X$, fix $E_{0+}, E_{0-b}, E_1, \cdots, E_{2\mu}$, and the values of the edges incident with the corresponding $X_0$.)

Theorem 1.1 follows. □

We should say something about the restriction $B(n) = o(n/\log\log n)$). We could replace $\omega$ in the definition of $p$ by a suitably large constant. We have allowed this "constant" to go to infinity, as it makes our inequalities more obvious and only at marginal loss of generality. The problem with smaller values of $p$ is with Phase 1. For suppose that $B(n)$ grows faster than $n/\log\log n$. Heuristic arguments suggest that $G^*$ has $m = ne^{-\alpha d}$ edges ($\alpha$ constant) and average degree growing with $m$. Thus $G^*$ will almost certainly have a "giant" component, making the construction of $F^*$ (presumably) hard.

On the positive side HAM has no problem, even for $B(n) = \alpha n$, $\alpha > 0$ sufficiently small. We are thus now able to give polynomial time algorithms which construct the large cycles in Fenner and Frieze [6], [7] and Frieze [9]. We report on this in a further paper [10].

Note finally that the method extends naturally to finding $k$ edge disjoint tours of minimum weight, $k$ constant. We optimally cover $X_0$ with $k$ sets of vertex disjoint paths and apply HAM $k$ times.

## REFERENCES

[1] D. ANGLUIN AND L. VALIANT, *Fast probabilistic algorithms for hamilton circuits and matchings*, J. Comput. System Sci., 18 (1979), pp. 155–193.

[2] M. W. BERN, E. L. LAWLER AND A. L. WONG, *Why certain subgraph computations require only linear time*, Proc. 26th IEEE Symposium on Foundations of Computer Science, 1985, pp. 117–125.

[3] B. BOLLOBÁS, *Random Graphs*, Academic Press, New York, 1985.

[4] B. BOLLOBÁS, T. I. FENNER AND A. M. FRIEZE, *An algorithm for finding hamilton cycles in random graphs*, Proc. 17th Annual ACM Symposium on Theory of Computing, 1985, pp. 430–439.

[5] P. ERDÖS AND A. RÉNYI, *On random graphs* 1, Publ. Math. Debrecen, 6 (1959), pp. 290–297.

[6] T. I. FENNER AND A. M. FRIEZE, *On the existence of hamilton cycles in a class of random graphs*, Discrete Math., 45 (1983), pp. 301–305.

[7] ———, *Hamilton cycles in random regular graphs*, J. Combin. Theory Ser. B, 37 (1984), pp. 103–112.

[8] C. M. FORTUIN, P. W. KASTELYN AND J. GINIBRE, *Correlation inequalities on some partially ordered sets*, Comm. Math. Phys., 22 (1971), pp. 89–103.

[9] A. M. FRIEZE, *On large matchings and cycles in sparse random graphs*, Discrete Math., 59 (1986), pp. 243–256.

[10] ———, *Finding hamilton cycles in sparse random graphs*, J. Combin. Theory, to appear.

[11] G. GRIMMETT AND D. STIRZAKER, *Probability and Random Processes*, Clarendon Press, Oxford, 1982.

[12] Y. GUREVICH AND S. SHELAH, *Expected computation time for hamiltonian path problem and clique problem*, to appear.

[13] J. H. HALTON AND R. TERADA, *A fast algorithm for the euclidean travelling salesman problem: optimal with probability* 1, this Journal, 11 (1982), pp. 28–46.

[14] R. M. KARP, *Probabilistic analysis of partitioning algorithms for the travelling salesman problem in the plane*, Math. Oper. Res., 2 (1977), pp. 209–224.

[15] ———, *A patching algorithm for the non-symmetric travelling salesman problem*, this Journal, 8 (1979), pp. 561–573.

[16] E. L. LAWLER, J. K. LENSTRA, A. H. G. RINOOY KAN AND D. B. SHMOYS, *The Travelling Salesman Problem*, John Wiley, New York,

[17] L. POSA, *Hamilton circuits in random graphs*, Discrete Math., 14 (1976), pp. 359–364.