Chapter 6

Some other approaches

In this chapter we describe some approximate counting problems that can be solved without the uuse of Markov chains.

6.1 Satisfiability

Here we are given a Boolean function F in *Disjunctive Normal Form* (DNF) e.g.

$$F = x_1x_2x_3 + \bar{x}_1x_2x_4 + x_3\bar{x}_5x_6x_7$$

and our task is to estimate the number of satisfying assignments of the variables.

Thus assume we have n Boolean variabes, $x_1, x_2, \ldots, x_n$ and

$$F = m_1 + m_2 + \cdots + m_r \text{ where } m_i = \prod_{j=1}^{n} x_j^{\alpha_{i,j}}$$

and $\alpha_{i,j} \in \{0, \pm 1\}$ and $x_j^0 = 1, x_j^1 = x_j, x_j^{-1} = \bar{x}_j$.

Let $A = \{0,1\}^n$ be the set of possible assignments of $0/1$ values to the variables and let $A^* = \{a \in A : \ F(a) = 1\}$. The task is then to estimate $|A^*|$.

Let $A_i = \{a \in A : \ m_i(a) = 1\}$, $i = 1, 2, \ldots, r$. Then we are faced with the following problem:
Cardinality of the Union Problem
Given sets $A_1, A_2, \ldots, A_r \subseteq A$, estimate $|A^*|$ where $A^* = \bigcup_{i=1}^{r} A_i$.

The approach given here is valid if

(i) r is small, i.e. polynomial in the description of the problem.

(ii) $|A_i|$ is known for each i.

(iii) It is possible to efficiently choose a random element of A_i for each i.

(iv) It is possible to efficiently decide whether a given $a \in A$ lies in A_i for each i.

For the problem to be interesting, the A_i need to be large, i.e. exponential in the description of the problem.
It is clear that the DNF problem satisfies the conditions (i)–(iv).

Another way of looking at this problem is that we have a $r \times |A|$ 0/1 matrix M where $M(i,a) = 1$ iff $a \in A_i$. Let ρ_i be the present the number of 1's in row i and let $\rho = \rho_1 + \rho_2 + \cdots + \rho_r$ be the total number of 1's in M. Let c_a denote the number of 1's in column a and let $A^* = \{a : c_a > 0\}$.. Now we want to estimate $\nu = |A^*|$. Consider the following algorithm: N is a parameter to be determined later.

MATRIX COLUMN WEIGHT ALGORITHM

begin
  Compute $p_i = \rho_i/\rho$ for $i = 1, 2, \ldots, r$.
  For $t = 1$ **to** N **do**
  begin
    (1a) Choose i_t randomly from $[r]$ according to distribution $p_1, p_2, \ldots, p_r$.
    (1b) Choose a_t randomly from $\{a : M(i_t, a) = 1\}$.
    (1c) Compute $Z_t = \rho/c_{a_t}$.
  Output $\bar{Z} = \frac{Z_1+Z_2+\cdots+Z_N}{N}$.
  end
end

The next lemma evaluates the accuracy of this procedure.

Lemma 6.1.1

$$\mathbf{Pr}(|\bar{Z} - \nu| \geq \epsilon\nu) \leq \frac{r}{\epsilon^2 N}.$$

Proof Fix $t = 1$. We claim that (t, a_t) is chosen uniformly at random from the set $\{(i,a) : M(i,a) = 1\}$. Indeed (1a) determines row i_t with probability proportional to the number of 1's in a row and then (1b) chooses a random member of the row. Thus we see that for $a \in A$

$$\mathbf{Pr}(a_1 = a) = \frac{c_a}{\rho}.$$

Thus

$$\mathbf{E}(Z_1) = \sum_{a \in A} \mathbf{E}(Z_1 \mid a_t = a)\mathbf{Pr}(a_t = a) = \sum_{a \in A^*} \frac{\rho}{c_a} \cdot \frac{c_a}{\rho} = |A^*|.$$

Now we estimate the variance.

$$\mathbf{E}(Z_1^2) = \sum_{a \in A^*} \frac{\rho^2}{c_a^2} \cdot \frac{c_a}{\rho} = \rho \sum_{a \in A^*} \frac{1}{c_a} \leq \rho\nu \leq r\nu^2.$$

Thus

$$\mathbf{Var}(Z_1) \leq r\mathbf{E}(Z_1)^2$$

which implies

$$\mathbf{Var}(\bar{Z}) \leq \frac{r}{N}\mathbf{E}(\bar{Z})^2$$

and the result follows from the Chebychef inequality. □

Putting $N = 4r\epsilon^{-2}$ we see that

$$\mathbf{Pr}(|\bar{Z} - \nu| \geq \epsilon\nu) \leq \frac{1}{4}.$$

This probability can be reduced to δ as we did in (1.2) by repeating the algorithm $\lceil 12 \ln(2/\delta) \rceil$ times and taking the median result.

6.1.1 Random assignments

We consider a generalisation of the DNF problem which will be useful in Section 6.2. We consider the problem of estimating the probability that a randomly generated assignment a satisfies a Boolean formula F, given in DNF. Thus we are given $0 < p < 1$ and suppose that assignment $a \in A$ is chosen by independently putting $x_j = 1$ with probability p. We wish to estimate

$$\Delta_p = \mathbf{Pr}(F(a) = 1).$$

If $p = 1/2$ then $\Delta_p = 2^{-n}|A^*|$ and so this problem generalises the problem of the previous section.

For $a \in A$ let $s(a) = |\{j : a_j = 1\}$. Let $A_k = \{a \in A : s(a) = k\}$ and $A_k^* = A^* \cap A_k$ for $k = 0, 1, 2, \ldots n$. Then

$$\Delta_p = \sum_{k=0}^{n} |A_k^*| p^k (1-p)^{n-k}$$

and so we can estimate Δ_p efficiently if we can estimate the $|A_k|$ efficiently. So let $A_{i,k} = \{a \in A_k : m_i(a) = 1\}$ so that $A_k = \bigcup_{i=1}^r A_{i,k}$. It remains to check that the sets $A_{i,k}$ satisfy requirements (i)–(iv) above. (i) holds. Let $n_i = |\{j : \alpha_{i,j} = 1\}|$ and $\bar{n}_i = |\{j : \alpha_{i,j} = -1\}|$ then $|A_{i,k}| = \binom{n-n_i-\bar{n}_i}{k-n_i}$ and so (ii) holds. This calculation shows that (iii) holds and (iv) still holds. It follows that the $|A_k^*|$ can be estimated, along with Δ_p.

It is more efficient to estimate Δ_p directly by modifying the MATRIX COLUMN WEIGHT ALGORITHM. We briefly indicate the steps. We write

$$\Delta_p = (1-p)^n \sum_{a \in A^*} x^{s(a)} \qquad \text{where } x = \frac{p}{1-p}.$$

Then define the $r \times |A|$ matrix M with $M(i,a) = x^{s(a)}$ if $m_i(a) = 1$ and $M(i,a) = 0$ otherwise. Now define $\rho_i = \sum_{a \in A} M(i,a)$ and p_i and ρ as above. The only change that we need to make to the algorithm is to replace (1b) by
(1b') Choose a_t from $\{a : M(i_t, a) = 1\}$ with probability proportional to $x^{s(a)}$.
With these changes, (i_t, a_t) is chosen with probability proportional to $x^{s(a)}$ and the output of the algorithm has expectation equal to $\sum_{a \in A^*} x^{s(a)}$. The proof of Lemma 6.1.1 goes through minor changes.

6.2 Reliability

Here we are given a graph $G = (V,E)$, $n = |V|, m = |E|$ and a probability $0 < p < 1$. Let $G_p = (V, E_p)$ be the random subgraph of G obtained by independently including each $e \in E$ with probability p. This models a network where each link (edge) *fails* independently with probability $q = 1-p$. The task is to estimate

$$FAIL(p) = \mathbf{Pr}(G_p \text{ is not connected}.$$

If $FAIL(p)$ is large then this can be estimated easily. We simply generate random copies $G_1, G_2, \ldots, G_N$ of G_p and let

$$\delta_i = \begin{cases} 1 & G_i \text{ is not connected} \\ 0 & G_i \text{ is connected} \end{cases}$$

We can then estimate $FAIL(p)$ by $\bar{\delta} = \frac{\delta_1 + \cdots + \delta_N}{N}$ where N is given in Lemma 6.2.1 below.

Let κ denote the minimum size of a cut in G i.e. $\min_{S \subseteq V} |S : \bar{S}|$ where $S : \bar{S}$ is the set of edges with one end in S and the other in $\bar{S} = V \setminus S$. We see immediately that

$$FAIL(p) \geq q^\kappa$$

since this is the probability that any given minimum cut *fails* in G_p i.e. contains no G_p edges.

Lemma 6.2.1 *Assume $q^\kappa \geq n^{-4}$ and let $N = 4n^4\epsilon^{-2}$. Then*

$$\mathbf{Pr}(|FAIL(p) - \bar{\delta}| \geq \epsilon FAIL(p)) \leq \frac{1}{4}.$$

Proof $\bar{\delta}$ has mean $\phi = FAIL(p)$ and variance $\phi(1-\phi)N^{-1}$ and so by the Chebychef inequality

$$\mathbf{Pr}(|FAIL(p) - \bar{\delta}| \geq \epsilon FAIL(p)) \leq \frac{\phi(1-\phi)}{\epsilon^2\phi^2 N}$$

and the result follows from $\phi \geq n^{-4}$. □

Note that each δ_i can be computed in $O(m)$ time and that κ can be computed in $O(n^3)$ time so that we can decide when to use Lemma 6.2.1.

The interesting case is of course when $FAIL(p)$ is small. A cut $S : \bar{S}$ with $\alpha\kappa$ edges is called an α-minimum cut. The algorithm we describe rests on the following two theorems:

Theorem 6.2.1 *G has at most $12n^{2\alpha}$ cuts of size at most $\alpha\kappa$.*

Proof (Deferred to Section 6.3).

Theorem 6.2.2 *Suppose $q^\kappa = n^{-(2+\delta)}$ for some $\delta > 0$. Then*

$$\mathbf{Pr}(\exists \text{ an } (\geq \alpha)\text{-minimum cut which fails}) \leq n^{-\alpha\delta}\gamma 12^\gamma$$

where $\gamma = 1 + 2/\delta$.

Proof (Deferred to Section 6.3).

Now consider the following algorithm:

$$q^\kappa = n^{-(2+\delta)} \text{ and } \alpha_0 = 2 - \frac{\log(\epsilon/1000)}{\log n}.$$

Reliability algorithm

1. Enumerate the cuts $S_i : \bar{S}_i$, $i = 1, 2, \ldots, \nu$ of size at most $\alpha_0\kappa$.
2. Compute an $\epsilon/2$-approximation Φ to

$$\Phi_0 = \mathbf{Pr}(\exists 1 \leq i \leq \nu : \; S_i : \bar{S}_i \text{ fails}).$$

3. Output Φ.

The cuts $S_i : \bar{S}_i$, $i = 1, 2, \ldots, \nu$ can be found in polynomial time (see Section 6.3).

Step 2 is executed as follows: Suppose we assign a set of Boolean variables $x_e.e \in E$. Consider the Boolean formula

$$F = F_1 + F_2 + \cdots + F_\nu$$

where

$$F_i = \prod_{e \in S_i : \bar{S}_i} x_e.$$

The edges of G_p define a (random) assignment of values to the x_e i.e. $x_e = 1$ iff edge e does *not* occur in G_p. Then $F_i = 1$ iff cut $S_i : \bar{S}_i$ fails and so $F = 1$ iff $\exists i \;\; S_i : \bar{S}_i$ fails. Thus we can use the algorithm of Section 6.1.1 to compute an $\epsilon/2$-approximation to $\Phi_0 = \mathbf{Pr}(F = 1)$ and so carry out Step 2. We execute this algorithm so the probability of failure is at most $\frac{1}{4}$. Thus the RELIABILITY ALGORITHM can be executed in polynomial time. It remains to prove

Theorem 6.2.3 *If* $q^\kappa = n^{-(2+\delta)}$, $\delta \geq 2$, *then*

$$\mathbf{Pr}(|\Phi - FAIL(p)| \geq \epsilon FAIL(p)) \leq \frac{1}{4}.$$

Proof Let

$$\Phi_1 = \mathbf{Pr}(\exists \text{ an } (\geq \alpha)\text{-minimum cut which fails}).$$

It follows from Theorem 6.2.2 that

$$\frac{\Phi_1}{FAIL(p)} \leq \frac{288 n^{-\alpha_0 \delta}}{n^{-(2+\delta)}} \leq \frac{\epsilon}{3}.$$

Now

$$\Phi_0 \leq FAIL(p) \leq \Phi_0 + \Phi_1$$

and so the $\frac{\epsilon}{2}$-approximation Φ to Φ_0 is an ϵ-approximation to $FAIL(p)$ and the theorem follows. □

6.3 Deferred Proofs

6.3.1 Proof of Theorem 6.2.1

We use the following CONTRACTION ALGORITHM to produce a cut. Each cut of size at most $\alpha\kappa$ will have probability at least $\frac{1}{12}n^{-2\alpha}$ of being chosen and the theorem follows.

CONTRACTION ALGORITHM

begin
    $k = \lceil 2\alpha \rceil$, $H \leftarrow G$.
    while $|V(H)| > k$ **do**
    begin
A   Choose e randomly from $E(H)$.

$H \leftarrow H \setminus e$ – contract e.

end

B **begin**

Let $K : \bar{K}$ be a random partition of $V(H)$ into 2 non-empty subsets.

"Expand" K into $S \subseteq V(G)$.

Output S

end

end

We need to explain "expand" K. When we contract edge $\{v, w\}$, the two vertices v, w are replaced by a single new vertex. Thus, in general, the vertices of H at Step B correspond to (disjoint) subsets of V. Thes $S = \bigcup_{v \in K} v$.

We note next that the minimum cut size of H is at least κ throughout. (H contains parallel edges.) This because the cutsets of H are a subset of the cutsets of G. In particular, H has minimum degree at leat κ.

We now consider a fixed α-minimum cut C of G. We will output S if (i) no edge of C is chosen at Step A and (ii) the contracted version of C is chosen at Step B.

After t executions of Step A, H will have $n - t$ vertices. Assume that no edge of C has been contracted. H has at least $\frac{1}{2}(n-t)\kappa$ edges and so the probability we do not choose $e \in C$ at the next iteration is at least $1 - \frac{2\alpha}{n-t}$. Thus the probability we choose C is at least

$$2^{1-k} \prod_{r=k+1}^{n} \left(1 - \frac{2\alpha}{r}\right) = \prod_{i=0}^{k-1} \frac{2k - 2\alpha - i}{n-i} \prod_{i=k}^{n-k-1} \frac{n - i + k - 2\alpha}{n-i}$$
$$\geq \frac{2^{1-k}k!}{n^k} \exp\left\{ \sum_{i=k}^{n-k-1} \frac{f}{n-i} - \frac{1}{2} \sum_{i=k}^{n-k-1} \frac{f^2}{(n-i)^2} \right\}$$

where $f = k - 2\alpha$. Now

$$\exp\left\{ \sum_{i=k}^{n-k-1} \frac{f}{n-i} - \frac{1}{2} \sum_{i=k}^{n-k-1} \frac{f^2}{(n-i)^2} \right\} \geq n^f (ek)^{-1} e^{-\pi^2/12}$$

and so the probability we choose C is at least

$$\frac{2^{1-k}k!}{ekn^{2\alpha}e^{\pi^2/12}} \geq \frac{1}{12n^{2\alpha}}.$$

□

We see immediately that if we run the CONTRACTION ALGORITHM $O(n^{2\alpha} \log n)$ times then **whp** we will produce all cuts of size $\alpha\kappa$ or less.

6.3.2 Proof of Theorem 6.2.2

Let $\kappa = \kappa_1 \leq \kappa_2 \leq \cdots \leq \kappa_r$ be an enumeration of the cut sizes in G. We bound

$$\Delta = \sum_{i=i_0}^{r} q^{\kappa_i} \qquad \text{where } i_0 = \min\{i : \ \kappa_i \geq \alpha\kappa\}$$

which bounds the probability that a large cut fails.

Theorem 6.2.1 implies that

$$\kappa_i \geq \max\left\{\alpha, \frac{\log(i/12)}{2\log n}\right\}\kappa \qquad \text{for } i \geq i_0.$$

Thus

$$\begin{aligned}
\Delta &\leq \sum_{i\leq 12n^{2\alpha}} q^{\alpha\kappa} + \sum_{i>12n^{2\alpha}} n^{-(2+\delta)\log(i/12)/(2\log n)} \\
&\leq 12n^{-\alpha\delta} + \int_{12n^{2\alpha}}^{\infty} \left(\frac{x}{12}\right)^{-(1+\delta/2)} dx \\
&\leq 12n^{-\alpha\delta} + 12^{1+\delta/2}(12n^{2\alpha})^{-\delta/2}
\end{aligned}$$

and the result follows.

6.4 Tutte Polynomial in Dense Graphs

For a graph G, the *Tutte Polynomial* $T_G(x,y)$ is a bivariate polynomial which for many values of x, y evaluates to interesting graph invariants e.g. $T_G(1,1)$ equals the number of spanning trees of G. We define it here by

$$T_G(x,y) = \sum_{A\subseteq E}(x-1)^{\kappa(A)-1}(y-1)^{|A|+\kappa(A)-n} \tag{6.1}$$

where $\kappa(A)$ is the number of components of $G_A = (V, A)$.

Some more interesting evaluations

- $T_G(2,1)$ is the number of forests of G.
- $T_G(1,2)$ is the number of forests of edge sets which contain a spanning tree of G.
- $T_G(2,0)$ is the number of orientations of the edges of G which do not contain a directed cycle.

- $(-1)^{n-\kappa(E)}\lambda^{\kappa(E)}T_G(1-\lambda)$ is the chromatic polynomial of G i.e. the coefficient of λ^k in this polynomial is the number of proper k-colourings of the vertices of G.

- $1 - FAIL(p) = q^{|E|-n+1}p^{n-1}T_G(1, 1/q)$ when G is connected.

It turns out that the hyperbolae H_α defined by

$$H_\alpha = \{(x, y) : (x-1)(y-1) = \alpha\}$$

play a special role in the theory.

- Along H_1, $T_G(x, y) = x^{|E|}(x-1)^{n-\kappa(E)-|E|}$.

- Along H_Q, for general positive integer Q, T_G specialises to the partition function of the Potts model of statistical physics.

There are several other important evaluations. Given the expressive power of this polynomial, it is not surprising that apart from a few special points and 2 special hyperbolae, the exact evaluation of T_G is $\#P$-hard even for the very restricted class of planar bipartite graphs. Here we consider *dense* graphs and prove the existence of an FPRAS for $T_G(x, y)$ whenever $x, y > 1$.

For $0 < \alpha < 1$, let $\mathcal{G}_\alpha$ denote the set of graphs $G = (V, E)$ with $|V| = n$ and minimum degree $\delta(G) \geq \alpha n$. A graph is α-dense if it is a member of $\mathcal{G}_\alpha$ or, somewhat loosely, dense if we omit the α.

A first easy, but essential, observation is the following. Let G_p denote the random graph obtained by selecting edges of G independently with probability p.

Lemma 6.4.1 *Assume G is connected with n vertices and m edges. Assume $x, y > 1$ and let $p = (y-1)/y$ and $Q = (x-1)(y-1)$. Let $\kappa = \kappa(G_p)$ be the number of components of G_p. Then*

$$T_G(x, y) = \frac{y^m}{(x-1)(y-1)^n}\mathbf{E}(Q^\kappa).$$

Proof It follows from (6.1) that

$$\begin{aligned} T_G(x, y) &= \frac{y^m}{(x-1)(y-1)^n}\sum_{A\subseteq E}\left(\frac{y-1}{y}\right)^{|A|}\left(\frac{1}{y}\right)^{m-|A|}((x-1)(y-1))^{\kappa(A)} \\ &= \frac{y^m}{(x-1)(y-1)^n}\sum_{A\subseteq E}Q^{\kappa(A)}\mathbf{Pr}\{G_p = G_A\}. \end{aligned}$$

$\square$

We now describe a property of dense graphs which is the key to much of the ensuing analysis. Let $N(v), v \in V$ denote the set of neighbours of v. Define $G^* = (V, E^*)$ by $(u, v) \in E^*$ if and only if $|N(u) \cap N(v)| \geq \alpha^2 n/2$. Let

$$s = \lceil 2/\alpha \rceil - 1.$$

Lemma 6.4.2 *Among any s+1 vertices of G, there are two which are adjacent in G^*.*

Proof Suppose there exist $v_1, v_2, \ldots v_{s+1}$ such that $|N(v_i) \cap N(v_j)| < \alpha^2 n/2$ if $i \neq j$. But then

$$\begin{aligned} \left| \bigcup_{i=1}^{s+1} N(v_i) \right| &\geq \sum_{i=1}^{s+1} |N(v_i)| - \sum_{i \neq j} |N(v_i) \cap N(v_j)| \\ &> (s+1)\alpha n - \binom{s+1}{2} \frac{\alpha^2 n}{2} \\ &= (s+1)\alpha n \left(1 - \frac{s\alpha}{4}\right) \\ &\geq n. \end{aligned}$$

□

Let $\hat{Q} = \max\{Q, Q^{-1}\}$ and $\zeta = y^m/((x-1)(y-1)^n)$.

We claim that the following algorithm estimates $T_G(x, y)$ for $G \in \mathcal{G}_\alpha$.

Algorithm EVAL
begin
  $p := \frac{y-1}{y}$; $Q := (x-1)(y-1)$; $t := \lceil 16\hat{Q}^{2s}\epsilon^{-2} \rceil$;
  for $i = 1$ **to** t **do**
  begin
    Generate G_p; $Z_i := Q^{\kappa(G_p)}$
  end
  $\tilde{Z} := \frac{Z_1 + Z_2 + \cdots + Z_t}{t}$;
  Output $Z = \zeta \tilde{Z}$
end

We first prove

Lemma 6.4.3 *In the notation of Lemma 6.4.1, let*

$$n_0 = \min \left\{ n : n \geq \max \left\{ \frac{24 \ln(n\hat{Q})}{\alpha^2 p^2}, Q^{20/\alpha^2} \right\} \right\}.$$

If $n \geq n_0$ then

$$Q \geq 1 \textit{ implies } \mathbf{E}(Q^{2\kappa}) \leq 2Q^{2s}.$$
$$Q < 1 \textit{ implies } \mathbf{E}(Q^{\kappa}) \geq Q^{s}/2.$$

Proof Let $\mathcal{E}_u$ denote the event $\{\kappa(G_p) \geq u+s+1\}$ for $1 \leq u \leq u_0 = \lfloor \alpha^2 n/8 \rfloor$. If $\mathcal{E}_u$ occurs we choose $X = \{x_1, x_2, \ldots, x_{u+s+1}\}$ with each x_i from a different component of G_p. Lemma 6.4.2 implies that we can choose $y_1, y_2 \in X$ such that y_1, y_2 are adjacent in G^*. Repeating the argument yields a matching $\{y_1, y_2\}, \ldots, \{y_{2t-1,2t}\}$ in G^* where $t = \lceil (u+1)/2 \rceil$ and $y_1, y_2, \ldots, y_{2t}$ each lie in different components of G_p. The probability that G_p contains no path of length 2 connecting y_{2i-1} to y_{2i} for each i, $1 \leq i \leq t$ is at most $(1-p^2)^K$, where $K = (\alpha^2 n/2 - 2u)t$. Hence for $u \leq u_0, n \geq n_0$

$$\mathbf{Pr}(\mathcal{E}_u) \leq n^{2t}(1-p^2)^K \leq (n^2 e^{-\alpha^2 p^2 n/8})^u.$$

Thus for $u \leq u_0, n \geq n_0$

$$\mathbf{Pr}(\mathcal{E}_u) \leq (n^2 \exp\{-3\ln(n\hat{Q})\})^u = n^{-u}\hat{Q}^{-3u}.$$

Suppose first that $Q \geq 1$. Then

$$\begin{aligned}
\mathbf{E}(Q^{2\kappa}) &\leq Q^{2s}\left(1 + Q^2 \sum_{u=1}^{u_0} Q^{2u}\mathbf{Pr}(\mathcal{E}_u)\right) + Q^{2n}\mathbf{Pr}(\mathcal{E}_{u_0}) \\
&\leq Q^{2s}\left(1 + Q^2 \sum_{u=1}^{u_0} (n^{-1}Q^{-1})^u\right) + Q^{2n} n^{-\alpha^2 n/8} \\
&\leq 2Q^{2s}.
\end{aligned}$$

Suppose now that $Q < 1$. Then

$$\begin{aligned}
\mathbf{E}(Q^{\kappa}) &\geq Q^s(1 - \mathbf{Pr}(\mathbf{E}_1)) \\
&\geq Q^s/2
\end{aligned}$$

for $n \geq n_0$. □

Theorem 6.4.1 *For fixed rational x, y, and $\epsilon > 0$, if $T = T_G(x,y)$ and Z is the output of Algorithm EVAL, then*

$$\mathbf{Pr}(|Z - T| \geq \epsilon T) \leq \frac{1}{4}.$$

Proof Since $Z = \zeta\left(\frac{Z_1 + \ldots + Z_t}{t}\right)$, from Lemma 6.4.1 we see that $T = \mathbf{E}(Z)$. From Chebychev's inequality

$$\mathbf{Pr}\{|Z - T| \geq \epsilon T\} \leq \frac{\mathbf{Var}(Z)}{\epsilon^2 T^2} \leq \frac{\zeta^2}{\epsilon^2 t}\frac{\mathbf{Var}(Z_i)}{T^2} \leq \frac{\zeta^2}{\epsilon^2 t}\frac{\mathbf{E}(Z_i^2)}{T^2}.$$

Case $Q < 1$

Lemma 6.4.3 gives

$$\mathbf{E}(Z_i^2) = \mathbf{E}(Q^{2\kappa(G_p)}) \leq 1.$$

$$T^2 = \zeta^2(\mathbf{E}(Z_i))^2 = \zeta^2(\mathbf{E}(Q^{\kappa(G_p)}))^2 \geq \zeta^2 Q^{2s}/4.$$

giving

$$\mathbf{Pr}\{|Z - T| \geq \epsilon T\} \leq \frac{4}{\epsilon^2 t Q^{2s}}.$$

Case $Q \geq 1$

$$\mathbf{Pr}\{|Z - T| \geq \epsilon T\} \leq \frac{\zeta^2}{\epsilon^2 t}\frac{\mathbf{E}(Q^{2\kappa})}{T^2} \leq \frac{2Q^{2s}}{\epsilon^2 t}$$

using Lemma 6.4.3, and noticing that for $Q \geq 1$, $T \geq \zeta$.

The result follows provided

$$t \geq \frac{16}{\epsilon^2 Q^{2s}} \quad (Q < 1) \text{ and } t \geq \frac{8Q^{2s}}{\epsilon^2} \quad (Q \geq 1),$$

which it is by choice of t in EVAL. □

Note: although polynomially bounded the running time grows when $(x-1)(y-1)$ or its inverse grow.

6.5 Permanent via Determinant

We consider here an algorithm for estimating the permanent of a 0-1 matrix based on evaluating a determinant. Let A be an $n \times n$ 0-1 matrix:

The KKLLL estimator
The estimator is defined to be the random variable Z that results from the simple experiment described below. The idea is due to Karmarker, Karp, Lipton, Lovász and Luby [**?**] and it is an improvement on a method due to Godsil and Gutman [**?**].

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

- **(a)** If $A_{i,j} = 0$ then set $B_{i,j}$ equal to 0;
- **(b)** If $A_{i,j} = 1$ then choose $B_{i,j}$ independently and randomly from the set $\{1, \omega, \omega^2\}$.

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

Let G be the bipartite graph on vertex set $U+V$, where $U=V=[n]$ and (i,j) is an edge of G iff $A_{i,j}=1$. Let $\mathcal{M}$ denote the set of all perfect matchings in G. Clearly, $\operatorname{per} A = |\mathcal{M}|$. For $M \in \mathcal{M}$ let $\operatorname{sgn}(M)$ be the sign of the associated permutation σ_M where $\sigma_M(i)=j$ iff $(i,j)\in\mathcal{M}$. Let $\beta(M)=\prod_{(i,j)\in M} B_{i,j}$. With these notational definitions we can prove that Z is an unbiassed estimator of $\operatorname{per} A$.

Theorem 6.5.1

$$\mathbf{E}(Z) = \operatorname{per} A.$$

Proof

$$\begin{aligned} Z &= \sum_{M,M'\in\mathcal{M}} \operatorname{sgn}(M)\operatorname{sgn}(M')\beta(M)\overline{\beta(M')} \\ &= \sum_{M\in\mathcal{M}} 1 + \sum_{M\neq M'\in\mathcal{M}} \operatorname{sgn}(M)\operatorname{sgn}(M')\beta(M)\overline{\beta(M')}. \end{aligned}$$

The result now follows from

$$\mathbf{E}(\beta(M)\overline{\beta(M')}) = 0 \qquad \text{for } M\neq M'\in\mathcal{M}. \tag{6.2}$$

If $M\neq M'$ then $M\oplus M'$ (the symmetric difference of M and M') contains at least one cycle C, say. Let $(i_1,j_1)\in M$ be an edge of C. Then we can write $\beta(M)\overline{\beta(M')} = B_{i_1,j_1}\Delta$ where Δ depends only on the values of $B_{i,j}$, $(i,j)\neq(i_1,j_1)$. Then, by the independence of the $B_{i,j}$'s,

$$\mathbf{E}(\beta(M)\overline{\beta(M')}) = \mathbf{E}(B_{i,j})\mathbf{E}(\Delta) = \frac{1}{3}(1+\omega+\omega^2)\mathbf{E}(\Delta) = 0$$

which confirmes (6.2). □

The efficiency of the KKLLL estimator will depend on its variance.

Let M and M' be perfect matchings in G. Denote by $c(M,M')$ the number of connected components (cycles) in $M\oplus M'$. 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 randomly from $\mathcal{M}$. (If G has no perfect matchings then define $\gamma(G)=1$.)

Theorem 6.5.2

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

Proof Let

$$\xi(M_1,M_2,M_3,M_4) = \prod_{i\in\{1,3\}} \operatorname{sgn}(M_i)\beta(M_i) \prod_{i\in\{2,4\}} \operatorname{sgn}(M_i)\overline{\beta(M_i)}.$$

Then

$$Z^2 = \sum_{\mathcal{M}^4} \xi(M_1, M_2, M_3, M_4).$$

If there exists (i_1, j_1) which appears an odd number of times in the product ξ then $\mathbf{E}(\xi) = 0$. Indeed, if it appears 3 times then it occurs at least once as $B_{i,j}$ and at least once as $\overline{B_{i,j}}$ and as $B_{i,j}\overline{B_{i,j}} = 1$ we can reduce to the case where (i_1, j_1) appears exactly once and then $\mathbf{E}(\xi) = 0$ as in the proof of (6.2).

So now assume that every (i, j) appears an even number of times in the product ξ. If there exists (i_1, j_1) which appears twice as $B_{i,j}^2$ or as $\overline{B_{i,j}}^2$ then $\mathbf{E}(\xi) = 0$ as $\mathbf{E}(B_{i,j}^2) = \mathbf{E}(\overline{B_{i,j}}^2) = 0$. (Here we see the advantage of taking ω as a cube root of unity, as opposed to -1 as in [**?**]).

We are left with the case where each $B_{i,j}$ occurs with an accompanying $\overline{B_{i,j}}$. But now we have $M_1 \oplus M_2 = M_3 \oplus M_4$, for if say $(i_1, j_1) \in M_1 \oplus M_2 \setminus M_3 \oplus M_4$ then (i_1, j_1) appears an odd number of times in ξ. Observe that in this case $\prod_{i=1}^4 \mathrm{sgn}(M_i) = 1$ since $\pi_{M_1}\pi_{M_2}^{-1}$ and $\pi_{M_3}\pi_{M_4}^{-1}$ have the same cycle structure (defined by $M_1 \oplus M_2$). Also $B_{i,j}\overline{B_{i,j}} = 1$ and so $\xi = 1$ here.

Now given M_1, M_2 there are $2^{c(M_1,M_2)}$ choices of M_3, M_4 which satisfy $M_1 \oplus M_2 = M_3 \oplus M_4$. Thus

$$\mathbf{E}(Z^2) = \sum_{\mathcal{M}^2} 2^{c(M_1,M_2)}$$

and the result follows from $\mathbf{E}(Z)^2 = |\mathcal{M}|^2$. $\square$

We will now restrict our attention to *dense* matrices. We assume that each row and column of A has at least $(\frac{1}{2} + \alpha)n$ non-zeros for some constant $\alpha > 0$.

Theorem 6.5.3 *Suppose $\alpha > 0$ is a constant, and that bipartite graph G has minimum vertex degree $\delta(G) \geq (\frac{1}{2} + \alpha)n$; then $\gamma(G) \leq O(n^{1+(2\ln 2)/\alpha})$.*

Proof Fix a perfect matching $M_0 \in \mathcal{M}$, and for $M \in \mathcal{M}$ let

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

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 6.5.1 *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 in Section 3.2.1 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,

$$\zeta(M) \geq \sum_{(a,b)} (\tfrac{1}{2} s_1 + s_2) \tag{6.3}$$
$$\geq \tfrac{1}{2} k\alpha n, \tag{6.4}$$

provided $n \geq \alpha^{-1}$. The $\frac{1}{2}$ in inequality (6.3) 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 $\eta(M') \le n$. Hence,

$$\tfrac{1}{2}k\alpha n N_{k,\ell} \le |E_{k,\ell}| \le \tfrac{1}{2}N_{k-2,\ell+1} + nN_{k-1,\ell},$$

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):

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

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

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

Finally,

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

and (b) follows. □

Proof of Theorem 6.5.3
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

$$k\alpha s_{k,\ell} \le \tfrac{1}{2}s_{k-2,\ell+1} + 2s_{k-1,\ell}. \tag{6.5}$$

Let $S_k = \sum_{\ell=0}^{n} s_{k,\ell}$. Then inequality (6.5) implies $k\alpha S_k \le \frac{1}{2}S_{k-2} + 2S_{k-1}$. It follows by an easy induction on k that for $k > k_0 = \lceil 4/\alpha \rceil$,

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

and hence

$$\sum_{k=k_0}^{n} S_k = O(S_{k_0} + S_{k_0-1}). \tag{6.6}$$

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

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

provided

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

Thus, for $k \le k_0$,

$$S_k \le ns_{k,\ell_0} + \sum_{\ell=0}^{\ell_0} s_{k,\ell} \le (n+\ell_0)2^{\ell_0}N. \tag{6.7}$$

Hence, from (6.6) and (6.7), $\Delta/N = \sum_{k=0}^{n} S_k/N = O(n^{1+(2\ln 2)/\alpha})$. □

It follows from Theorem 6.5.3 that $\mathbf{Var}(Z) = O(n^{1+(2\ln 2)/\alpha})$ and so by taking the average of $t = O(\epsilon^{-2}n^{1+(2\ln 2)/\alpha})$ independently generated values of Z we can estimate per A to within $1+\epsilon$, with probability at least $3/4$ as required.

Bibliography

[1] D. Aldous also J. Fill Reversible Markov chains and random walks on graphs Monograph in preparation, available from `http://www.stat.berkeley.edu/users/aldous/book.html`, Department of Statistics, University of California, Berkeley.

[2] T. Lindvall Lectures on the Coupling Method Wiley-IntersciencieNew York 1992

[3] C. H. Papadimitriou Computational Complexity Addison–WesleyReading, MA 1994

[Bro86] A.Z. Broder, How hard is it to marry at random? (On the approximation of the permanent), *Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, ACM Press, 1986, 50–58. Erratum in *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, 1988, p. 551.

[Friez89] A.M. Frieze, *A note on computing random permanents* (unpublished manuscript), 1989.

[FS92] A. Frieze and S. Suen, Counting the number of Hamiltonian cycles in random digraphs, *Random Structures and algorithms*, 3:235–241, 1992.

[HL72] O.J. Heilmann and E.H. Lieb, Theory of monomer-dimer systems, *Communications in Mathematical Physics*, 25:190–232, 1972.

[JMS92] M. Jerrum, B. McKay and A. Sinclair, When is a graphical sequence stable?, *Random Graphs* 2 (A. Frieze and T. Łuczak, eds), Wiley, 1992, 101–115.

[JS89] M.R. Jerrum and A.J. Sinclair, Approximating the permanent, *SIAM Journal on Computing*, 18:1149–1178, 1989.

[JS90a] M.R. Jerrum and A.J. Sinclair, Fast Uniform Generation of Regular Graphs, *Theoretical Computer Science*, 73:91–100, 1990.

[4] Mark Jerrum and Alistair Sinclair, The Markov chain Monte Carlo method: an approach to approximate counting and integration. In *Approximation Algorithms for NP-hard Problems* (Dorit Hochbaum, ed.), PWS, 1996, 482–520.

[KRS93] C. Kenyon, D. Randall and A. Sinclair, Matchings in lattice graphs, *Proceedings of the 25th Annual ACM Symposium on Theory of Computing*, ACM Press, 1993, 738–746. Full version to appear in *Journal of Statistical Physics*, 1996.

[Met53] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, Equation of state calculation by fast computing machines, *Journal of Chemical Physics*, 21:1087–1092, 1953.

[Mih89a] M. Mihail, On coupling and the approximation of the permanent, *Information Processing Letters*, 30:91–95, 1989.

[MW91] M. Mihail and P. Winkler, On the number of Eulerian orientations of a graph, *Proceedings of the 3rd Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM Press, 1992, 138–145.

[Mot89] R. Motwani, Expanding graphs and the average-case analysis of algorithms for matchings and related problems, *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, ACM Press, 1989, 550–561.

[Val79a] L.G. Valiant, The complexity of computing the permanent, *Theoretical Computer Science*, 8:189–201, 1979.