LFCS

Laboratory for Foundations of Computer Science
Department of Computer Science - The University of Edinburgh

An Analysis of a Monte Carlo Algorithm for Estimating the Permanent

by

Mark Jerrum

LFCS Report Series ECS-LFCS-91-164

LFCS
Department of Computer Science
University of Edinburgh
The King's Buildings
Edinburgh EH9 3JZ

June 1991

An Analysis of a Monte Carlo Algorithm for Estimating the Permanent

Mark Jerrum†

June 1991

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 are necessary to obtain a reliable approximation that is within a constant factor of the correct value. Nevertheless, the same authors conjectured that for almost every 0,1-matrix the variance of the estimator is small. The conjecture is shown to be true; indeed for almost every 0,1-matrix, $O(n\,\omega(n)\epsilon^{-2})$ trials suffice to obtain a reliable approximation that is within a factor $(1+\epsilon)$ of the correct value. Here $\omega(n)$ is any function tending to infinity as $n \to \infty$.

---

† Department of Computer Science, University of Edinburgh, The King's Buildings, Edinburgh EH9 3JZ, UK. Part of the work described here was 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

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

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, $\mathrm{per}(A)$ 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 [11] 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 [13], 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* [8] for f is a probabilistic algorithm that takes as input a string x and a real number $0 < \epsilon < 1$, and produces as output a number Y (a random variable) such that $(1-\epsilon)f(x) \leq Y \leq (1+\epsilon)f(x)$ 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 [6, 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 on this front, it seems reasonable to weaken the requirements further, and ask whether there exists an fpras for $\mathrm{per}(A)$ 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 [5, 12], 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, 10]. A brief discussion of this result, including a description of the class $\mathcal{G}$, can be found in the final section of the paper. The polynomial bounding the execution time of the algorithm of Jerrum and Sinclair was not explicitly computed in [5], but its degree is not small. It is not yet clear whether this approach could 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 [7]. Their algorithm is based on an unbiased estimator for $\mathrm{per}(A)$, 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 $\mathrm{per}(A)$.

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, $\mathrm{O}(n\epsilon^{-2})$ trials are sufficient to accomplish the same task when G is the complete bipartite graph K_{nn} [7].

Karmarkar et al. conjecture that it is the second of these two examples that is the more characteristic of graphs in general, and that $\mathrm{O}(n\epsilon^{-2})$ trials suffice for a.e. G. It is a consequence of the main result of this paper that something very close to the conjecture is true: namely that $n\omega(n)\epsilon^{-2}$ trials suffice for a.e. G, where $\omega(n)$ is any function tending to infinity as $n \to \infty$. A more precise statement of the result 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 $0 \leq i, j \leq n-1$:

   (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 [3], 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 [7].

Theorem 1. $\mathbf{E}(Z) = \mathrm{per}(A)$. □

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 $\gamma(G) = \mathbf{E}\left(2^{c(M,M')}\right)$ 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.)

$$\frac{\mathbf{E}(Z^2)}{\mathbf{E}(Z)^2} = \gamma(G).$$

Proof. The theorem is essentially a restatement of Theorem 4 of [7]. 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 $O(\epsilon^{-2}\gamma(G))$ 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 $t = \lceil 4\epsilon^{-2}\gamma(G) \rceil$ trials with the KKLLL estimator, letting the results be $Z_0, Z_1, \ldots, Z_{t-1}$. Using Theorem 2,

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

Hence, by Chebychev's inequality,

$$\Pr\left((1-\epsilon)\,\mathbf{E}(Z) \leq \frac{1}{t}\sum_{i=0}^{t-1} Z_i \leq (1+\epsilon)\,\mathbf{E}(Z)\right) \geq \frac{3}{4}. \quad \square$$

The important point about Corollary 3 is that it reduces the analysis of the KKLLL approximation scheme on random inputs to the analysis of $\gamma(G)$ for randomly chosen G. The latter will be our goal for the remainder of the paper.

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 the uniform one. 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 $\gamma(G) = \mathbf{E}\left(2^{c(M,M')}\right)$, 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 $0 \leq k \leq n$ 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'.

We shall see that 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. It is worth remarking that no analogous theorem holds for the random graph model $\mathcal{B}(n,p)$, in which potential edges are selected independently and with probability p; it is for precisely this reason that we have chosen to work with the former model.

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

$$\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2} = 1 + O\Big(\frac{n^3}{m^2}\Big).$$

Proof. 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

$$\mathbf{E}(X) = \sum_M \mathbf{E}(X_M), \tag{1}$$

and

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

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 $t \leq 2n$. 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

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

and hence, from equation (1),

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

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 [4, 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:

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

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

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

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 main result

We are now ready to tackle the main result, which states 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 [9, 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. 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 $\Pr\left(X \leq \frac{3}{4}\,\mathbf{E}(X)\right) = \mathrm{O}(n^3 m^{-2})$, which is clearly equivalent to $\Pr(X^2 \leq \frac{9}{16}\,\mathbf{E}(X)^2) = \mathrm{O}(n^3 m^{-2})$. A second application of Theorem 4 then yields

$$\Pr\left(X^2 \leq \tfrac{1}{2}\,\mathbf{E}(X^2)\right) = \mathrm{O}\Big(\frac{n^3}{m^2}\Big). \tag{11}$$

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 $\left[\delta - \mathrm{O}(n^3 m^{-2})\right]N$ graphs simultaneously satisfy the conditions $\gamma(G) > n\delta^{-1}$ and $X(G)^2 > \frac{1}{2}\,\mathbf{E}(X^2)$. Now observe that inequality (10) may be recast in the form

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

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

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

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

It should be clear that the event $\gamma(G) \leq n\delta^{-1}$ appearing in the statement of Theorem 5 may be replaced by $\gamma(G) \leq an\delta^{-1}$, 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 $\Pr\left(\gamma(G) \leq n\delta^{-1}\right) \geq 1 - \delta$.

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 $\gamma(G) \leq n\,\omega(n)$. □

5. 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 [5]. 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 [5]) 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 8. Almost every $G \in \mathcal{B}(n, p{=}\frac{1}{2})$ satisfies $\rho(G) \leq 4n$.

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

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

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

$$\Pr\left(X < \tfrac{4}{5}\,\mathsf{E}(X)\right) \to 0, \quad \text{as } n \to \infty, \tag{12}$$

and

$$\Pr\left(\widehat{X} > \tfrac{5}{4}\,\mathsf{E}(\widehat{X})\right) \to 0, \quad \text{as } n \to \infty. \tag{13}$$

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

$$\mathsf{E}(\widehat{X}) = (n+1)!\left(\frac{m}{n^2}\right)^{n-1} \exp\left\{-\frac{n^2}{m} + 1 + \mathrm{O}\left(\frac{n^3}{m^2}\right)\right\};$$

comparing this formula with the existing one for $\mathsf{E}(X)$, we see that

$$\frac{m\,\mathsf{E}(\widehat{X})}{n^3\,\mathsf{E}(X)} \to 1, \quad \text{as } n \to \infty. \tag{14}$$

Combining (12), (13), and (14), we obtain

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

(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 8 (but formally incomparable with it) may be found in [5].

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 [5] 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 real numbers $0 < \delta, \epsilon < 1$, 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) $\Pr\big(Y = \perp \text{ or } (1-\epsilon)f(x) \leq Y \leq (1+\epsilon)f(x)\big) \geq 1-\delta$, for *every* x;

(2) $\Pr(Y \neq \perp) \geq 1-\delta$, for every n and *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, ϵ^{-1}, and δ^{-1}. (Since there is no obvious 'powering lemma' for failure probabilities under this definition, δ must appear as an explicit input parameter.)

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 [5].) 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 $\gamma(G)$.

Acknowledgements

I am indebted to Alan Frieze, Marek Karpinski, László Lovász, Joel Spencer, and Mario Szegedy for useful discussions and advice, and to Alistair Sinclair for reading and commenting on a preliminary 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] 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.

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

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

[6] 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.

[7] N. KARMARKAR, R. KARP, R. LIPTON, L. LOVÁSZ, and M. LUBY, *A Monte-Carlo Algorithm for Estimating the Permanent*, preprint 1988.

[8] 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.

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

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

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

[12] Alistair SINCLAIR, *Randomised Algorithms for Counting and Generating Combinatorial Structures*, PhD Thesis, University of Edinburgh, June 1988.

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