

ON TESTING HAMILTONICITY OF GRAPHS

ALEXANDER BARVINOK

July 15, 2014

ABSTRACT. Let us fix a function $f(n) = o(n \ln n)$ and reals $0 \leq \alpha < \beta \leq 1$. We present a polynomial time algorithm which, given a directed graph G with n vertices, decides either that one can add at most βn new edges to G so that G acquires a Hamiltonian circuit or that one cannot add αn or fewer new edges to G so that G acquires at least $e^{-f(n)} n!$ Hamiltonian circuits, or both.

1. INTRODUCTION AND MAIN RESULTS

Let $G = (V, E)$ be a directed graph with set V of vertices and set E of edges. A *Hamiltonian circuit* is a closed walk $i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_n \rightarrow i_1$ that visits every vertex of G exactly once. It is a classical NP-complete problem to determine whether a given directed graph contains a Hamiltonian circuit (in which case the graph is called *Hamiltonian*). In what follows, n denotes the number of vertices of the graph, $n = |V|$.

The following version of the problem is also known to be NP-complete: Given $0 < \beta < 1$, is it true that one can add at most βn new edges to a given directed graph with n vertices so that the graph becomes Hamiltonian? In fact, for any fixed $\beta < 1/320$, the problem is NP-complete, see [PY93] and [EK06].

Anastasios Sidiropoulos pointed out to the author that testing Hamiltonicity does not become any easier if we are promised that should the directed graph be Hamiltonian, it contains at least $\exp\{-n^\epsilon\} n!$ Hamiltonian circuits for some fixed $\epsilon > 0$. Indeed, let G be a given directed graph with m vertices. Let us choose $k > 2/\epsilon$ and construct a new directed graph \widehat{G} by attaching a complete directed graph with m^k vertices by two edges to two selected vertices u and v of G . Hence the new graph \widehat{G} with $n = m + m^k$ vertices contains at least $(m^k - 2)!$ Hamiltonian circuits if and only if G contains a Hamiltonian path with endpoints u and v . If there is no such path in G then \widehat{G} contains no Hamiltonian circuits.

1991 *Mathematics Subject Classification.* 05C45, 15A15, 68E10.

Key words and phrases. permanent, graph, Hamiltonian circuit, algorithm.

This research was partially supported by NSF Grant DMS 0856640.

Let us choose a function $f(n) = o(n \ln n)$ and fix two numbers $0 \leq \alpha < \beta \leq 1$. We present a polynomial time algorithm, which, given a directed graph G with n vertices, outputs at least one of the following two statements a) and b):

a) one can add at most βn new edges to G so that G acquires a Hamiltonian circuit;

b) one cannot add αn or fewer new edges to G so that G acquires at least $e^{-f(n)}n!$ Hamiltonian circuits.

For example, confronted with two directed graphs on n vertices one of which contains at least $10^{-3n}n!$ Hamiltonian circuits and the other doesn't become Hamiltonian unless more than $10^{-3}n$ new edges added to the graph, our algorithm will be able to tell which graph is which in polynomial time. On the other hand, testing whether one needs to add at least $10^{-3}n$ new edges to a given directed graph on n vertices so that the graph becomes Hamiltonian is an NP-hard problem and testing whether a given directed graph on n vertices contains at least $10^{-3n}n!$ Hamiltonian circuits is also an NP-hard problem.

It may happen though that while the statements a) and b) are both true, the algorithm outputs only one of them.

We note that even if we are told that the graph contains at least $e^{-f(n)}n!$ Hamiltonian circuits, it is not obvious how to construct any of the circuits efficiently (deterministically or probabilistically). We also note some vague similarity with property testing questions [G+98].

Our algorithm is based on computing permanents and their Hamiltonian versions.

(1.1) Permanents and Hamiltonian permanents. Let $A = (a_{ij})$ be an $n \times n$ real matrix. The *permanent* of A is defined as

$$\text{per } A = \sum_{\sigma \in S_n} \prod_{i=1}^n a_{i\sigma(i)},$$

where the sum is taken over the symmetric group S_n of permutations of the set $\{1, \dots, n\}$. As is known, the problem of computing the permanent exactly is $\#P$ -hard, even if the entries of A are restricted to be 0 and 1 [Va79]. For non-negative matrices a fully polynomial randomized approximation scheme is available [J+04]. We, however, are interested in computing permanents of a rather restricted class of matrices. Namely, let us suppose that

$$(1.1.1) \quad \frac{1}{n^{0.1}} \leq a_{ij} \leq 1 \quad \text{for all } i, j.$$

Then a version of the scaling algorithm of [L+00], see also [BS11], approximates $\text{per } A$ in polynomial in n time within an $O(\exp\{n^{0.35}\})$ factor. The algorithm is deterministic and easy to implement. We review the algorithm in Section 3.

Let $H_n \subset S_n$ be the subset of $(n-1)!$ permutations consisting of a single cycle. We define the *Hamiltonian permanent* by

$$\text{ham } A = \sum_{\sigma \in H_n} \prod_{i=1}^n a_{i\sigma(i)}.$$

If A is a 0-1 matrix then it is an NP-complete problem to tell $\text{ham } A$ from 0, as the problem is equivalent to testing Hamiltonicity of the directed graph with adjacency matrix A . It turns out, however, that when (1.1.1) holds, $\text{per } A$ and $\text{ham } A$ have the same logarithmic order.

(1.2) Theorem. *Let $A = (a_{ij})$ be an $n \times n$ matrix such that*

$$\epsilon \leq a_{ij} \leq 1$$

for some $\epsilon > 0$ and all i, j . Let

$$r = \left\lfloor \frac{4 \ln n}{\epsilon^2} \right\rfloor + 6.$$

Then

$$\frac{1}{2^r} \left(\frac{\epsilon}{n} \right)^r \text{per } A \leq \text{ham } A \leq \text{per } A.$$

In particular, if we choose $\epsilon = n^{-0.1}$ then $\text{per } A$ approximates $\text{ham } A$ within an $O(\exp\{n^{0.3}\})$ factor.

We prove Theorem 1.2 in Section 2.

In a different setting, the relation between the permanent and Hamiltonian permanent of the adjacency matrix of a k -regular graph was used in [Vi12] while the first use of permanents to bound the number of Hamiltonian circuits in tournaments goes back to [A90].

(1.3) Testing Hamiltonicity of graphs. Let us fix a function $f(n) = o(n \ln n)$ and real numbers $0 \leq \alpha < \beta \leq 1$. Given a directed graph $G = (V, E)$, we identify $V = \{1, \dots, n\}$ and construct an $n \times n$ matrix $A = A(G)$, $A = (a_{ij})$, as follows:

$$a_{ij} = \begin{cases} 1 & \text{if } (i \rightarrow j) \in E \\ n^{-0.1} & \text{otherwise.} \end{cases}$$

Using Theorem 1.2 and the algorithm of Section 3, we compute $\text{ham } A$ within a factor of $O(\exp\{n^{0.4}\})$. If G does not become Hamiltonian unless more than βn new edges are added to G , then

$$(1.3.1) \quad \text{ham } A \leq n^{-0.1\beta n} n!.$$

If, however, one can add αn or fewer edges to G so that G acquires at least $e^{-f(n)}n!$ Hamiltonian circuits then

$$(1.3.2) \quad \text{ham } A \geq n^{-0.1\alpha n} e^{-f(n)} n!.$$

The ratio of the lower bound in (1.3.2) to the upper bound in (1.3.1) is

$$n^{0.1(\beta-\alpha)n} e^{-f(n)} = n^{0.1(\beta-\alpha)n - o(n)}$$

and hence for all sufficiently large $n \geq N(f, \alpha, \beta)$ we will be able to conclude that either (1.3.1) is violated, or (1.3.2) is violated, or both are violated. For finitely many $n < N(f, \alpha, \beta)$ one can enumerate the Hamiltonian circuits in G directly, in constant time. If (1.3.1) is violated, then the statement

a) one can add at most βn new edges to G so that G becomes Hamiltonian is true and if (1.3.2) is violated, then the statement

b) one cannot add αn or fewer new edges to G so that G acquires at least $e^{-f(n)}n!$ Hamiltonian circuits

is true. Hence the algorithm will output at least one of the statements a) and b). On the other hand, if a) is true we may still have (1.3.1) and if b) is true we may still have (1.3.2). Hence it may happen that although a) and b) are both true, the algorithm outputs only one of a) and b).

2. PROOF OF THEOREM 1.2

Let us fix an $n \times n$ matrix $A = (a_{ij})$ such that

$$\epsilon \leq a_{ij} \leq 1$$

for some $\epsilon > 0$ and all i, j . We consider the symmetric group S_n as a finite probability space, where we let

$$\mathbf{P}(\sigma) = (\text{per } A)^{-1} \left(\prod_{i=1}^n a_{i\sigma(i)} \right) \quad \text{for } \sigma \in S_n.$$

(2.1) Lemma. *Let us define random variables*

$$l_i : S_n \longrightarrow \mathbb{R} \quad \text{for } i = 1, \dots, n,$$

where $l_i(\sigma)$ is the length of the cycle of $\sigma \in S_n$ containing i . Then

$$\mathbf{P}(\sigma \in S_n : l_i(\sigma) = m) \leq \frac{1}{\epsilon^2(n-m)} \quad \text{for } i = 1, \dots, n$$

and $m = 1, \dots, n-1$.

Proof. Without loss of generality, we assume that $i = 1$. With the set of permutations $\sigma \in S_n$ such that $l_1(\sigma) = m$ we associate a set $\Sigma \subset S_n$. Each permutation σ with $l_1(\sigma) = m$ contributes $n - m$ permutations τ to Σ as follows.

We write the cycle of σ containing 1 as

$$1 = j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_m \rightarrow 1.$$

We pick any of the $n - m$ numbers, say r , not in the cycle. We write the cycle containing r as

$$r = j_{m+1} \rightarrow j_{m+2} \rightarrow \dots \rightarrow j_{m+k} \rightarrow r$$

and produce a permutation $\tau \in \Sigma$ by merging the two cycles together:

$$1 = j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_m \rightarrow r = j_{m+1} \rightarrow j_{m+2} \rightarrow \dots \rightarrow j_{m+k} \rightarrow 1.$$

Since $\epsilon \leq a_{ij} \leq 1$, we have

$$(2.1.1) \quad \mathbf{P}(\tau) \geq \epsilon^2 \mathbf{P}(\sigma).$$

Thus each permutation σ contributes $n - m$ permutations $\tau \in \Sigma$ depending on the choice of r . The set Σ consists of all permutations τ thus obtained from all permutations σ with $l_1(\sigma) = m$.

We observe that every $\tau \in \Sigma$ is obtained from a unique permutation σ . To reconstruct σ from τ , we choose the cycle of τ containing 1, write it as

$$1 \rightarrow j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_{m+k} \rightarrow 1$$

for some $k > 0$ and split it into the two cycles,

$$1 \rightarrow j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_m \rightarrow 1 \quad \text{and} \quad j_{m+1} \rightarrow \dots \rightarrow j_{m+k} \rightarrow j_{m+1}.$$

Since every permutation $\sigma \in S_n$ with $l_1(\sigma) = m$ gives rise to $n - m$ permutations $\tau \in \Sigma$, using (2.1.1) we obtain

$$\mathbf{P}\left(\sigma \in S_n : l_1(\sigma) = m\right) \leq \frac{1}{\epsilon^2(n - m)} \mathbf{P}\left(\tau \in \Sigma\right) \leq \frac{1}{\epsilon^2(n - m)},$$

as desired. □

(2.2) Corollary. *We have*

$$\mathbf{E}(l_i^{-1}) \leq \frac{2 \ln n}{n \epsilon^2} + \frac{3}{5n} \quad \text{for } i = 1, \dots, n.$$

Proof. We have

$$\begin{aligned}
\mathbf{E} (l_i^{-1}) &= \sum_{m=1}^n \frac{1}{m} \mathbf{P} (\sigma : l_i(\sigma) = m) \\
&= \sum_{m: m \leq n/3} \frac{1}{m} \mathbf{P} (\sigma : l_i(\sigma) = m) + \sum_{m: m > n/3} \frac{1}{m} \mathbf{P} (\sigma : l_i(\sigma) = m) \\
&\leq \frac{2}{n\epsilon^2} \sum_{m: m \leq n/3} \frac{1}{m} + \frac{3}{n} \sum_{m: m > n/3} \mathbf{P} (\sigma : l_i(\sigma) = m) \\
&\leq \frac{2 \ln n}{n\epsilon^2} + \frac{3}{n}.
\end{aligned}$$

□

(2.3) Lemma. Let $c(\sigma)$ be the number of cycles in a permutation $\sigma \in S_n$. Then

$$\sum_{\substack{\sigma \in S_n: \\ c(\sigma) < 4\epsilon^{-2} \ln n + 6}} \prod_{i=1}^n a_{i\sigma(i)} \geq \frac{1}{2} \text{per } A.$$

Proof. We consider $c : S_n \rightarrow \mathbb{R}$ as a random variable. Let l_i be the random variables of Lemma 2.1 and Corollary 2.2. Then

$$c(\sigma) = \sum_{i=1}^n l_i^{-1}(\sigma),$$

since the sum of $l_i^{-1}(\sigma)$ over all i in a particular cycle of σ is equal to 1. Therefore, by Corollary 2.2,

$$\mathbf{E} c = \sum_{i=1}^n \mathbf{E} (l_i^{-1}) \leq \frac{2 \ln n}{\epsilon^2} + 3.$$

By the Markov inequality

$$\mathbf{P} \left(\sigma \in S_n : c(\sigma) \geq \frac{4 \ln n}{\epsilon^2} + 6 \right) \leq \frac{1}{2}$$

and the proof follows. □

(2.4) Proof of Theorem 1.2. Clearly,

$$\text{ham } A \leq \text{per } A.$$

Let

$$X = \left\{ \sigma \in S_n : c(\sigma) < 4\epsilon^{-2} \ln n + 6 \right\},$$

so by Lemma 2.3

$$(2.4.1) \quad \sum_{\sigma \in X} \prod_{i=1}^n a_{i\sigma(i)} \geq \frac{1}{2} \text{per } A.$$

Let

$$r = \left\lfloor \frac{4 \ln n}{\epsilon^2} + 6 \right\rfloor$$

and for $k = 1, \dots, r$ let

$$X_k = \left\{ \sigma \in S_n : c(\sigma) = k \right\}$$

and let $H_n \subset S_n$ be the set of Hamiltonian cycles.

With every permutation $\sigma \in X_k$, we associate a set $T(\sigma) \subset H_n$ as follows. We order the k cycles of a permutation $\sigma \in X_k$ arbitrarily, choose arbitrarily elements j_1, \dots, j_k , one in each cycle, and patch the cycles into a Hamiltonian cycle τ by replacing the edges

$$(2.4.2) \quad i_1 \rightarrow j_1, \quad i_2 \rightarrow j_2, \quad \dots, \quad i_k \rightarrow j_k$$

in σ by the edges

$$(2.4.3) \quad i_1 \rightarrow j_2, \quad i_2 \rightarrow j_3, \quad \dots, \quad i_k \rightarrow j_1.$$

Because of $\epsilon \leq a_{ij} \leq 1$, we have

$$\mathbf{P}(\tau) \geq \epsilon^k \mathbf{P}(\sigma) \geq \epsilon^r \mathbf{P}(\sigma).$$

Let $T(\sigma)$ be the set of all possible Hamiltonian cycles τ obtained in this way.

We observe that any $\tau \in H_n$ is obtained from at least one and at most $\binom{n}{k}$ permutations $\sigma \in X_k$, since for any choice of j_1, \dots, j_k in τ ,

$$\tau = j_1 \rightarrow \dots \rightarrow i_1 \rightarrow j_2 \rightarrow \dots \rightarrow i_2 \rightarrow j_3 \rightarrow \dots \rightarrow i_{k-1} \rightarrow j_k \rightarrow \dots \rightarrow i_k \rightarrow j_1$$

corresponds to at most one permutation $\sigma \in X_k$ in which the edges (2.4.3) are replaced by the edges (2.4.2). This proves that

$$n^r \mathbf{P}(H_n) \geq \binom{n}{k} \mathbf{P}(H_n) \geq \epsilon^r \mathbf{P}(X_k)$$

and the proof follows by (2.4.1). □

3. ESTIMATING THE PERMANENT

(3.1) Scaling a matrix to doubly stochastic. An $n \times n$ matrix $B = (b_{ij})$ is called *doubly stochastic* if

$$\sum_{j=1}^n b_{ij} = 1 \quad \text{for } i = 1, \dots, n, \quad \sum_{i=1}^n b_{ij} = 1 \quad \text{for } j = 1, \dots, n$$

and

$$b_{ij} \geq 0 \quad \text{for all } i, j.$$

Let $A = (a_{ij})$ be $n \times n$ matrix such that $a_{ij} > 0$ for all i, j . Sinkhorn's Theorem [Si64] states that there exists a doubly stochastic matrix $B = (b_{ij})$ and positive numbers $\lambda_1, \dots, \lambda_n$ and μ_1, \dots, μ_n such that

$$a_{ij} = \lambda_i \mu_j b_{ij} \quad \text{for all } i, j.$$

The matrix B is unique whereas the numbers λ_i and μ_j are unique up to a rescaling $\lambda_i := \lambda_i \tau$, $\mu_j := \mu_j \tau^{-1}$ for some $\tau > 0$ and all i, j . We say that B is the *doubly stochastic scaling* of A . We observe that

$$(3.1.1) \quad \text{per } A = \left(\prod_{i=1}^n \lambda_i \right) \left(\prod_{j=1}^n \mu_j \right) \text{per } B.$$

The problem of computing the product of multipliers λ_i and μ_j in (3.1.1) within relative error $\epsilon > 0$ is a convex programming problem which can be solved in time polynomial in n and $\ln(1/\epsilon)$ [KK96].

In [L+00], formula (3.1.1) was used to compute permanents in strongly polynomial time within a factor of e^{-n} . We need more precision for matrices A with more structure.

(3.2) Scaling of δ -balanced matrices. Let $A = (a_{ij})$ be an $n \times n$ positive matrix and let $0 < \delta < 1$. We say that A is *δ -balanced* if

$$\delta a_{kl} \leq a_{ij} \quad \text{for all } i, j, k, l.$$

If B is a doubly stochastic scaling of a δ -balanced matrix A then B is δ^3 -balanced (Lemma 4.1 of [BS11]).

We also need the following estimate (Lemma 5.1 of [B+10]).

(3.3) Lemma. *Let B be an $n \times n$ doubly stochastic matrix such that*

$$b_{ij} \leq \frac{\alpha}{n} \quad \text{for all } i, j$$

and some $\alpha > 1$ Then

$$\frac{n!}{n^n} \leq \text{per } B \leq \frac{n!}{n^n} (2\pi n)^{\alpha/2} e^{\alpha^2/12n}.$$

The lower bound is, of course, the famous van der Waerden bound for permanents of doubly stochastic matrices, cf. [L+00]. In particular, if B is δ^3 -balanced then

$$(3.2.1) \quad \frac{n!}{n^n} \leq \text{per } B \leq \frac{n!}{n^n} (2\pi n)^{1/2\delta^3} e^{1/12\delta^6 n}.$$

The matrix $A = A(G)$ of Section 1.3 is δ -balanced for $\delta = n^{-0.1}$ and hence its doubly stochastic scaling B is $n^{-0.3}$ -balanced. Therefore, formula (3.2.1) approximates $\text{per } B$ within an $O(\exp\{n^{0.35}\})$ factor. Hence the scaling algorithm of Section 3.1 approximates $\text{per } A$ within an $O(\exp\{n^{0.4}\})$ factor.

ACKNOWLEDGMENT

The author is grateful to Anastasios Sidiropoulos for a valuable conversation.

REFERENCES

- [A90] N. Alon, *The maximum number of Hamiltonian paths in tournaments*, *Combinatorica* **10** (1990), no. 4, 319–324.
- [B+10] A. Barvinok, Z. Luria, A. Samorodnitsky, and A. Yong, *An approximation algorithm for counting contingency tables*, *Random Structures & Algorithms* **37** (2010), no. 1, 25–66.
- [BS11] A. Barvinok and A. Samorodnitsky, *Computing the partition function for perfect matchings in a hypergraph*, *Combinatorics, Probability and Computing* **20** (2011), no. 6, 815–835.
- [EK06] L. Engebretsen and M. Karpinski, *TSP with bounded metrics*, *Journal of Computer and System Sciences* **72** (2006), no. 4, 509–546.
- [G+98] O. Goldreich, S. Goldwasser and D. Ron, *Property testing and its connection to learning and approximation*, *Journal of the ACM* **45** (1998), no. 4, 653–750.
- [J+04] M. Jerrum, A. Sinclair and E. Vigoda, *A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries*, *Journal of the ACM* **51** (2004), no. 4, 671–697.
- [KK96] B. Kalantari and L. Khachiyan, *On the complexity of nonnegative-matrix scaling*, *Linear Algebra and its Applications* **240** (1996), 87–103.
- [L+00] N. Linial, A. Samorodnitsky and A. Wigderson, *A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents*, *Combinatorica* **20** (2000), no. 4, 545–568.
- [PY93] C.H. Papadimitriou and M. Yannakakis, *The traveling salesman problem with distances one and two*, *Mathematics of Operations Research* **18** (1993), no. 1, 1–11.
- [Si64] R. Sinkhorn, *A relationship between arbitrary positive matrices and doubly stochastic matrices*, *Annals of Mathematical Statistics* **35** (1964), 876–879.
- [Va79] L.G. Valiant, *The complexity of computing the permanent*, *Theoretical Computer Science* **8** (1979), no. 2, 189–201.
- [Vi12] N.K. Vishnoi, *A permanent approach to the traveling salesman problem*, 53rd IEEE Annual Symposium on Foundations of Computer Science (FOCS) (2012), 76–80.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109-1043, USA

E-mail address: barvinok@umich.edu