A polynomial-time approximation algorithm for the number of k-matchings in bipartite graphs

Shmuel Friedland and Daniel Levy
Department of Mathematics, Statistics, and Computer Science,
University of Illinois at Chicago
Chicago, Illinois 60607-7045, USA

July 28, 2006

Abstract

We show that the number of k-matching in a given undirected graph G is equal to the number of perfect matching of the corresponding graph G_k on an even number of vertices divided by a suitable factor. If G is bipartite then one can construct a bipartite G_k. For bipartite graphs this result implies that the number of k-matching has a polynomial-time approximation algorithm. The above results are extended to permanents and hafnians of corresponding matrices.

2000 Mathematics Subject Classification: 05A15, 05C70, 68A10.
Keywords and phrases: Perfect matchings, k-matchings, permanents, hafnians, polynomial-time approximation algorithm.

1 Introduction

Let $G = (V, E)$ be an undirected graph, (with no self-loops), on the set of vertices V and the set of edges E. A set of edges $M \subseteq E$ is called a *matching* if no two distinct edges $e_1, e_2 \in M$ have a common vertex. M is called a *k-matching* if $\#M = k$. For $k \in \mathbb{N}$ let $\mathcal{M}_k(G)$ be the set of k-matchings in G. ($\mathcal{M}_k(G) = \emptyset$ for $k > \lfloor \frac{\#V}{2} \rfloor$.) If $\#V = 2n$ is even then an n-matching is called a *perfect matching.* $\phi(k, G) := \#\mathcal{M}_k(G)$ is number of k-matchings, and let $\phi(0, G) := 1$. Then $\Phi(x, G) := \sum_{k=0}^{\infty} \phi(k, G)x^k$ is the *matching polynomial* of G. It is known that a nonconstant matching polynomial of G has only real negative roots [6].

Let G be a bipartite graph, i.e., $V = V_1 \cup V_2$ and $E \subset V_1 \times V_2$. In the special case of a bipartite graph where $n = \#V_1 = \#V_2$, it is well known that $\phi(n, G)$ is given as $\operatorname{perm} B(G)$, the permanent of the incidence matrix $B(G)$ of the bipartite graph G. It was shown by Valiant that the computation of the permanent of a $(0,1)$ matrix is $\#\mathbf{P}$-complete [8]. Hence, it is believed that the computation of the number of perfect matching in a general bipartite graph satisfying $\#V_1 = \#V_2$ cannot be polynomial.

In a recent paper Jerrum, Sinclair and Vigoda gave a *fully-polynomial randomized approximation scheme* (*fpras*) to compute the permanent of a nonnegative matrix [7]. (See also Barvinok [1] for computing the permanents within a simply exponential factor, and Friedland, Rider and Zeitouni [5] for concentration of permanent estimators for certain large positive matrices.)

[7] yields the existence a fpras to compute the number of perfect matchings in a general bipartite graph satisfying $\#V_1 = \#V_2$. The aim of this note is to show that there exists fpras to compute the number of k-matchings for any bipartite graph G and any integer $k \in [1, \frac{\#V}{2}]$. In particular, the generating matching polynomial of any bipartite graph G has a fpras. This observation can be used to find a fast computable approximation to the *pressure* function, as discussed in [4], for certain families of infinite graphs appearing in many models of statistical mechanics, like the integer lattice $\mathbb{Z}^d$.

More generally, there exists a fpras for computing $\mathrm{perm}_k B$, the sum of all $k \times k$ subpermanents of an $m \times n$ matrix B, for any nonnegative B. This is done by showing that $\mathrm{perm}_k B = \frac{\mathrm{perm}\, B_k}{(m-k)!(n-k)!}$ for a corresponding $(m+n-k) \times (m+n-k)$ matrix B_k.

It is known that for a nonbipartite graph G on $2n$ vertices, the number of perfect matchings is given by $\mathrm{haf}\, A(G)$, the *hafnian* of the incidence matrix $A(G)$ of G. The existence of a fpras for computing the number of perfect matching for any undirected graph G on even number of vertices is an open problem. (The probabilistic algorithm suggested in [7] applies to the computation of perfect matchings in G, however it is not known if this algorithm is fpras.) The number of k-matchings in a graph G is equal to $\mathrm{haf}_k A(G)$, the sum of the hafnians of all $2k \times 2k$ principle submatrices of $A(G)$. We show that that for any $m \times m$ matrix A there exists a $(2m-2k) \times (2m-2k)$ matrix A_k such that $\mathrm{haf}_k A = \frac{\mathrm{haf}\, A_k}{(2m-k)!}$. Hence the computation of the number of k-matching in an arbitrary G, where $n = O(k)$, has fpras if and only if the number of perfect matching in G has fpras.

2 The equality $\mathrm{perm}_k B = \frac{\mathrm{perm}\, B_k}{(m-k)!(n-k)!}$

Recall that for a square matrix $A = [a_{ij}]_{i,j=1}^n \in \mathbb{R}^{n\times n}$, the permanent of A is given as $\mathrm{perm}\, A := \sum_{\sigma\in \mathrm{S_n}} a_{1\sigma(1)} \dots a_{n\sigma(n)}$, where $\mathrm{S_n}$ is the permutation group on $\langle n\rangle := \{1, \dots, n\}$. Let $\mathrm{Q_{k,m}}$ denote the set of all subset of cardinality k of $\langle m\rangle$. Identify $\alpha \in \mathrm{Q_{k,m}}$ with the subset $\{\alpha_1, \dots, \alpha_k\}$ where $1 \le \alpha_1 < \dots < \alpha_k \le m$. Given an $m \times n$ matrix $B = [b_{ij}]_{i,j=1}^{m,n} \in \mathbb{R}^{m\times n}$ and $\alpha \in \mathrm{Q_{k,m}}, \beta \in \mathrm{Q_{l,n}}$ we let $B[\alpha, \beta] := [b_{\alpha_i\beta_j}]_{i,j=1}^{k,l} \in \mathbb{R}^{k\times l}$ to be the corresponding $k \times l$ submatrix of

B. Then

$$\operatorname{perm}_k B := \sum_{\alpha\in \mathrm{Q_{k,m}}, \beta\in \mathrm{Q_{k,n}}} \operatorname{perm} B[\alpha,\beta].$$

Let $G = (V_1 \cup V_2, E)$ be a bipartite graph on two classes of vertices V_1 and V_2. For simplicity of notation we assume that $E \subset V_1 \times V_2$. It would be convenient to assume that $\#V_1 = m, \#V_2 = n$. So G is presented by $(0,1)$ matrix $B(G) \in \{0,1\}^{m\times n}$. That is $B(G) = [b_{ij}]_{i,j=1}^{m,n}$ and $b_{ij} = 1 \iff (i,j) \in E$. Let $k \in [1, \min(m,n)]$ be an integer. Then k-matching is a choice of k edges in $E_k := \{e_1,\dots,e_k\} \subset E$ such that E_k covers $2k$ vertices in G. That is, no two edges in E_k have a common vertex. It is straightforward to show that $\operatorname{perm}_k B(G)$ is the number of k-matching in G.

More generally, let $B = [b_{ij}] \in \mathbb{R}_+^{m\times n}$, $\mathbb{R}_+ := [0,\infty)$ be an $m \times n$ nonnegative matrix. We associate with B the following bipartite graph $G(B) = (V_1(B)\cup V_2(B), E(B))$. Identify $V_1(B), V_2(B)$ with $\langle m\rangle, \langle n\rangle$ respectively. Then for $i \in \langle m\rangle, j \in \langle n\rangle$ the edge (i,j) is in $E(B)$ if and only if $b_{ij} > 0$. Let $G_w := (V_1(B) \cup V_2(B), E_w(B))$ be the weighted graph corresponding to B. I.e., the weight of the edge $(i,j) \in E(B)$ is $b_{ij} > 0$. Hence $B(G_w)$, the representation matrix of the *weighted bipartite graph* G_w, is equal to B. Let $M \in \mathcal{M}_k(G(B))$. Then $\prod_{(i,j)\in M} b_{ij}$ is the *weight* of the matching M in G_w. In particular, $\operatorname{perm}_k B$ is the *total weight* of *weighted* k-matchings of G_w. The weighted matching polynomial corresponding to $B \in \mathbb{R}_+^{m\times n}$, or G_w induced by B, is defined as:

$$\Phi(x,B) := \sum_{k=0}^{\min(m,n)} \operatorname{perm}_k B\, x^k,\ B \in \mathbb{R}_+^{m\times n},\ \operatorname{perm}_0 B := 0.$$

$\Phi(x,B)$ can be viewed as the *grand partition function* for the monomer-dimer model in statistical mechanics [6]. (See §3 for the case of a nonbipartite graph.) In particular, all roots of $\Phi(x,B)$ are negative.

Theorem 2.1 *Let $B \in \mathbb{R}_+^{m\times n}$ and $k \in \langle \min(m,n)\rangle$. Let $B_k \in \mathbb{R}_+^{(m+n-k)\times(m+n-k)}$ be the following 2×2 block matrix $B_k := \begin{bmatrix} B & \mathbf{1}_{m,m-k} \\ \mathbf{1}_{n-k,n} & \mathbf{0}\end{bmatrix}$, where $\mathbf{1}_{p,q}$ is a $p \times q$ matrix whose all entries are equal to 1. Then*

$$\operatorname{perm}_k B = \frac{\operatorname{perm} B_k}{(m-k)!(n-k)!}. \tag{2.1}$$

Proof. For simplicity of the exposition we assume that $k < \min(m,n)$. (In the case that $k = \min(m,n)$ then B_k has one of the following block structure: $1\times 1,\ 1\times 2,\ 2\times 1$.) Let $G_w = (V_1(B) \cup V_2(B), E_w(B)), G_{w,k} = (V_1(B_k) \cup V_2(B), E_w(B_k))$ be the weighted graphs corresponding to B, B_k respectively. Note that G_w is a weighted subgraph of $G_{w,k}$ induced by $V_1(B) = \langle m\rangle \subset \langle m+n-k\rangle = V_1(B_k), V_2(B) = \langle n\rangle \subset \langle n+m-k\rangle = V_2(B_k)$. Furthermore, each vertex in $V_1(B_k)\backslash V_1(B)$ is connected exactly to each vertex in $V_2(B)$, and each vertex in $V_2(B_k)\backslash V_2(B)$ is connected exactly to each vertex in $V_1(B)$. The weights of each of these edges is 1. These are all edges in $G(B_k)$. A perfect match in $G(B_k)$ correspond to:

- An $n-k$ match between the set of vertices $V_1(B_k)\backslash V_1(B)$ and the set of vertices $\beta' \in \mathrm{Q_{n-k,n}}$, viewed as a subset of $V_2(B)$.
- An $m-k$ match between the set of vertices $V_2(B_k)\backslash V_2(B)$ and the set of vertices $\alpha' \in \mathrm{Q_{m-k,m}}$, viewed as a subset of $V_1(B)$.
- A k match between the set of vertices $\alpha := \langle m\rangle\backslash\alpha' \subset V_1(B)$ and $\beta := \langle n\rangle\backslash\beta' \subset V_2(B)$.

Fix $\alpha \in Q_{k,m}, \beta \in Q_{k,n}$. Then the total weight of k-matchings in $G_w(B_k)$ using the set of vertices $\alpha \subset V_1(B_k), \beta \subset V_2(B_k)$ is given by $\operatorname{perm} B[\alpha,\beta]$. The total weight of $n-k$ matchings using $V_1(B_k)\backslash V_1(B)$ and $\beta' \subset V_2(B_k)$ is $(n-k)!$. The total weight of $m-k$ matchings using $V_2(B_k)\backslash V_2(B)$ and $\alpha' \subset V_1(B_k)$ is $(m-k)!$. Hence the total weight of perfect matchings in $G_w(B_k)$, which matches the set of vertices $\alpha \subset V_1(B_k)$ with the set $\beta \subset V_2(B_k)$ is given by $(m-k)!(n-k)! \operatorname{perm} B[\alpha,\beta]$. Thus $\operatorname{perm} B_k = (m-k)!(n-k)! \operatorname{perm}_k B$. $\square$

We remark that the special case of Theorem 2.1 where $m = n$ appears in an equivalent form in [2].

Proposition 2.2 *The complexity of computing the number of k-matchings in a bipartite graph $G = (V_1 \cup V_2, E)$, where* $\min(\#V_1, \#V_2) \geq k \geq c\max(\#V_1, \#V_2)^\alpha$ *and* $c, \alpha \in (0,1]$, *is polynomially equivalent to the complexity of computing the number of perfect matching in a bipartite graph $G' = (V_1' \cup V_2', E')$, where $\#V_1' = \#V_2'$.*

Proof. Assume first that $G = (V_1 \cup V_2, E), m = \#V_1, n = \#V_2$ and $k \in [c\max(\#V_1, \#V_2)^\alpha, \min(m,n)]$ are given. Let $G' = (V_1' \cup V_2', E')$ be the bipartite graph constructed in the proof of Theorem 2.1. Theorem 2.1 yields that the number of perfect matching in G' determines the number of k-matching in G. Note that $n' := \#V_1' = \#V_2' = O(k^{\frac{1}{\alpha}})$. So the k-matching problem is a special case of the perfect matching problem.

Assume second that $G' = (V_1' \cup V_2', E')$ is a given bipartite graph with $k = \#V_1 = \#V_2$. Let $m, n \geq k$ and denote by $G = (V_1 \cup V_2, E'), \#V_1 = m, \#V_2 = n$ the graph obtained from G by adding $m-k, n-k$ isolated vertices to V_1', V_2' respectively, $(E' = E)$. Then a perfect matching in G' is a k-matching in G, and the number of perfect matching in G' is equal to the number of k-matchings in G. Furthermore if $k \geq c\max(m,n)^\alpha$ it follows that $m, n = O(k^{\frac{1}{\alpha}})$. $\square$

The results of [7] yield.

Corollary 2.3 *Let $B \in \mathbb{R}_+^{m\times n}$ and $k \in \langle\min(m,n)\rangle$. Then there exists a fully-polynomial randomized approximation scheme to compute* $\operatorname{perm}_k B$. *Furthermore for each $x \in \mathbb{R}$ there exists a fully-polynomial randomized approximation scheme to compute the matching polynomial $\Phi(x,B)$.*

3 Hafnians

Let $G = (V, E)$ be an undirected graph on $m := \#V$ vertices. Identify V with $\langle m \rangle$. Let $A(G) = [a_{ij}]_{i,j=1}^m \in \{0,1\}^{m\times m}$ be the incidence matrix of G, i.e. $a_{ij} = 1$ if and only if $(i,j) \in E$. Since we assume that G ia undirected and has no self-loops, $A(G)$ is a symmetric $(0,1)$ matrix with a zero diagonal. Denote by $\mathrm{S_m}(\mathcal{T}) \supset \mathrm{S_{m,0}}(\mathcal{T})$ the set of symmetric matrices and the subset of symmetric matrices with zero diagonal respectively, whose nonzero entries are in the set $\mathcal{T} \subseteq \mathbb{R}$. Thus any $A = [a_{ij}] \in \mathrm{S_{m,0}}(\mathbb{R}_+)$ induces $G(A) = (V(A), E(A))$, where $V(A) = \langle m \rangle$ and $(i,j) \in E(A)$ if and only if $a_{ij} > 0$. Such an A induces a weighted graph $G_w(A)$, where the edge $(i,j) \in E(A)$ has the weight $a_{ij} > 0$. Let $M \in \mathcal{M}_k(G(A))$ be a k-matching in $G(A)$. Then the weight of M in $G_w(A)$ is given by $\prod_{(i,j)\in M} a_{ij}$.

Assume that m is even, i.e. $m = 2n$. It is well known that the number of perfect matchings in G is given by $\mathrm{haf}\, A(G)$, the *hafnian* of $A(G)$. More general, the total weight of all weighted perfect matchings of $G_w(A), A \in \mathrm{S_{2n,0}}(\mathbb{R}_+)$ is given by $\mathrm{haf}\, A$, the hafnian of A.

Recall the definition of the hafnian of $2n \times 2n$ real symmetric matrix $A = [a_{ij}] \in \mathbb{R}^{2n\times 2n}$. Let K_{2n} be the complete graph on $2n$ vertices, and denote by $\mathcal{M}(K_{2n})$ the set of all perfect matches in K_{2n}. Then $\alpha \in \mathcal{M}(K_{2n})$ can be represented as $\alpha = \{(i_1, j_1), (i_2, j_2), .., (i_n, j_n)\}$ with $i_k < j_k$ for $k = 1, \ldots$. Denote $a_\alpha := \prod_{k=1}^n a_{i_k j_k}$. Then $\mathrm{haf}\, A := \sum_{\alpha\in\mathcal{M}(K_{2n})} a_\alpha$. Note that $\mathrm{haf}\, A$ does not depend on the diagonal entries of A. Hafnian of A is related to the *pfaffian* of the skew symmetric matrix $B = [b_{ij}] \in \mathbb{R}^{2n\times 2n}$, where $b_{ij} = a_{ij}$ if $i < j$, the same way the permanent of $C \in \mathbb{R}^{n\times n}$ is related to the determinant of C. Recall $\mathrm{pfaf}\, B = \sum_{\alpha\in\mathcal{M}(K_{2n})} \mathrm{sgn}(\alpha) b_\alpha$, where $\mathrm{sgn}(\alpha)$ is the signature of the permutation $\alpha \in \mathcal{S}_{2n}$ given by $\alpha = \begin{bmatrix} 1 & 2 & 3 & 4 & .. & 2n \\ i_1 & j_1 & i_2 & j_2 & .. & j_n \end{bmatrix}$. Furthermore $\det B = (\mathrm{pfaf}\, B)^2$.

Let $A \in \mathrm{S_m}(\mathbb{R})$. Then

$$\mathrm{haf}_k\, A := \sum_{\alpha\in Q_{2k,m}} \mathrm{haf}\, A[\alpha,\alpha],\ k = 1, \ldots, \lfloor \frac{m}{2} \rfloor.$$

For $A \in \mathrm{S_{m,0}}(\mathbb{R}_+)$ $\mathrm{haf}_k\, A$ is the total weight of all weighted k-matchings in $G_w(A)$. Let $\mathrm{haf}_0(A) := 1$. Then the weighted matching polynomial of $G_w(A)$ is given by $\Phi(x, A) := \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} \mathrm{haf}_k\, A\, x^k$. It is known that a nonconstant $\Phi(x, A), A \in \mathrm{S_{m,0}}(\mathbb{R}_+)$ has only real negative roots [6].

Theorem 3.1 *Let $A \in S_{m,0}(\mathbb{R}_+)$ and $k \in \langle \lfloor \frac{m}{2} \rfloor \rangle$. Let $A_k \in \mathrm{S_{2m-2k,0}}(\mathbb{R}_+)$ be the following 2×2 block matrix $A_k := \begin{bmatrix} A & \mathbf{1}_{m,m-2k} \\ \mathbf{1}_{m-2k,m} & \mathbf{0} \end{bmatrix}$. Then*

$$\mathrm{haf}_k\, A = \frac{\mathrm{haf}\, A_k}{(m-2k)!}. \tag{3.1}$$

Proof. It is enough to consider the nontrivial case $k < \frac{m}{2}$. Let $G_w = (V(A), E_w(A)), G_{w,k} = (V(A_k), E_w(A_k))$ be the weighted graphs corresponding to A, A_k respectively. Note that G_w is a weighted subgraph of $G_{w,k}$ induced by $V(A) = \langle m \rangle \subset \langle 2m-2k \rangle = V(A_k)$. Furthermore, each vertex in $V(A_k)\backslash V(A)$ is connected exactly to each vertex in $V(A)$. The weights of each of these edges is 1. These are all edges in $G(A_k)$. A perfect match in $G(A_k)$ correspond to:

- An $m-2k$ match between the set of vertices $V(A_k)\backslash V(A)$ and the set of vertices $\alpha' \in \mathrm{Q}_{\mathrm{m-2k,m}}$, viewed as a subset of $V(A)$.
- A k match between the set of vertices $\alpha := \langle m \rangle \backslash \alpha' \subset V(B)$.

Fix $\alpha \in Q_{2k,m}$. Then the total weight of k-matchings in $G_w(A_k)$ using the set of vertices $\alpha \subset V(A_k)$ is given by $\mathrm{haf}\, A[\alpha, \alpha]$. The total weight of $m-2k$ matchings using $V(A_k)\backslash V(A)$ and $V(A)\backslash\alpha$ is $(m-2k)!$. Hence the total weight of perfect matchings in $G_w(A_k)$, which matches the set of vertices $\alpha \subset V(A_k)$ is given by $(m-2k)!\, \mathrm{haf}\, A[\alpha, \alpha]$. Thus $\mathrm{haf}\, A_k = (m-2k)!\, \mathrm{haf}_k\, A$. $\square$

It is not known if the computation of the number of perfect matching in an arbitrary undirected graph on an even number of vertices, or more generally the computation of $\mathrm{haf}\, A$ for an arbitrary $A \in \mathrm{S}_{2\mathrm{n},0}(\mathbb{R}_+)$, has a fpras. The probabilistic algorithm outlined in [7] carries over to the computation of $\mathrm{haf}\, A$, however it is an open problem if this algorithm is a fpras. Theorem 3.1 shows that the computation of $\mathrm{haf}_k\, A$, for $A \in \mathrm{S}_{\mathrm{m},0}(\mathbb{R}_+)$, has the same complexity as the computation of $\mathrm{haf}\, A$, for $A \in \mathrm{S}_{2\mathrm{n},0}(\mathbb{R}_+)$.

4 Remarks

In this section we offer an explanation, using the recent results in [3], why $\mathrm{perm}\, A$ is a nicer function than $\mathrm{haf}\, A$. Let $A = [a_{ij}] \in \mathrm{S}_\mathrm{n}(\mathbb{R}), \mathrm{B} = [\mathrm{b_{ij}}] \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$. For $\mathbf{x} := (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$ let

$$p(\mathbf{x}) := \prod_{i=1}^{n} (\sum_{j=1}^{n} b_{ij} x_j), \quad q(\mathbf{x}) := \frac{1}{2} \mathbf{x}^\top A \mathbf{x}.$$

Then $\mathrm{perm}\, B = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} p(\mathbf{x})$ and $\mathrm{haf}\, A = ((\frac{n}{2})!)^{-1} \frac{\partial^n}{\partial x_1 \ldots \partial x_n} q(\mathbf{x})^{\frac{n}{2}}$ if n is even. Assume that $B \in \mathbb{R}_+^{n\times n}$ has no zero row. Then $p(\mathbf{x})$ is a positive hyperbolic polynomial. (See the definition in [3].) Assume that $A \in \mathrm{S}_{2\mathrm{m},0}(\mathbb{R}_+)$ is irreducible. Then $q(\mathbf{x})$, and hence any power $q(\mathbf{x})^i, i \in \mathbb{N}$, is positive hyperbolic if and only if all the eigenvalues of A, except the Perron-Frobenius eigenvalue, are nonpositive.

References

[1] A. Barvinok, Polynomial time algorithms to approximate permanents and mixed discriminants within a simply exponential factor, *Random Structures Algorithms* 14 (1999), 29-61.

[2] S. Friedland, A proof of a generalized van der Waerden conjecture on permanents, *Linear Multilin. Algebra* 11 (1982), 107-120.

[3] S. Friedland and L. Gurvits, Generalized Friedland-Tverberg inequality: applications and extensions, *submitted.*

[4] S. Friedland and U.N. Peled, The pressure associated with multidimensional SOFT, *in preparation.*

[5] S. Friedland, B. Rider and O. Zeitouni, Concentration of permanent estimators for certain large matrices, *Annals of Applied Probability*, 14(2004), 1559-1576.

[6] O.J. Heilman and E.H. Lieb, Theory of monomer-dimer systems, *Comm. Math. Phys.* 25 (1972), 190–232; Errata 27 (1972), 166.

[7] M. Jerrum, A. Sinclair and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries, *J. ACM* 51 (2004), 671-697.

[8] L.G. Valiant, The complexity of computing the permanent, *Theoretical Computer Science* 8 (1979), 189-201.