An analysis of a Monte Carlo algorithm for estimating the permanent

Alan Frieze*
Department of Mathematics, Carnegie Mellon University
Pittsburgh PA15213, U.S.A.

Mark Jerrum†
Department of Computer Science, University of Edinburgh
The King's Buildings, Edinburgh EH9 3JZ, United Kingdom

ABSTRACT Karmarkar, Karp, Lipton, Lovász, and Luby proposed a Monte Carlo algorithm for approximating the permanent of a non-negative $n \times n$ matrix, which is based on an easily computed, unbiased estimator. It is not difficult to construct 0,1-matrices for which the variance of this estimator is very large, so that an exponential number of trials is necessary to obtain a reliable approximation that is within a constant factor of the correct value.

Nevertheless, the same authors conjectured that for a random 0,1-matrix the variance of the estimator is typically small. The conjecture is shown to be true; indeed, for almost every 0,1-matrix A, just $\mathrm{O}(n\,\omega(n)\varepsilon^{-2})$ trials suffice to obtain a reliable approximation to the permanent of A within a factor $1 \pm \varepsilon$ of the correct value. Here $\omega(n)$ is any function tending to infinity as $n \to \infty$. This result extends to random 0,1-matrices with density at least $n^{-1/2}\omega(n)$.

It is also shown that polynomially many trials suffice to approximate the permanent of any dense 0,1-matrix, i.e., one in which every row- and column-sum is at least $(\frac{1}{2} + \alpha)n$, for some constant $\alpha > 0$. The degree of the polynomial bounding the number of trials is a function of α, and increases as $\alpha \to 0$.

---

* Supported by NSF grant CCR-9225008.

† The work described here was partly carried out while the author was visiting Princeton University as a guest of DIMACS (Center for Discrete Mathematics and Computer Science).

1. Summary

The *permanent* of an $n \times n$ matrix $A = (a_{ij} : 0 \leq i, j \leq n-1)$ is defined by

$$\operatorname{per} A = \sum_{\pi} \prod_{i=0}^{n-1} a_{i,\pi(i)} ,$$

where the sum is over all permutations π of $[n] = \{0, \ldots, n-1\}$. In this paper, A will usually be a 0,1-matrix, in which case the permanent of A has a simple combinatorial interpretation: namely, $\operatorname{per} A$ is equal to the number of perfect matchings (1-factors) in the bipartite graph $G = (U, V, E)$, where $U = V = [n]$, and $(i, j) \in E$ iff $a_{ij} = 1$. The permanent function arises naturally in a number of fields, including algebra, combinatorial enumeration, and the physical sciences, and has been an object of study by mathematicians since first appearing in 1812 in the work of Cauchy and Binet. (See Minc [16] for background material.) Despite considerable effort, and in contrast with the syntactically very similar determinant, no efficient procedure for computing this function is known.

Convincing evidence for the inherent intractability of the permanent was provided in the late 1970s by Valiant [18], who demonstrated that it is complete for the class #P of enumeration problems, and thus as hard as counting the number of satisfying assignments to a CNF formula, or the number of accepting computations of a polynomial-time-bounded nondeterministic Turing machine. Interest has therefore turned to finding computationally feasible approximation algorithms for the permanent.

The notion of "computationally feasible approximation algorithm" can be formalised as follows. Let f be a function from input strings to natural numbers. A *randomised approximation scheme* [13] for f is a probabilistic algorithm that takes as input a string x and a real number $0 < \varepsilon < 1$, and produces as output a number Y (a random variable) such that $(1-\varepsilon) f(x) \leq Y \leq (1+\varepsilon) f(x)$ with high probability. For definiteness we take the phrase "with high probability" to mean with probability at least $\frac{3}{4}$. The success probability may be boosted to $1 - \delta$ for any desired $\delta > 0$ by running the algorithm $O(\lg \delta^{-1})$ times and taking the median of the results [8, Lemma 6.1]. A randomised approximation scheme is said to be *fully polynomial* if its execution time is bounded by a polynomial in $|x|$ and ε^{-1}. We shall contract the rather unwieldy phrase "fully-polynomial randomised approximation scheme" to *fpras*.

The question of whether there exists an fpras for the permanent of a 0,1-matrix has received much attention, but for the time being remains open. Given the apparent lack of progress, it seems reasonable to weaken the requirements further, and ask whether

there exists an fpras for $\operatorname{per} A$ that works for "almost all" inputs.* In order to make this statement precise, it is convenient to switch to a graph-theoretic viewpoint. Our new question, then, is whether there exist a randomised algorithm $\mathcal{A}$ and a family $\mathcal{G}$ of bipartite graphs, satisfying the following two conditions:

(1) When restricted to inputs of the form (G, ε) where $G \in \mathcal{G}$, the algorithm $\mathcal{A}$ constitutes an fpras for the number of perfect matchings in G.

(2) Almost every (a.e.) bipartite graph is a member of $\mathcal{G}$. That is, the fraction of $2n$-vertex bipartite graphs that are *not* members of $\mathcal{G}$ tends to zero as n tends to infinity.

The modified question was answered affirmatively by Jerrum and Sinclair [10], who presented a randomised approximation scheme based on the simulation of an appropriately defined Markov chain, an approach that had earlier been proposed by Broder [2, 15]. The polynomial bounding the execution time of the algorithm of Jerrum and Sinclair was not explicitly computed in [10], but its degree is not small. It is not yet clear whether this approach could ever form the basis of a truly practical algorithm, despite the undoubted scope that exists for optimising the algorithm and tightening its analysis. For this reason alone, it is worth investigating alternative approaches.

A promising Monte Carlo algorithm for approximating the permanent of a 0,1-matrix was proposed by Karmarkar, Karp, Lipton, Lovász, and Luby [12]. Their algorithm is based on an unbiased estimator for $\operatorname{per} A$, which will be described in the next section. The KKLLL estimator may be computed relatively efficiently, the most computationally demanding step being the evaluation of a single $n \times n$ determinant. A randomised approximation scheme can be obtained from the KKLLL estimator as follows. Choose t sufficiently large, and make a sequence of t trials with the KKLLL estimator, letting the results be $Z_0, Z_1, \ldots, Z_{t-1}$; then return $(Z_0 + Z_1 + \cdots + Z_{t-1})t^{-1}$ as the estimate of $\operatorname{per} A$.

The efficiency of the above approximation scheme depends on the chosen value of t and hence on the variance of the KKLLL estimator. Reverting once more to the graph-theoretic viewpoint, suppose that the KKLLL estimator is being used to provide an approximation to the number of perfect matchings in a specified bipartite graph G. The number of trials necessary to obtain a reliable and close approximation is greatly influenced by the structure of G. To illustrate this point, consider first the graph G that is the disjoint union of $\frac{1}{2}n$ copies of $K_{2,2}$. In this case, exponentially many trials are necessary to obtain an approximation that satisfies the conditions of a randomised approximation scheme. In stark contrast, $\mathrm{O}(n\varepsilon^{-2})$ trials are sufficient to accomplish the same task when G is the

---

* There is persuasive circumstantial evidence that no efficient algorithm exists that computes the permanent *exactly* on almost all inputs; see, e.g., Gemmell and Sudan [6].

complete bipartite graph K_{nn} [12]. Karmarkar et al. conjecture that it is the second of these two examples that is the more characteristic of graphs in general, and that $O(n\varepsilon^{-2})$ trials suffice for a.e. G. The first of our two main results shows that something very close to the conjecture is true: namely that $n\,\omega(n)\varepsilon^{-2}$ trials suffice for a.e. G, where $\omega(n)$ is any function tending to infinity as $n \to \infty$.

We also consider the performance of the KKLLL estimator when applied to dense 0,1-matrices, i.e., those in which every row- and column-sum is at least $(\frac{1}{2}+\alpha)n$ (equivalently, to bipartite graphs with minimum vertex degree at least $(\frac{1}{2}+\alpha)n$). It was already known [10] that the execution time of the Jerrum-Sinclair approximation scheme is uniformly bounded by a polynomial in n, over the whole range of α. However, for reasons given earlier, it is still of interest to know how other, simpler, and perhaps more practical approaches perform. Our second main result concerning the KKLLL estimator is that a polynomial number of trials suffice to estimate the permanent of a dense matrix for any fixed $\alpha > 0$, though the degree of the polynomial in question increases as $\alpha \to 0$.

A more precise statement of the results will be possible after we have reviewed the properties of the KKLLL estimator.

2. The KKLLL estimator

The estimator is defined to be the random variable Z that results from the simple experiment described below.

(1) Form a matrix $B = (b_{ij})$ from A as follows. Let $\{1, \omega, \omega^2\}$ be the cube roots of unity. For each pair i, j in the range $0 \le i, j \le n-1$:

  (a) if $a_{ij} = 0$ then set b_{ij} equal to 0;

  (b) if $a_{ij} = 1$ then choose b_{ij} independently and u.a.r. from the set $\{1, \omega, \omega^2\}$.

(2) Set Z equal to $|\det B|^2$, where $|z|$ denotes the modulus of complex number z.

The KKLLL estimator is a simple modification of an earlier estimator of Godsil and Gutman [7], which used square rather than cube roots of unity. At first sight, it may seem surprising that the KKLLL estimator should be unbiased. Nevertheless, the following theorem can be established with little difficulty [12].

Theorem 1. $\operatorname{Exp} Z = \operatorname{per} A$.

As we have noted, the efficiency of the KKLLL estimator will depend on its variance. Karmarkar et al. derive a useful expression for the variance, which is best formulated in graph-theoretic terms. Let G be a bipartite graph on vertex set $U + V$, where $U = V = [n]$, and let M and M' be perfect matchings in G. Denote by $c(M, M')$ the number of connected components (cycles) in $M \oplus M'$, the symmetric difference of M and M'.

Define $\gamma(G) = \operatorname{Exp}\left(2^{c(M,M')}\right)$ to be the expected value of $2^{c(M,M')}$ when M and M' are selected u.a.r. from the set of all perfect matchings in G. (If G has no perfect matchings then define $\gamma(G) = 1$.)

Theorem 2. *(Karmarkar, Karp, Lipton, Lovász, and Luby.)*

$$\frac{\operatorname{Exp}(Z^2)}{(\operatorname{Exp} Z)^2} = \gamma(G).$$

Proof. The theorem is essentially a restatement of Theorem 4 of [12]. However, it may be helpful to point out the precise correspondence between the two versions of the theorem.

The set D that appears in the original version of the theorem can be interpreted as the set of all subgraphs of G that can be expressed as a union of two perfect matchings in G. Note that any subgraph in D is a disjoint union of single edges and cycles; further note that the number of ways of expressing the subgraph as a union of two perfect matchings is 2^c, where c is the number of cycles in the subgraph. With this correspondence in mind, it can be seen that the denominator appearing on the right hand side of the identity in the original statement of the theorem is simply the square of the number of matchings in G. (Note that the G appearing in the original theorem is *not* the same as the one used here.) Using the same correspondence, the numerator can be seen to be equal to $\sum_{M,M'} 2^{c(M,M')}$, where the summation is over all pairs (M, M') of matchings in G. Thus the quotient is the expected value of $2^{c(M,M')}$ when M and M' are perfect matchings selected u.a.r. from G. By definition, this expectation is $\gamma(G)$.

Corollary 3. *A sequence of $\mathrm{O}\left(\varepsilon^{-2}\gamma(G)\right)$ trials with the KKLLL estimator suffices to obtain an approximation to the number of perfect matchings in G that satisfies the conditions of a randomised approximation scheme.*

Proof. Perform $t = \left\lceil 4\varepsilon^{-2}\gamma(G)\right\rceil$ trials with the KKLLL estimator, letting the results be $Z_0, Z_1, \ldots, Z_{t-1}$. Using Theorem 2,

$$\operatorname{Var}\left(\frac{1}{t}\sum_{i=0}^{t-1} Z_i\right) = \frac{\operatorname{Var} Z}{t} \leq \frac{\gamma(G)(\operatorname{Exp} Z)^2}{t}.$$

Hence, by Chebychev's inequality,

$$\Pr\left((1-\varepsilon)\operatorname{Exp} Z \leq \frac{1}{t}\sum_{i=0}^{t-1} Z_i \leq (1+\varepsilon)\operatorname{Exp} Z\right) \geq \frac{3}{4}.$$

The important point about Corollary 3 is that it reduces the analysis of the KKLLL approximation scheme on random inputs, or some restricted class of inputs, to the analysis of $\gamma(G)$ for randomly chosen G, or for a graph G selected from the given class. A detailed analysis of $\gamma(G)$ for random and dense graphs G forms the content of Sections 4 and 5.

3. The permanent of a random matrix

For reasons that will be explained later, we choose to work with the random graph model $\mathcal{B}(n,m)$; thus our sample space is the set of all m-edge bipartite graphs on vertex set $U+V$, where $U=V=[n]$, and the probability distribution is uniform. The formula "select $G \in \mathcal{B}(n,m)$" is thus a shorthand for "select u.a.r. an m-edge bipartite graph on vertex set $U+V$." We have noted that the performance of the KKLLL approximation scheme on input G depends crucially on the quantity $\gamma(G) = \mathrm{Exp}\left(2^{c(M,M')}\right)$, where M and M' are matchings in G selected u.a.r., and $c(M,M')$ denotes the number of cycles in $M \oplus M'$. An analysis of the behaviour of the approximation scheme on a random input will therefore rest on an estimation of $\gamma(G)$ when G is selected according to the random graph model $\mathcal{B}(n,m)$. The natural route is via an experiment (A) of the form:

(A1) select $G \in \mathcal{B}(n,m)$;

(A2) select M, M' u.a.r. from the set of all matchings in G.

Unfortunately, it seems impossible to argue about the behaviour of $c(M,M')$ when M and M' are generated in this way. Instead we consider a related experiment (B) of the form:

(B1) select k in the range $0 \le k \le n$ from an "appropriate" distribution;

(B2) select M, M' u.a.r. from the set of pairs of matchings on vertex set $U+V$ that satisfy $|M \cap M'| = k$;

(B3) select G u.a.r. from the set of all m-edge bipartite graphs on vertex set $U+V$ that contain M and M'.

Intuitively (and in fact) these two experiments are not too dissimilar provided the number of perfect matchings in a random $G \in \mathcal{B}(n,m)$ is fairly tightly clustered. Theorem 4 assures us that this is indeed the case.

Theorem 4. *Suppose the function $m = m(n)$ satisfies $m^2 n^{-3} \to \infty$ as $n \to \infty$. For $G \in \mathcal{B}(n,m)$, denote by $X(G)$ the number of perfect matchings in G. Then*

$$\frac{\mathrm{Exp}(X^2)}{(\mathrm{Exp}\, X)^2} = 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big).$$

The key decision here is to work with $\mathcal{B}(n,m)$ rather than the more usual random graph model $\mathcal{B}(n,p)$, in which potential edges are selected independently and with probability p. Theorem 4 fails badly in the latter model; indeed, for $p = n^{-\varepsilon}$, the ratio $\mathrm{Exp}(X^2)/(\mathrm{Exp}\, X)^2$ grows faster than any polynomial in n. Informally, one could say that the permanent of a random 0,1-matrix (determined by a sequence of n^2 Bernoulli trials) depends strongly on the number of 1s in the matrix, but only rather weakly on their disposition. Note that more precise information about the distribution of X has recently been obtained by Janson [9].

Proof of Theorem 4. Let M be a perfect matching on $U + V$, i.e., a set of n independent edges spanning U and V. For $G \in \mathcal{B}(n,m)$, define the random variable $X_M(G)$ to be 1 if M is contained in G, and 0 otherwise. Note that by linearity of expectation

$$\mathrm{Exp}\, X = \sum_M \mathrm{Exp}\, X_M, \tag{1}$$

and

$$\mathrm{Exp}(X^2) = \sum_{M,M'} \mathrm{Exp}(X_M X_{M'}), \tag{2}$$

where M and M' range over all $n!$ matchings on $U + V$.

To estimate the above sums, we need to compute the probability that a particular graph appears as a subgraph of a randomly selected $G \in \mathcal{B}(n,m)$. Let H be any t-edge bipartite graph on vertex set $U + V$, where $t \le 2n$. The probability $q = q(t)$ that H is a subgraph of $G \in \mathcal{B}(n,m)$ is given by

$$q = \binom{n^2 - t}{m - t}\binom{n^2}{m}^{-1} = \frac{m(m-1)\cdots(m-t+1)}{n^2(n^2-1)\cdots(n^2-t+1)}.$$

Taking logarithms, and expanding $\ln(1-x)$ as $-x + \mathrm{O}(x^2)$, we have:

$$\begin{aligned}
\ln q &= \sum_{i=0}^{t-1} \big[\ln(m-i) - \ln(n^2-i)\big] \\
&= t\ln\Big(\frac{m}{n^2}\Big) + \sum_{i=0}^{t-1}\Big[\ln\Big(1-\frac{i}{m}\Big) - \ln\Big(1-\frac{i}{n^2}\Big)\Big] \\
&= t\ln\Big(\frac{m}{n^2}\Big) - \sum_{i=0}^{t-1}\Big[\frac{i}{m} - \frac{i}{n^2} + \mathrm{O}\Big(\frac{i^2}{m^2}\Big)\Big] \\
&= t\ln\Big(\frac{m}{n^2}\Big) - \frac{t(t-1)}{2}\Big(\frac{1}{m} - \frac{1}{n^2}\Big) + \mathrm{O}\Big(\frac{t^3}{m^2}\Big).
\end{aligned}$$

Thus, noting that $tm^{-1} \leq 2nm^{-1} = \mathrm{O}(n^3 m^{-2})$,

$$q = \Big(\frac{m}{n^2}\Big)^t \exp\Big\{ -\frac{t^2}{2}\Big(\frac{1}{m} - \frac{1}{n^2}\Big) + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\}. \tag{3}$$

Specialising to the case $t = n$, we obtain

$$\mathrm{Exp}\, X_M = \Big(\frac{m}{n^2}\Big)^n \exp\Big\{ -\frac{n^2}{2m} + \frac{1}{2} + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\},$$

and hence, from equation (1),

$$(\mathrm{Exp}\, X)^2 = (n!)^2 \Big(\frac{m}{n^2}\Big)^{2n} \exp\Big\{ -\frac{n^2}{m} + 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\}. \tag{4}$$

In order to deal with sum (2), we need to estimate the number of pairs of matchings M, M' as a function of the overlap $k = |M \cap M'|$. This is essentially the *problème des rencontres*, which asks for the number of permutations of $[n]$ that leave precisely k elements fixed. Let $D(n)$ denote the solution to the *problème des rencontres* in the special case $k = 0$; thus $D(n)$ is the number of "derangements" of n elements. An elementary application of the principle of inclusion-exclusion establishes that $D(n)$ is equal to $\mathrm{e}^{-1} n!$, rounded to the nearest integer [8, p. 9]. The number of pairs of matchings M, M' with $|M \cap M'| = k$ has a simple expression in terms of $D(\cdot)$, namely

$$n!\binom{n}{k} D(n-k). \tag{5}$$

(To make sense of this formula, we should take $D(0) = 1$.)

We are now ready to tackle sum (2). Letting $\alpha = 2n(m^{-1} - n^{-2})$, and using estimates (3) and (5) we have:

$$\begin{aligned}
\mathrm{Exp}(X^2) &= \sum_{k=0}^{n} \sum_{\substack{M, M': \\ |M \cap M'| = k}} \mathrm{Exp}(X_M X_{M'}) \\
&= \sum_{k=0}^{n} n!\binom{n}{k} D(n-k) \Big(\frac{m}{n^2}\Big)^{2n-k} \exp\Big\{ -\frac{(2n-k)^2}{2}\Big(\frac{1}{m} - \frac{1}{n^2}\Big) + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\} \\
&\leq n!\Big(\frac{m}{n^2}\Big)^{2n} \sum_{k=0}^{n} \binom{n}{k} D(n-k) \Big(\frac{n^2}{m}\Big)^k \exp\Big\{ -\alpha n + \alpha k + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\} \\
&= n!\Big(\frac{m}{n^2}\Big)^{2n} \exp\Big\{ -\alpha n + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\} \sum_{k=0}^{n} \binom{n}{k} D(n-k) \Big[\frac{\mathrm{e}^{\alpha} n^2}{m}\Big]^k.
\end{aligned} \tag{6}$$

Noting that $D(n-k) \le \mathrm{e}^{-1}(n-k)! + 1$ and $\mathrm{e}^{\alpha} = 1 + \mathrm{O}(nm^{-1})$, we obtain the following bound on the sum appearing in (6):

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n}{k} D(n-k) \Big[\frac{\mathrm{e}^{\alpha} n^2}{m}\Big]^k &\le \frac{n!}{\mathrm{e}} \sum_{k=0}^{\infty} \frac{1}{k!} \Big[\frac{\mathrm{e}^{\alpha} n^2}{m}\Big]^k + \sum_{k=0}^{n} \binom{n}{k} \Big[\frac{\mathrm{e}^{\alpha} n^2}{m}\Big]^k \\
&= n! \exp\Big\{\frac{\mathrm{e}^{\alpha} n^2}{m} - 1\Big\} + \Big[1 + \frac{\mathrm{e}^{\alpha} n^2}{m}\Big]^n \\
&\le n! \exp\Big\{\frac{n^2}{m} - 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\} + \big[1 + \mathrm{O}(\sqrt{n}\,)\big]^n \\
&= n! \exp\Big\{\frac{n^2}{m} - 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\}. \qquad (7)
\end{aligned}$$

(The second term in the penultimate line is much smaller than the first, and can be absorbed within the $\mathrm{O}(\cdot)$ of the first term.) Substituting (7) for the sum in (6) we obtain

$$\mathrm{Exp}(X^2) = (n!)^2 \Big(\frac{m}{n^2}\Big)^{2n} \exp\Big\{ - \frac{n^2}{m} + 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\}.$$

The theorem follows from this estimate combined with the earlier one (4).

It is perhaps worth remarking that there is a rudimentary approximation algorithm for the permanent implicit in Theorem 4. Suppose A is a 0,1-matrix chosen uniformly at random. Let m be the number of ones appearing in A, and compute the expectation of $\mathrm{per}\, A$ conditional on A having precisely m ones. Theorem 4 assures us that the probability that this expectation differs from $\mathrm{per}\, A$ by more than say 1% tends to zero as n tends to infinity.

4. The performance of the KKLLL estimator: random matrices

We are now ready to tackle the first main result, to the effect that $\gamma(G)$ is small for almost every bipartite graph G.

Theorem 5. *Let $m = m(n)$ and $\delta = \delta(n)$ be functions satisfying $0 < \delta < 1$, and $m^2 \delta n^{-3} \to \infty$ as $n \to \infty$. Assume n is sufficiently large, and select $G \in \mathcal{B}(n,m)$. Then* $\Pr\big(\gamma(G) \le n\delta^{-1}\big) \ge 1 - \delta$.

Proof. We begin with some preliminary computations concerned with the number of cycles in a random derangement. Denote by S_n the set of all permutations on $[n]$, and by $D_n \subset S_n$ the set of all derangements, i.e., permutations with no fixed points. For $\pi \in S_n$, let $c(\pi)$ be the number of cycles in π, including those of length one. Consider the sums

$s(n) = \sum_{\pi \in S_n} 2^{c(\pi)}$ and $d(n) = \sum_{\pi \in D_n} 2^{c(\pi)}$; the latter may be expressed in terms of the former by applying the principle of inclusion-exclusion:

$$\begin{aligned} d(n) &= s(n) - \binom{n}{1} 2^1 s(n-1) + \binom{n}{2} 2^2 s(n-2) - \cdots + (-1)^n \binom{n}{n} 2^n s(0) \\ &= \sum_{k=0}^{n} \binom{n}{k} (-2)^k s(n-k). \end{aligned} \tag{8}$$

(The first term corresponds to unrestricted permutations; the second to permutations that fix specified single elements; the third to permutations that fix specified pairs of elements; and so on.) Now it is known (see [11, Ex. 3.12]) that $s(n) = (n+1)!$ Substituting for $s(n)$ in equation (8) and simplifying, we obtain

$$\begin{aligned} d(n) &= n! \sum_{k=0}^{n} \frac{(-2)^k (n-k+1)}{k!} \\ &= (n+1)! \sum_{k=0}^{n} \frac{(-2)^k}{k!} \;-\; n! \sum_{k=1}^{n} \frac{(-2)^k}{(k-1)!} \\ &= \mathrm{e}^{-2}(n+1)! + \mathrm{O}(2^n) + 2\mathrm{e}^{-2} n! + \mathrm{O}(2^n) \\ &= \mathrm{e}^{-2}(n+3)n! + \mathrm{O}(2^n). \end{aligned}$$

Since the total number of derangements of n elements is $\mathrm{e}^{-1}n! + \mathrm{O}(1)$, the expectation of $2^{c(\pi)}$ over all derangements π is

$$\mathrm{e}^{-1}(n+3) + \mathrm{O}\Big(\frac{2^n}{n!}\Big) = \mathrm{e}^{-1}n + \mathrm{O}(1). \tag{9}$$

Let Ω denote the set of triples (G, M, M'), where G is an m-edge bipartite graph on vertex set $U + V$, and M, M' are matchings in G. Recall experiment (B) from the previous section. Observe that the number of ways of extending M, M' to a graph G in step (B3) is a function only of the overlap $k = |M \cap M'|$. Thus it is clear that the probability distribution on k in step (B1) can be chosen so that the result of the experiment is a triple (G, M, M') chosen u.a.r. from Ω. Also observe that, for given k, the expected value of $2^{c(M,M')}$ after step (B2) is the same as the expected value of $2^{c(\pi)}$, where π is selected u.a.r. from the set of all derangements on $n-k$ elements. Thus the expected value of $2^{c(M,M')}$ for a triple (G, M, M') selected u.a.r. from Ω is bounded above by $\mathrm{e}^{-1}n + \mathrm{O}(1)$, that is:

$$\frac{1}{|\Omega|} \sum_{(G,M,M') \in \Omega} 2^{c(M,M')} \leq \mathrm{e}^{-1}n + \mathrm{O}(1). \tag{10}$$

Choose $G \in \mathcal{B}(n,m)$, and recall that $X(G)$ denotes the number of perfect matchings in G. Theorem 4 and Chebychev's inequality together imply $\Pr\big(X \le \frac{3}{4}\operatorname{Exp}(X)\big) = \mathrm{O}(n^3m^{-2})$, which is clearly equivalent to $\Pr(X^2 \le \frac{9}{16}\operatorname{Exp}(X)^2) = \mathrm{O}(n^3m^{-2})$. A second application of Theorem 4 then yields

$$\Pr\big(X^2 \le \tfrac{1}{2}\operatorname{Exp}(X^2)\big) = \mathrm{O}\Big(\frac{n^3}{m^2}\Big). \tag{11}$$

Let N be the number of m-edge bipartite graphs on vertex set $U+V$. To complete the proof of the theorem, we shall assume that there are more than δN graphs G with $\gamma(G) > n\delta^{-1}$, and obtain a contradiction. Note that the assumption, taken together with (11), would imply that at least $\big[\delta - \mathrm{O}(n^3m^{-2})\big]N$ graphs simultaneously satisfy the conditions $\gamma(G) > n\delta^{-1}$ and $X(G)^2 > \frac{1}{2}\operatorname{Exp}(X^2)$. Now observe that inequality (10) may be recast in the form

$$\frac{1}{|\Omega|}\sum_G X(G)^2\gamma(G) \le \mathrm{e}^{-1}n + \mathrm{O}(1).$$

According to our calculations, the left hand side of this inequality is bounded below by

$$\frac{1}{|\Omega|}\left[\frac{1}{2} - \mathrm{O}\Big(\frac{n^3}{m^2\delta}\Big)\right]nN\operatorname{Exp}(X^2) = \left[\frac{1}{2} - \mathrm{O}\Big(\frac{n^3}{m^2\delta}\Big)\right]n.$$

But since $n^3m^{-2}\delta^{-1} \to 0$ as $n \to \infty$, this provides a contradiction when n is sufficiently large.

It should be clear that the event $\gamma(G) \le n\delta^{-1}$ appearing in the statement of Theorem 5 may be replaced by $\gamma(G) \le an\delta^{-1}$, where a is any constant exceeding e^{-1}. The result easily translates to the random graph model $\mathcal{B}(n,p)$.

Corollary 6. *Let $p = p(n)$ and $\delta = \delta(n)$ be functions satisfying $0 < p,\delta < 1$, and $p^2\delta n \to \infty$ as $n \to \infty$. Assume n is sufficiently large, and select $G \in \mathcal{B}(n,p)$. Then $\Pr\big(\gamma(G) \le n\delta^{-1}\big) \ge 1-\delta$.*

Proof. The result follows from Theorem 5, using standard techniques for translating between the two random graph models. See Theorem 2 on page 34 of [1].

Specialising to the case $p = \frac{1}{2}$, we obtain:

Corollary 7. *Let $\omega(n)$ be any function tending to infinity as $n \to \infty$. Then a.e. $G \in \mathcal{B}(n,p{=}\frac{1}{2})$ satisfies $\gamma(G) \le n\,\omega(n)$.*

Thus, as claimed at the outset, $O(n\,\omega(n)\varepsilon^{-2})$ trials using the KKLLL estimator suffice to obtain a reliable estimate — to within a factor $1+\varepsilon$ of the correct value — of the permanent of a.e. 0,1-matrix.

5. The performance of the KKLLL estimator: dense matrices

We now analyse the variance of the KKLLL estimator when applied to inputs satisfying a simple *deterministic* criterion.

Theorem 8. *Suppose $\alpha > 0$ is a constant, and let G be an $(n+n)$-vertex bipartite graph of minimum vertex degree $\delta(G) \geq (\frac{1}{2}+\alpha)n$; then $\gamma(G) \leq O(n^{1+(2\ln 2)/\alpha})$.*

Assume G is as in the statement of the theorem. Let U, V be the vertex bipartition of G, and $\mathcal{M}$ be the set of perfect matchings of G. Fix a perfect matching $M_0 \in \mathcal{M}$, and for $M \in \mathcal{M}$ let

$$\iota(M) = |M \cap M_0|, \text{ and}$$
$$c(M) = \text{number of cycles in } M \oplus M_0.$$

Let $\mathcal{M}_{k,\ell} = \{M \in \mathcal{M} : \iota(M) = k,\, c(M) = \ell\}$, and $N_{k,\ell} = |\mathcal{M}_{k,\ell}|$. We show that perfect matchings of G are concentrated in sets $\mathcal{M}_{k,\ell}$ with k and ℓ small.

Lemma 9. *Let $N_{k,\ell}$ be as defined above. Then*

(a) $k\alpha N_{k,\ell} \leq N_{k-2,\ell+1} + 2N_{k-1,\ell}$, *and*

(b) $(2\alpha\ell - 1 - k\ell/n)\, N_{k,\ell} \leq 2(\ln n) N_{k,\ell-1}$.

Proof. We use a quantitative version of Dirac's [3] argument for demonstrating the existence of a Hamilton cycle in a dense graph; the same basic technique was used by Dyer, Frieze, and Jerrum [4] to verify an fpras for counting Hamilton cycles in a dense graph.

We first show part (a) of the lemma. Fix k, ℓ and consider pairs (M, M') with $M \in \mathcal{M}_{k,\ell}$ and $M' \in \mathcal{M}_{k-2,\ell+1} \cup \mathcal{M}_{k-1,\ell}$ such that for some $a_1, a_2 \in U$ and $b_1, b_2 \in V$,

$$M \setminus M' = \{(a_1, b_1), (a_2, b_2)\},$$
$$M' \setminus M = \{(a_1, b_2), (a_2, b_1)\},$$

and

$$(a_1, b_1) \in M \cap M_0.$$

There are two types of pair satisfying these conditions:

(i) If $(a_2, b_2) \in M_0$, then $M' \in \mathcal{M}_{k-2,\ell+1}$; moreover, $M' \cap M_0$ is obtained from $M \cap M_0$ by deleting the two edges (a_1, b_1) and (a_2, b_2), and $M' \oplus M_0$ is obtained from $M \oplus M_0$ by adding the 4-cycle $(a_1, b_1, a_2, b_2, a_1)$.

(ii) If $(a_2, b_2) \notin M_0$, then $M' \in \mathcal{M}_{k-1,\ell}$; moreover, $M' \cap M_0$ is obtained from $M \cap M_0$ by deleting the single edge (a_1, b_1), and $M' \oplus M_0$ is obtained from $M \oplus M_0$ by replacing the edge (a_2, b_2) of some cycle by the path (a_2, b_1, a_1, b_2) of length three.

Let $E_{k,\ell}$ denote the set of all such pairs (M, M'). For $M \in \mathcal{M}_{k,\ell}$, let $\zeta(M)$ denote the number of perfect matchings $M' \in \mathcal{M}_{k-2,\ell+1} \cup \mathcal{M}_{k-1,\ell}$ such that $(M, M') \in E_{k,\ell}$. For $M' \in \mathcal{M}_{k-2,\ell+1} \cup \mathcal{M}_{k-1,\ell}$, let $\eta(M')$ denote the number of perfect matchings $M \in \mathcal{M}_{k,\ell}$ such that $(M, M') \in E_{k,\ell}$.

Fix $M \in \mathcal{M}_{k,\ell}$ and $(a, b) \in M \cap M_0$. There are $s \geq 2\alpha n - 1$ edges (a', b') of M, other than (a, b) itself, such that both (a, b') and (a', b) are edges of G. Suppose s_1 are such that $(a', b') \in M \cap M_0$, and let $s_2 = s - s_1$. Then (a, b) contributes to s_1 type (i) pairs and s_2 type (ii) pairs involving M. Hence,

$$\begin{aligned} \zeta(M) &\geq \sum_{(a,b)} (\tfrac{1}{2} s_1 + s_2) \\ &\geq \tfrac{1}{2} k\alpha n, \end{aligned} \tag{12}$$

provided $n \geq \alpha^{-1}$. The $\frac{1}{2}$ in inequality (12) comes from the fact that two edges of $M \cap M_0$ contribute to the same type (i) pair.

On the other hand, if $M' \in \mathcal{M}_{k-2,\ell+1}$ then $\eta(M')$ is at most the number of 4-cycles in $M' \oplus M_0$, and so $\eta(M) \leq \frac{1}{2} n$. If $M' \in \mathcal{M}_{k-1,\ell}$ then $\eta(M')$ is at most the number of paths of length three in $M' \oplus M_0$ with middle edge in M_0, and so $\eta(M') \leq n$. Hence,

$$\begin{aligned} \tfrac{1}{2} k\alpha n N_{k,\ell} &\leq |E_{k,\ell}| \\ &\leq \tfrac{1}{2} N_{k-2,\ell+1} + n N_{k-1,\ell}, \end{aligned}$$

and (a) follows.

We now turn to part (b) of the lemma. Let $E'_{k,\ell}$ denote the set of pairs $(M, M') \in \mathcal{M}_{k,\ell} \times \mathcal{M}_{k,\ell-1}$ such that, for some $a_1, a_2 \in U$ and $b_1, b_2 \in V$,

$$\begin{aligned} M \setminus M' &= \{(a_1, b_1), (a_2, b_2)\}, \\ M' \setminus M &= \{(a_1, b_2), (a_2, b_1)\}, \end{aligned}$$

and

$$(a_1, b_1), (a_2, b_2), (a_2, b_1), (a_1, b_2) \notin M_0.$$

Here $M' \cap M_0 = M \cap M_0$ and $M' \oplus M_0$ is obtained from $M \oplus M_0$ as follows: take two disjoint cycles, C_1 containing (a_1, b_1) and C_2 containing (a_2, b_2). Replace the edges $(a_1, b_1), (a_2, b_2)$ by $(a_1, b_2), (a_2, b_1)$ creating one large cycle out of the vertices of C_1 and C_2. If C_i has $2m_i$ vertices, for $i = 1, 2$, we define $w(M, M') = m_1^{-1} + m_2^{-1}$.

For $M \in \mathcal{M}_{k,\ell}$, let

$$\mu(M) = \sum_{M':(M,M')\in E'_{k,\ell}} w(M,M'),$$

and for $M' \in \mathcal{M}_{k,\ell-1}$, let

$$\nu(M') = \sum_{M:(M,M')\in E'_{k,\ell}} w(M,M').$$

Fix $M \in \mathcal{M}_{k,\ell}$ and $(a,b) \in M \setminus M_0$, and suppose the cycles of $M \oplus M_0$ have size $2m_i$, for $1 \le i \le \ell$. If (a,b) is in a cycle of size $2m$ then there are $s \ge 2\alpha n - m - k$ edges (a',b') of $M \setminus M_0$ such that (a,b') and (a',b) are edges of G, and (a',b') and (a,b) are in different cycles. Putting $(a_1,b_1) = (a,b)$ and $(a_2,b_2) = (a',b')$ yields a member of $E'_{k,\ell}$. Apportioning weight m^{-1} to (a,b):

$$\begin{aligned}\mu(M) &\ge \sum_{i=1}^{\ell} m_i(2\alpha n - m_i - k)m_i^{-1} \\ &\ge (2\alpha\ell - 1)n - k\ell.\end{aligned}$$

Now fix $M' \in \mathcal{M}_{k,\ell-1}$ and suppose the cycles of $M' \oplus M_0$ have size $2m_i$, for $1 \le i \le \ell-1$. Fix a cycle C of size $2m$ in $M' \oplus M_0$. At worst, each pair of edges of $C \setminus M_0$ could contribute a pair (M,M') to $E'_{k,\ell}$. This observation gives

$$\begin{aligned}\nu(M) &\le \sum_{i=1}^{\ell-1} m_i \left(\sum_{j=2}^{m_i-2} \frac{1}{j} + \frac{1}{m_i - j} \right) \\ &\le \sum_{i=1}^{\ell-1} 2m_i \ln m_i \\ &\le 2n \ln n.\end{aligned}$$

Finally,

$$\begin{aligned}((2\alpha\ell - 1)n - k\ell)N_{k,\ell} &\le \sum_{(M,M')\in E'_{k,\ell}} w(M,M') \\ &\le 2n(\ln n)N_{k,\ell-1},\end{aligned}$$

and (b) follows.

Proof of Theorem 8. Let $N = |\mathcal{M}|$, and

$$\Delta = \sum_{k=0}^{n} \sum_{\ell=0}^{n} N_{k,\ell} 2^{\ell}.$$

Our aim is to find a uniform bound on Δ/N, which will also be a bound on $\gamma(G)$. Let $s_{k,\ell} = N_{k,\ell} 2^{\ell}$. It follows from Lemma 9(a) that

$$k\alpha s_{k,\ell} \leq \tfrac{1}{2} s_{k-2,\ell+1} + 2 s_{k-1,\ell}. \tag{13}$$

Let $S_k = \sum_{\ell=0}^{n} s_{k,\ell}$. Then inequality (13) implies $k\alpha S_k \leq \frac{1}{2} S_{k-2} + 2 S_{k-1}$. It follows by an easy induction on k that for $k > k_0 = \lceil 4/\alpha \rceil$,

$$S_k \leq \left(\frac{1+\sqrt{3}}{4} \right)^{k-k_0} (S_{k_0} + S_{k_0-1}),$$

and hence

$$\sum_{k=k_0}^{n} S_k = O(S_{k_0} + S_{k_0-1}). \tag{14}$$

Now assume $k \leq k_0$. From Lemma 9(b),

$$\frac{N_{k,\ell}}{N_{k,\ell-1}} \leq \frac{2 \ln n}{(2\alpha - k/n)\ell - 1} \leq \frac{1}{2},$$

provided

$$\ell \geq \ell_0 = \left\lceil \frac{4 \ln n + 1}{2\alpha - k_0/n} \right\rceil.$$

Thus, for $k \leq k_0$,

$$S_k \leq n s_{k,\ell_0} + \sum_{\ell=0}^{\ell_0} s_{k,\ell} \leq (n + \ell_0) 2^{\ell_0} N. \tag{15}$$

Hence, from (14) and (15), $\Delta/N = \sum_{k=0}^{n} S_k/N = \mathrm{O}(n^{1+(2 \ln 2)/\alpha})$.

6. Trustworthy approximation

We have seen that the KKLLL estimator provides an fpras for the permanent of a.e. 0,1-matrix, which is more efficient than the one proposed by Jerrum and Sinclair [10]. However, there is an important sense in which the results obtained by the latter approach are more "trustworthy" than those of the former. The aim of this section is to assign a precise meaning to this informal claim.

As usual, let G be a bipartite graph on vertex set $U + V$, and let $X(G)$ be the number of perfect matchings in G. Denote by $\widehat{X}(G)$ the number of "near-perfect matchings" in G, i.e., matchings that have precisely $n-1$ edges. Define

$$\rho(G) = \frac{\widehat{X}(G)}{X(G)}$$

provided $X(G) > 0$, and adopt the convention that $\rho(G) = \infty$ when $X(G) = 0$. The approximation scheme of Jerrum and Sinclair is known (Corollary 5.3 of [10]) to provide a reliable approximation to the number of perfect matchings in G in time polynomial in n, $\rho(G)$, and ε^{-1}. (Here, ε is the parameter controlling the accuracy of the approximation, and $\rho(G)$ is assumed to be known in advance.) Although it is possible to construct graphs G for which $\rho(G)$ is very large, it can be shown, using tools from Section 3, that such graphs are exceptional.

Corollary 10. *Almost every $G \in \mathcal{B}(n, p{=}\frac{1}{2})$ satisfies $\rho(G) \le 4n$.*

Proof. Assume the function $m = m(n)$ satisfies $m^2 n^{-3} \to \infty$ as $n \to \infty$, and select $G \in \mathcal{B}(n, m)$. The estimate

$$\frac{\operatorname{Exp}(\widehat{X}^2)}{(\operatorname{Exp}\widehat{X})^2} = 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)$$

is akin to that provided by Theorem 4 and can be proved by a similar argument. By applying Chebychev's inequality to X and $\widehat{X}$ in turn, we obtain

$$\Pr\big(X < \tfrac{4}{5}\operatorname{Exp} X\big) \to 0, \quad \text{as } n \to \infty, \tag{16}$$

and

$$\Pr\big(\widehat{X} > \tfrac{5}{4}\operatorname{Exp}\widehat{X}\big) \to 0, \quad \text{as } n \to \infty. \tag{17}$$

The expectation of $\widehat{X}$, obtained by computations similar to those appearing in the proof of Theorem 4, is

$$\operatorname{Exp}\widehat{X} = (n+1)!\Big(\frac{m}{n^2}\Big)^{n-1} \exp\Big\{-\frac{n^2}{m} + 1 + \mathrm{O}\Big(\frac{n^3}{m^2}\Big)\Big\};$$

comparing this formula with the existing one for $\operatorname{Exp} X$, we see that

$$\frac{m \operatorname{Exp}\widehat{X}}{n^3 \operatorname{Exp} X} \to 1, \quad \text{as } n \to \infty. \tag{18}$$

Combining (16), (17), and (18), we obtain

$$\Pr\left(\rho(G) \leq \frac{7n^3}{4m}\right) \to 1, \quad \text{as } n \to \infty.$$

(The constant $\frac{7}{4}$ here has no significance beyond its lying strictly between $(\frac{5}{4})^2$ and 2.) The corollary is obtained by translating this result to the $\mathcal{B}(n, p{=}\frac{1}{2})$ model using standard techniques.

A result related to Corollary 10 (but formally incomparable with it) may be found in [10].

So far we have seen nothing that distinguishes the two approaches in a qualitative sense. The efficiencies of the two approximation schemes depend on parameters, γ and ρ, which are large in the worst case, but small on average. However, the crucial point is that the condition "$\rho(G)$ is small" can be verified by a randomised polynomial-time algorithm with small error probability, whereas no such verification procedure is known for the condition "$\gamma(G)$ is small." (See the discussion following Theorem 5.3 of [10] for a precise precise statement of this claim.)

Following a suggestion of Joel Spencer, we may formalise the consequences of this apparent distinction. Let f be a function from input strings to natural numbers, and let $\mathcal{A}$ be a probabilistic algorithm that takes an input string x together with a real number $0 < \varepsilon < 1$, and returns a result Y (a random variable) that is either an approximation to $f(x)$ or a special "undefined symbol" $\perp$. For each n, the input strings of length n are assumed to be drawn from some specified probability distribution. A strong notion of what it means for $\mathcal{A}$ to work for almost every input is encapsulated in the following two conditions:

(1) $\Pr\big(Y = \perp \text{ or } (1-\varepsilon)f(x) \leq Y \leq (1+\varepsilon)f(x)\big) \geq 3/4$, for *every* x;

(2) $\Pr(Y \neq \perp) \to 1$, as $n \to \infty$, for *randomly selected* x with $|x| = n$.

The idea here is to separate the twin concerns of *reliability* and *range of applicability*, and give the former a higher status. Thus condition (1) demands that the response must be correct with high probability for *arbitrary* inputs, while condition (2) merely asks that an informative response should be provided with high probability for *random* inputs. As before, we may call such an algorithm *fully polynomial* if it runs in time polynomial in n and ε^{-1}.

The above definition crystalises an apparent distinction between the two known approximation schemes for the number of perfect matchings in a random graph. The approach via Markov chain simulation *does* lead to an approximation scheme that satisfies conditions (1) and (2) above, where x is interpreted as the encoding of a bipartite graph, $f(x)$ as the number of perfect matchings in x, and the probability distribution on inputs x is

given by the random graph model $\mathcal{B}(n, p{=}\frac{1}{2})$. (Full details may be found in the discussion following Theorem 5.3 of [10].) However, it is not known whether the same end could be achieved using the KKLLL estimator. The question is of some interest, since the latter approach is more likely to lead to a practical algorithm. The barrier appears to be the difficulty of obtaining estimates for the crucial parameter $\gamma(G)$.

7. Related results

Theorem 4 asserts that the permanent of a random 0,1-matrix is surprisingly tightly concentrated, provided we condition on the total number of 1s in the matrix. This result suggests that the task of estimating the permanent of a random matrix is not really as difficult as it seems at first sight. This feeling is reinforced by work of Rasmussen [17] in which it is shown that a particularly simple Monte Carlo estimator for the permanent performs well, at least for matrices with constant density.

Other enumeration problems on random inputs may also be less daunting than we might have supposed. Frieze and Suen [5] have presented a Monte Carlo algorithm for estimating the number of Hamiltonian cycles in a digraph. Using methods similar in spirit to those described here, they show that the algorithm efficiently produces reliable estimates when the input graph is selected randomly.

Acknowledgements

The second author is indebted to Marek Karpinski, László Lovász, Joel Spencer, and Mario Szegedy for useful discussions and advice, and to Alistair Sinclair for reading and commenting on an earlier version of the paper.

References

[1] Béla Bollobás, *Random Graphs*, Academic Press, 1985.

[2] Andrei Z. Broder, How hard is it to marry at random? (On the approximation of the permanent), *Proceedings of the 18th ACM Symposium on Theory of Computing*, 1986, pp. 50–58. Erratum in *Proceedings of the 20th ACM Symposium on Theory of Computing*, 1988, p. 551.

[3] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society **2** (1952) pp. 69–81.

[4] Martin Dyer, Alan Frieze, and Mark Jerrum, Approximately counting Hamilton cycles in dense graphs, *Proceedings of the 4th ACM-SIAM Symposium on Discrete Algorithms*, Society for Industrial and Applied Mathematics, 1994, pp. 336-343.

[5] Alan Frieze and Stephen Suen, Counting the number of Hamiltonian cycles in random digraphs, *Random Structures and Algorithms* **3** (1992), pp. 235–241.

[6] Peter GEMMELL and Madhu SUDAN, Highly resilient correctors for polynomials, *Information Processing Letters* **34** (1992), pp. 169–174.

[7] C. D. GODSIL and I. GUTMAN, On the matching polynomial of a graph, *Algebraic Methods in Graph Theory, I* (L. Lovász and V. T. Sós, editors), Colloquia Mathematica Societatis János Bolyai **25**, North-Holland, 1981.

[8] Marshall HALL Jr, *Combinatorial Theory*, Blaisdell, Waltham Massachusetts, 1967.

[9] Svante JANSON, The Number of Spanning Trees, Hamilton Cycles and Perfect Matchings in a Random Graph, *Combinatorics, Probability and Computing* **3** (1994), pp. 97–126.

[10] Mark JERRUM and Alistair SINCLAIR, Approximating the permanent, *SIAM Journal on Computing* **18** (1989), pp. 1149–1178.

[11] Mark R. JERRUM, Leslie G. VALIANT, and Vijay V. VAZIRANI, Random generation of combinatorial structures from a uniform distribution, *Theoretical Computer Science* **43** (1986), pp. 169–188.

[12] N. KARMARKAR, R. KARP, R. LIPTON, L. LOVÁSZ, and M. LUBY, *A Monte-Carlo Algorithm for Estimating the Permanent. SIAM Journal on Computing* **22** (1993), pp. 284–293.

[13] R. M. KARP and M. LUBY, Monte-Carlo algorithms for enumeration and reliability problems, *Proceedings of the 24th IEEE Symposium on Foundations of Computer Science*, 1983, pp. 56–64.

[14] László LOVÁSZ, *Combinatorial Problems and Exercises*, North-Holland, 1979.

[15] Milena MIHAIL, On coupling and the approximation of the permanent, *Information Processing Letters* **30** (1989), pp. 91–95.

[16] Henryk MINC, *Permanents*, Addison Wesley, 1978.

[17] Lars Eilstrup RASMUSSEN, Approximating the Permanent: a Simple Approach, *Random Structures and Algorithms* **5** (1994), pp. 349–361.

[18] L. G. VALIANT, The complexity of computing the permanent, *Theoretical Computer Science* **8** (1979), pp. 189–201.