Sampling Contingency Tables

Martin Dyer* Ravi Kannan† John Mount ‡

February 3, 1995

1 Introduction

Given positive integers $r_1, r_2, \ldots r_m$ and $c_1, c_2, \ldots c_n$, let $I(r, c)$ be the set of $m \times n$ arrays with nonnegative integer entries and row sums $r_1, r_2, \ldots r_m$ respectively and column sums $c_1, c_2, \ldots c_n$ respectively. Elements of $I(r, c)$ are called *contingency tables* with these row and column sums.

We consider two related problems on contingency tables. Given $r_1, r_2, \ldots r_m$ and $c_1, c_2, \ldots c_n$,

1) Determine $|I(r, c)|$.

2) Generate randomly an element of $I(r, c)$, each with probability $1/|I(r, c)|$.

The counting problem is of combinatorial interest in many contexts. See, for

*School of Computer Studies, Leeds University, Leeds, England

†School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA 15213, USA. Partially supported by NSF Grant CCR 9208597.

‡School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA 15213, USA. Partially supported by NSF Grant CCR 9208597.

example, the survey by Diaconis and Gangolli [5]. We show that even in the case when m or n is 2, this problem is #P hard.

The random sampling problem is of interest in Statistics. In particular, it comes up centrally in an important new method formulated by Diaconis and Efron [4] for testing independence in two-way tables. (See also Mehta and Patel [21] for a more classical test of independence). We show that this problem can be solved in approximately in polynomial time provided the row and column sums are sufficiently large, in particular, provided $r_i \in \Omega(n^2 m)$ and $c_j \in \Omega(m^2 n)$. This then implies under the same conditions, a polynomial time algorithm to solve the counting problem (1) approximately.

The algorithm we give relies on the random walk approach, and is closely related to those for volume computation and sampling from log-concave distributions [10, 2, 9, 20, 15, 22].

2 Preliminaries and Notation

The number of rows will always be denoted by m and the number of columns by n. We will let N denote $(n-1)(m-1)$. For convenience, we number the coordinates of any vector in $\mathbf{R}^{nm}$ by a pair of integers (i,j) or just ij where i runs from 1 through m and j from 1 through n. Suppose u is an m vector of reals (the row sums) and v is an n vector of reals (the column sums). We define

$$V(u,v) = \{x \in \mathbf{R}^{nm} : \sum_j x_{ij} = u_i \text{ for } i = 1,2,\ldots m;\ \sum_i x_{ij} = v_j \text{ for } j = 1,2,\ldots n\}.$$

$V(u,v)$ can be thought of as the set of real $m \times n$ matrices with row and column sums specified by u and v respectively. Let

$$P(u,v) = V(u,v) \cap \{x : x_{ij} \geq 0 \text{ for } i = 1,2,\ldots m,\ j = 1,2,\ldots n\}.$$

We call such a polytope a "contingency polytope".

Then, $I(u,v)$ is the set of vectors in $P(u,v)$ with all integer coordinates.

Note that the above sets are all trivially empty if $\sum_i u_i \neq \sum_j v_j$. So we will assume throughout that $\sum_i u_i = \sum_j v_j$.

We will need another quantity denoted by $\alpha(u,v)$ defined for u,v satisfying $u_i > 2n \forall i$ and $v_j > 2m \forall j$:

$$\alpha(u,v) = \max_{i,j} \left(\frac{u_i + 2n}{u_i - 2n}, \frac{v_j + 2m}{v_j - 2m} \right)^{(n-1)(m-1)}.$$

We will abbreviate $\alpha(u,v)$ to α when u,v are clear from the context. An easy calculation shows that if u_i is $\Omega(n^2 m)\ \forall i$ and v_j is $\Omega(nm^2)\ \forall j$, then α is $O(1)$.

On input u,v (we assume that $u_i > 2n \forall i$ and $v_j > 2m \forall j$), our algorithm runs for time bounded by $\alpha(u,v)$ times a polynomial in n , m , $\max_{ij}(\log u_i, \log v_j)$, and $(1/\epsilon)$ where ϵ will be an error parameter specifying the desired degree of accuracy. (See section 5, first paragraph for a description of the input/output specifications of the algorithm.)

3 Hardness of counting

We will show that exactly counting contingency tables is hard, even in a very restricted case. We will use the familiar notion of # P-completeness introduced by Valiant [25].

Theorem 1 *The problem of determining the exact number of contingency tables with prescribed row and column sums is is #P-complete, even in the* $2 \times n$ *case.*

Proof It is easy to see that this problem is in #P. We simply guess mn integers x_{ij} in the range $[0, J]$ (where J is the sum of the row sums) and check whether they satisfy the constraints. The number of accepting computations is the number of tables.

For proving hardness, we proceed as follows: Given positive integers $a_1, a_2, \ldots, a_{n-1}, b$, it is shown in [8] that it is #P-hard to compute the $(n-1)$-dimensional volume of the polytope

$$\sum_{j=1}^{n-1} a_j y_j \leq b, \qquad 0 \leq y_j \leq 1 \; (j = 1, 2, \ldots, n-1).$$

It follows that it is #P-hard to compute the $(n-1)$-dimensional volume of the polytope

$$\sum_{j=1}^{n} a_j y_j = b, \qquad 0 \leq y_j \leq 1 \; (j = 1, 2, \ldots, n),$$

where $a_n = b$. Hence, by substituting $x_{1j} = a_j y_j$, $x_{2j} = a_j(1 - y_j)$, that it follows that it is #P-hard to compute the $(n-1)$-dimensional volume of the polytope $P(r, c)$ (with 2 rows and n columns) where

$$r = (b, \sum_{j=1}^{n-1} a_j), \quad c = (a_1, \ldots, a_{n-1}, b).$$

Now the number $\nu(r, c)$ of integer points in $P(r, c)$ is clearly the number of contingency tables with the required row and column sums. But, for integral r, c, $P(r, c)$ is a polytope with integer vertices. (It is the polytope of a $2 \times n$ *transportation problem* [11].) Consider the family of polytopes $P(tr, tc)$ for

$t = 1, 2, \ldots$. It is well known [12, Chapter 12] that the number of integer points in $P(tr, tc)$ will be a polynomial in t, the *Ehrhart polynomial*, of degree $(n-1)$, the dimension of $P(r, c)$. Moreover the coefficient of t^{n-1} will be the $(n-1)$-dimensional volume of $P(r, c)$. It is also straightforward to show that the coefficients of this polynomial are of size polynomial in the length of the description of $P(r, c)$.

Suppose therefore we could count the number of $2 \times n$ contingency tables for arbitrary r, c. Then we could compute $\nu(r, c)$ for arbitrary integral r, c. Hence we could compute $\nu(tr, tc)$ for $t = 1, 2, \ldots, n$ and thus determine the coefficients of the Ehrhart polynomial. But this would allow us to compute the volume of $P(r, c)$, which we have seen is a #P-hard quantity.

$\square$

Remark If both n, m are fixed, we can apply a recent result of Barvinok's [3] (that the number of lattice points in a fixed dimensional polytope can be counted exactly in polynomial time) to compute $|I(u, v)|$. Diaconis and Efron [4] give an explicit formula for $|I(u, v)|$ in case $n = 2$. Their formula has exponentially many terms (as expected from our hardness result). Mann has a similar formula for $n = 3$. Sturmfels demonstrates a structural theorem which allows for an effective counting scheme that runs in polynomial time for n, m fixed.[24] Sturmfels has used the technique to count 4×4 contingency tables in the literature. We will, in later papers, discuss our work counting 4×4 and 5×4 contingency tables.

4 Sampling Contingency tables : Reduction to continuous sampling

This section reduces the problem of sampling from the discrete set of contingency tables to the problem of sampling with near-uniform density from a contingency polytope.

To this end, we first take a natural basis for the lattice of all integer points in $V(u,v)$ and associate with each lattice point a parallelepiped with respect to this basis. Then we produce a convex set $P^{\odot}$ which has two properties : a) each parallelepiped associated with a point in $I(u,v)$ is fully contained in $P^{\odot}$ and b) the volume of $P^{\odot}$ divided by the total volume of the parallelepipeds associated with points in $I(u,v)$ is at most $\alpha(r,c)$. Now the algorithm is simple : pick a random point y from $P^{\odot}$ with near uniform density (this can be done in polynomial time), find the lattice point x in whose parallelepiped y lies and accept x if it is in $I(u,v)$ [acceptance occurs with probability at least $1/\alpha(u,v)$]; otherwise, reject and rerun. The volume of each parallelepiped is the same, so the probability distribution of the x in $I(u,v)$ is near uniform. This general approach has been used also for the simpler context of sampling from the 0-1 solutions to a knapsack problem [13].

[While $P^{\odot}$ is a convex set and by the general methods, we can sample from it in polynomial time with near uniform density, it will turn out that $P^{\odot}$ is also isomorphic to a contingency polytope. In the next section, we show how to exploit this to get a better polynomial time algorithm than the general one.]

We will have to build up some geometric facts before producing the $P^{\odot}$.

Let U be the lattice:

$$\{x \in \mathbf{R}^{nm} : \sum_j x_{ij} = 0 \text{ for } i = 1, 2, \ldots m;\ \sum_i x_{ij} = 0 \text{ for } j = 1, 2, \ldots n;\ x_{ij} \in \mathbf{Z}\}.$$

For $1 \leq i \leq m - 1$ and $1 \leq j \leq n - 1$, let $b(ij)$ be the vector in $\mathbf{R}^{nm}$ given by $b(ij)_{ij} = 1,\ b(ij)_{i+1,j} = -1,\ b(ij)_{i,j+1} = -1,\ b(ij)_{i+1,j+1} = 1$ and $b(ij)_{kl} = 0$ for kl other than the 4 above.

Any vector x in $V(0, 0)$ can be expressed as a linear combination of the $b(ij)$'s as follows (the reader may check this by direct calculation)

$$x = \sum_{k=1}^{m-1} \sum_{l=1}^{n-1} \left(\sum_{i=1}^{k} \sum_{j=1}^{l} x_{ij} \right) b(kl). \tag{1}$$

It is also easy to check that the $b(ij)$ are all linearly independent. This implies that the subspace $V(0, 0)$ has dimension $(n - 1)(m - 1)$ and so does the affine set $V(u, v)$ which is just a translate of $V(0, 0)$. Also, we see that if x is an integer vector, then the above linear combination has integer coeffients; so, the $b(ij)$ form a basis of the lattice U.

It is easy to see that if u, v are positive vectors, then the dimension of $P(u, v)$ is also N. To see this, it suffices to come up with an $x \in \mathbf{R}^{nm}$ with row and column sums given by u, v and with each entry x_{ij} strictly positive because then we can add any small real multiples of the $b(ij)$ to x and still remain in $P(u, v)$. Such an x is easy to obtain : for example, we can choose $x_{11}, x_{12}, \ldots x_{1n}$ to satisfy $0 < x_{1j} < min(u_1, v_j)$ and summing to u_1. Then we subtract this amount from the column sums and repeat the process on the second row etc.

We will denote by Vol($P(u, v)$) the (N dimensional) volume of $P(u, v)$.

Obviously, if either u or v has a non-integer coordinate, then $I(u, v)$ is empty.

Lemma 1 *: If p, q are m vectors of positive reals and s, t are n vectors of positive reals with $q \geq p$ and $t \geq s$ (componentwise), then*

$$\textit{Vol}(P(q,t)) \geq \textit{Vol}(P(p,s)).$$

Proof By induction on the number of coordinates of $(q\ \ t)$ that are strictly greater than the corresponding coordinates of $(p\ \ s)$. After changing row and column numbers if necessary, assume without loss of generality that $q_1 - p_1$ is the least among all POSITIVE components of $(q\ \ t) - (p\ \ s)$. After permuting columns if necessary, we may also assume that we have $t_1 - s_1 > 0$. Let $P_1 = P(p,s)$. Let P_2 be defined by

$$P_2 = \{x \in \mathbf{R}^{nm} : x_{11} \geq -(q_1 - p_1); x_{ij} \geq 0 \text{ for } (ij) \neq (11)\} \cap V(p,s).$$

Let p' be an $m-$vector defined by $p'_1 = q_1; p'_i = p_i$ for $i = 2, 3, \ldots m$. Let s' be defined by $s'_1 = s_1 + q_1 - p_1; s'_j = s_j$ for $j = 2, 3, \ldots n$. Denote by P_3 the set $P(p', s')$. Then P_2 and P_3 have the same volume as seen by the isomorphism $x_{11} \to x_{11} + (q_1 - p_1)$. Clearly, P_2 contains P_1. Also the row sums and column sums in P_3 are no greater than the corresponding row and column sums in $(q\ \ t)$. Applying the inductive assumption to P_3 and $P(q,t)$, we have the lemma. $\square$

Let u, v be fixed and consider $V(u,v)$. Let

$$U(u,v) = V(u,v) \cap \mathbf{Z}^{nm}.$$

We may partition $V(u,v)$ into fundamental parallelepipeds, one corresponding to each point of $U(u,v)$. Namely,

$$V(u,v) = \cup_{x \in U(u,v)} \{x + \sum_{i,j} \lambda_{ij} b(ij) : 0 \leq \lambda_{ij} < 1 \text{ for } i = 1, 2, \ldots m;\ j = 1, 2, \ldots n\}.$$

We call $\{x + \sum_{i,j} \lambda_{ij} b(ij) : 0 \leq \lambda_{ij} < 1$ for $i = 1, 2, \ldots m;\ j = 1, 2, \ldots n\}$, "the parallelepiped" associated with x and denote it by $F(x)$.

Claim : For any $x \in I(u, v)$, and for any y belonging to $F(x)$ we have $x_{ij} - 2 < y_{ij} < x_{ij} + 2$.

Proof : This follows from the fact that at most four of the $b(ij)$'s have a nonzero entry in each position and two of them have a +1, the other two a -1. $\square$

Lemma 2 *: For nonnegative vectors $u \in \mathbf{R}^m$ and $v \in \mathbf{R}^n$ and any $l \in \mathbf{R}^{nm}$, let*

$$P(u, v, l) = V(u, v) \cap \{x \in \mathbf{R}^{nm} : x_{ij} \geq l_{ij}\}.$$

Then,

(i) for any x in $I(u, v)$, we have $F(x) \subseteq P(u, v, -\underline{2})$ where $\underline{2}$ is the nm-vector of all 2 's.

(ii) $P(u, v, +\underline{2}) \subseteq \cup_{x \in I(u,v)} F(x)$. [$P(u, v, -\underline{2})$ will play the role of $P^{\odot}$ in the brief description of the algorithm given earlier.]

Proof : (i) Since x is in $I(u, v)$, we have $x_{ij} \geq 0$ and so for any $y \in F(x)$, we have $y_{ij} \geq -2$ proving (i). For (ii), observe that for any $y \in V(u, v)$, there is a unique $x \in U(u, v)$ such that $y \in F(x)$. If $y_{ij} \geq 2$, then we must have $x_{ij} \geq 0$, so the corresponding x belongs to $I(u, v)$ as claimed.

Lemma 3 *: For $u \in \mathbf{R}^m$ and $v \in \mathbf{R}^n$ with $u_i > 2n$ for all i and $v_j > 2m$ for all j, we have*

$$\frac{\textit{Vol}(\cup_{x \in I(u,v)} F(x))}{\textit{Vol}(P(u, v, -\underline{2}))} \geq \frac{1}{\alpha(u, v)}.$$

Proof : By (ii) of the lemma above, Vol($\cup_{x \in I(u,v)} F(x)$) is at least Vol($P(u, v, +\underline{2})$). But $P(u, v, +\underline{2})$ is isomorphic to $P(u - 2n\underline{1}, v - 2m\underline{1})$ where $\underline{1}$'s are vectors of all 1 's of suitable dimensions as seen from the substitution $x'_{ij} = x_{ij} - 2$. Similarly, we have that $P(u, v, -\underline{2})$ is isomorphic to $P(u + 2n\underline{1}, v + 2m\underline{1})$. Let $\rho = \alpha^{1/N}$. Consider the set $\rho P(u - 2n\underline{1}, v - 2m\underline{1}) = \{\rho x : x \in P(u - 2n\underline{1}, v - 2m\underline{1})\}$. This is precisely the set $P(\rho(u - 2n\underline{1}), \rho(v - 2m\underline{1}))$. By the definition of ρ, we have $\rho(u_i - 2n) \geq u_i + 2n \forall i$ and $\rho(v_j - 2m) \geq v_j + 2m \forall j$. So by lemma 1, the volume of $\rho P(u - 2n\underline{1}, v - 2m\underline{1})$ is at least the volume of $P(u + 2n\underline{1}, v + 2m\underline{1})$. But $\rho P(u - 2n\underline{1}, v - 2m\underline{1})$ is a dilation of the N dimensional object $P(u - 2n\underline{1}, v - 2m\underline{1})$ by a factor of ρ and thus has volume precisely equal to $\rho^N = \alpha$ times the volume of $P(u - 2n\underline{1}, v - 2m\underline{1})$ completing the proof.

□

Since $P(u, v, -2)$ is isomorphic to $P(u + 2n\underline{1}, v + 2m\underline{1})$, the essential problem we have is one of sampling from a contingency polytope with uniform density.

Proposition 1 *For every $x \in U(r, c)$, we have $vol_N(F(x)) = 1$.*

Proof : Note that if $z = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \lambda_{ij} b(ij)$, then we have $z_{ij} = \lambda_{ij} - \lambda_{i-1,j} - \lambda_{i,j-1} + \lambda_{i-1,j-1}$ (where if one of the subscripts is 0, we take the λ to be zero too. Thus for each i, $1 \leq i \leq m - 1$ and each j, $1 \leq j \leq n - 1$, the "height" of the parallelepiped $F(x)$ in the direction of z_{ij} perpendicular to the "previous" ($(a, b) < (i, j)$ if $a \leq i$ and $b \leq j$ and ($a < i$ or $b < j$)) coordinates is 1 proving our proposition.

□

5 The Algorithm

We now describe the algorithm which given row sums $u_1, u_2, \dots u_m$ and column sums $v_1, v_2, \dots v_n$, and an error parameter $\epsilon > 0$, draws a multiset W of samples from $I(u, v)$ (the set of integer contingency tables with the given row and column sums) with the following property :

if $g : I(u, v) \to \mathbf{R}$ is any function, then the true mean $\bar{g}$ of g and the sample mean $\tilde{g}$ (defined below) satisfy

$$E((|\tilde{g} - \bar{g}|)^2) \leq \epsilon g_0^2$$

where

$$\bar{g} = \frac{\sum_{w \in I(u,v)} g(w)}{|I(u, v)|} \quad \tilde{g} = \frac{\sum_{w \in W} g(w)}{|W|} \quad g_0 = \max_{w, x \in I(u,v)} g(w) - g(x).$$

As stated in the last section, we will first draw samples from $P(u + 2n\underline{1}, v + 2m\underline{1})$. We let $r = u + 2n\underline{1}$ and $c = v + 2m\underline{1}$.

We assume that $u_i > 2n \forall i$ and $v_j > 2m \forall j$ throughout this section.

The function g often takes on values 0 or 1. This is so in the motivating application described by Diaconis and Efron [4] where one wants to estimate the proportion of contingency tables that have (the so-called) χ^2 measure of deviation from the independent greater than the observed table. So, in this case, g_0 will be 1 for a table if its χ^2 measure is greater than that of the observed, otherwise, it will be zero. The evaluation of g for a particular table is in general simple as it is in this case.

At the end of the section, we describe how to use the sampling algorithm to estimate $|I(u, v)|$.

For reasons that will become clear later, we rearrange the rows and columns so that

$$r_1, r_2, \ldots r_{m-1} \leq r_m \quad c_1, c_2, \ldots c_{n-1} \leq c_n. \tag{2}$$

We assume the above holds from now on. An $m \times n$ table x in $P(r, c)$ is obviously specified completely by its entries in all but the last row and column. From this, it is easy to see that $P(r, c)$ is isomorphic to the polytope $Q(r, c)$ in $\mathbf{R}^N$ (recall the notation $N = (m-1)(n-1)$) which is defined as the set of x satisfying the following inequalities :

$$\sum_{j=1}^{n-1} x_{ij} \leq r_i \forall i \quad \sum_{i=1}^{m-1} x_{ij} \leq c_j \forall j \quad \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} x_{ij} \geq \sum_{i=1}^{m-1} r_i - c_n \quad x_{ij} \geq 0.$$

In the above, as well as in the rest of the section, i will run through $1, 2, \ldots m-1$ and j will run through $1, 2, \ldots n-1$ unless otherwise specified.

The algorithm will actually pick points in $Q(r, c)$ with near uniform density. To do so, the algorithm first scales each coordinate so that $Q(r, c)$ becomes "well-rounded". Let

$$\rho_{ij} = \text{ Min } (\frac{r_i}{n-1}, \frac{c_j}{m-1}) \text{ for } i = 1, 2, \ldots m-1 \text{ and } j = 1, 2, \ldots n-1. \tag{3}$$

The scaling is given by

$$y_{ij} = \frac{N}{\rho_{ij}} x_{ij} \text{ for } i = 1, 2, \ldots m-1 \text{ and } j = 1, 2, \ldots n-1. \tag{4}$$

This transformation transforms the polytope $Q(r, c)$ to a polytope $QQ = QQ(r, c)$ "in y space". Note that $QQ(r, c)$ is the set of y that satisfy the following constraints :

$$\frac{1}{N} \sum_j \rho_{ij} y_{ij} \leq r_i \forall i \qquad \frac{1}{N} \sum_i \rho_{ij} y_{ij} \leq c_j \forall j \tag{5}$$

$$\frac{1}{N}\sum_{i,j}\rho_{ij}y_{ij} \geq \sum_i r_i - c_n \qquad y_{ij} \geq 0.$$

Instead of working in $QQ(r,c)$, the algorithm will work with a larger convex set. To describe the set, first consider the convex set $P'(r,c)$ obtained from $P(r,c)$ by discarding the constraint $x_{mn} \geq 0$. The corresponding set $Q'(r,c)$ in N space is given by upper bounds on the row and column sums and nonnegativity constraints. The scaling above transforms $Q'(r,c)$ to the following set $QQ'(r,c)$:

$$QQ'(r,c) = \{y : \frac{1}{N}\sum_j \rho_{ij}y_{ij} \leq r_i \forall i \quad \frac{1}{N}\sum_i \rho_{ij}y_{ij} \leq c_j \forall j \quad y_{ij} \geq 0\} \tag{6}$$

We will impose the following log-concave function $F : P'(r,c) \to \mathbf{R}_+$:

$$F(x) = \text{ Min } (1, e^{Mx_{mn}}) \qquad \text{where } M = 2(\frac{n-1}{r_m} + \frac{m-1}{c_n}) \tag{7}$$

which is 1 on $P(r,c)$ and falls off exponentially outside. The corresponding function on QQ' is given by :

$$G(y) = \text{ Min } \left(1, e^{M\left(\frac{1}{N}\sum_{i,j}\rho_{ij}y_{ij} - \sum_i r_i + c_n\right)}\right). \tag{8}$$

We call a cube of the form $\{y : 0.4s_{ij} \leq y_{ij} < 0.4(s_{ij}+1)\}$ in y space where s_{ij} are all integers a "lattice cube" (it is a cube of side 0.4); its center p where $p_{ij} = 0.4s_{ij} + 0.2$ will be called a lattice cube center (lcc). Let L be the set of lattice cubes that intersect $QQ'(r,c)$. We will interchangeably use L to denote the set of lcc's of such cubes. (Note that the lcc itself may not be in QQ'.) If y is an lcc, we denote by $C(y)$ the lattice cube of which it is the center. Note that for a particular lcc y, it is easy to check whether it is in L - we just round down all coordinates (to an integer multiple of 0.4) and check if the point so obtained is in QQ'. In our algorithm, each step will only modify one coordinate of y; in this case, by keeping running row and column sums, we can check in $O(1)$ time whether the new y is in L.

A simple calculation shows :

Proposition 2 *: For any lcc y and any $z \in C(y)$, we have $e^{-0.8}G(y) \leq G(z) \leq e^{0.8}G(y)$.*

Now we are ready to present the algorithm. There are two steps of the algorithm of which the first is the more important. Some details of execution of the algorithm are given after the algorithm. In the second step, we may reject in one of three places. If we do so, then, this trial does not produce a final sample.

The Algorithm

The first paragraph of this section describes the input / output behavior of the algorithm. See Remark following the Theorem for the choice of T_0, T_1.

I. Let $T_0, T_1 \geq 1$. Their values are to determined by the accuracy needed.

Run the random walk $Y^{(1)}, Y^{(2)}, \ldots Y^{(T_0+T_1)}$ with L as state space described below for $T_0 + T_1$ steps starting at any state of L.

II. for $i = T_0 + 1$ to $i = T_0 + T_1$ do the following :

a. Pick $Z^{(i)}$ with uniform density from $C(Y^{(i)})$. Reject with probability $1 - \frac{G(Z^{(i)})}{e^{\cdot 8} G(Y^{(i)})}$. [Reject means end the execution for this i.]

b. Reject if $Z^{(i)}$ is not in $QQ(r, c)$.

c. Let $p^{(i)}$ be the point in $P(r, c)$ corresponding to $Z^{(i)}$. If $p^{(i)}$ is in a $F(w)$ for some $w \in I(u, v)$, then add $w^{(i)} = w$ to the multiset W of samples to be output; otherwise reject.

$\square$

Random walk for step I

Step I is a random walk on L. Two lcc 's in L will be called "adjacent" if they differ in one coordinate by 0.4 and are equal on all the other coordinates. The transition probabilities for the random walk for step I are given by :

$$\text{Pr}(y \to y') = \frac{1}{2N} \text{ Min } (1, \frac{G(y')}{G(y)}) \qquad \text{for } y, y' \text{ adjacent lcc 's in } L$$

$$\text{Pr}(y \to y') = 0 \text{ for } y, y' \text{ not adjacent and}$$

$$\Pr(y \to y) = 1 - \sum_{y' \in L \setminus \{y\}} \Pr(y \to y').$$

This is an irreducible time-reversible Metropolis Markov Chain and from the identity

$$G(y) \Pr(y \to y') = G(y') \Pr(y' \to y),$$

it follows that the steady state probabilities, π, exist and are proportional to $G(\cdot)$.

An execution of the random walk produces $Y^{(i)}, i = 1, 2, \ldots$ where $Y^{(i+1)}$ is chosen from $Y^{(i)}$ according to the transition probababilities.

In the next section, we will use the results of Frieze, Kannan, Polson [[15]] to show that the second largest absolute value of an eigenvalue of the Markov Chain (we denote this quantity by θ) satisfies (see Theorem 3):

$$(1-\theta)^{-1} \leq N^4(n+m-2)(35+o(1)),$$

where the $o(1)$ term goes to zero as $n+m, N \to \infty$. As pointed out there, for $N \geq 10$ and $n+m \geq 11$, we have $(1-\theta)^{-1} \leq 36N^4(n+m-2)$. The quantity $1-\theta$ is called the "spectral gap".

The following theorem is proved using a result of Aldous [[1]]. It involves a quantity A_1 which is the limiting probability of acceptance in the above algorithm. A sequence of lemmas (4-7) is then used to bound A_1 from below. This lower bound may be just plugged into the Theorem. See Remark below for a more practical way to estimate A_1.

Theorem 2 *The multiset of samples W produced by the algorithm satisfies :*

if $g : I(u,v) \to \mathbf{R}$ is any function, with $g_0 = \max_{w,x \in I(u,v)} g(w) - g(x)$, and the *mean* $\bar{g} = \frac{\sum_{w\in I(u,v)} g(w)}{|I(u,v)|}$*, we have*

$$E((\frac{1}{|W|}\sum_{w\in W} g(w)-\bar{g})^2) \leq \frac{1}{T_1A_1}\frac{(11.5-1806\alpha(u,v)\log\theta)}{-\log\theta}\left(1+\theta^{T_0}e^{5N}N^{2N}e^4\right)g_0^2.$$

Further,

$$A_1 \geq e^{-4.4}/\alpha(u,v).$$

Remark To use the theorem, we must choose T_0, T_1 to be large enough. This can be done in a fairly standard manner : suppose we want

$$E((\frac{1}{|W|}\sum_{w\in W} g(w)-\bar{g})^2)) \leq \epsilon.$$

Then, note that using Theorem 3, we can choose

$$T_0 = (5N + 2N\log(N) + \frac{4}{N})(35 + o(1))N^4(n+m-2)$$

$$T_1 = \frac{(840 + o(1))}{A_1}g_0^2N^4(n+m-2)/\epsilon.$$

We may estimate A_1 from the run of the algorithm which potentially fares much better than the lower bound we have given.

Proof of the Theorem

Let $S = \sum_{w\in W}(g(w)-\bar{g})$. Let $f : L \times I(u,v) \to [0\ \ 1]$ be such that $f(y,w)$ is the probability that given $Y^{(i)} = y$ in step I of the algorithm, we pick $w^{(i)} = w$. Letting τ denote the linear transformation such that for each $x \in P(r,c)$, $\tau(x) = z$ gives us the corresponding point in QQ, (the exact description of τ is not needed here), we have

$$f(y,w) = (2.5)^N \int_{z\in\tau(F(w))\cap C(y)} \frac{G(z)}{e^{.8}G(y)}dz.$$

Let $h(y) = \sum_{w \in I(u,v)} f(y,w)(g(w) - \bar{g})$. [This is the expected value of $g(w^{(i)}) - \bar{g}$ given $Y^{(i)} = y$ where we count a rejected $Y^{(i)}$ as giving us the value 0.]

Let

$$S' = \sum_{i=T_0+1}^{T_0+T_1} h(Y^{(i)}).$$

[Note that the expectation of S given the $Y^{(i)}$ is precisely S'.] Consider the expectation of h with respect to π:

$$\sum_y \pi(y)h(y) = \sum_y \sum_w \pi(y) f(y,w)(g(w) - \bar{g}) = \sum_y \frac{G(y)}{\sum_L G} \sum_w \int_{z \in \tau(F(w)) \cap C(y)} \frac{G(z)}{e^{\cdot 8} G(y)} (g(w) - \bar{g}) dz$$

$$= \frac{1}{e^{\cdot 8} \sum_L G} \sum_w (g(w) - \bar{g}) \int_{z \in \tau(F(w))} G(z) dz = \sum_w \frac{\text{vol}(\tau(F(w)))}{e^{\cdot 8} \sum_L G} (g(w) - \bar{g}) = 0$$

because $G(z)$ is 1 on all of QQ which contains $\tau(F(w))$; also the volume of $\tau(F(w))$ is the same for all $w \in I$.

If we had started the chain in the steady state, the expected value of S' would be $T_1 \sum_y \pi(y)h(y) = 0$. Thus as T_1 tends to infinity, the limit of S'/T_1 is 0. Let $A(y)$ be the probability of accepting y. Let $A_1 = \sum_y \pi(y)A(y)$. We also need the variance of h (wrt π) denoted $h_2 = \sum_y \pi(y)(h(y))^2$.

$$h_2 = \sum_y \pi(y)(h(y))^2 \le g_0^2 \sum_y \pi(y)(A(y))^2 \le g_0^2 A_1$$

It follows from Aldous's Proposition 4.2 (using the fact that his τ_e/T_1 will be less than 1 here which implies that his $\alpha(\tau_e/T_1)$ is at most $2(\tau_e/T_1)(1 + e^{-1})$)

$$E((S')^2) \le T_1 \left(1 + \frac{\theta^{T_0}}{\pi_*}\right) \frac{2(1 + e^{-1})}{(-\log\theta)} g_0^2 A_1,$$

where, π_* is the minimum of all $\pi(y), y \in L$. [We note that Aldous lets T_0 actually be a random variable to avoid the dependence on negative eigenvalues of the chain.

In our case, we will be able to get an upper bound on θ, the second largest *absolute value* of an eigenvalue, so this is not necessary. A simple modification of Aldous's argument which we do not give here implies the above inequality even though our T_0 is a deterministic quantity.]

It also follows by similar arguments that

$$E((|W| - T_1 A_1)^2) \leq T_1 \left(1 + \frac{\theta^{T_0}}{\pi_*}\right) \frac{2(1+e^{-1})}{(-\log\theta)} A_1. \tag{9}$$

Consider now S. Recall that

$$E(S|Y^{(i)}, i = 1, 2, \dots T_0 + T_1) = S'.$$

Given $Y^{(i)}, i = 1, 2, \dots T_0 + T_1$, the process of producing each $w^{(i)}$ is independent. For some Y $E((S - S')^2) \leq E((S - S')^2 | Y^{(i)}, i = 1, 2, \dots T_0 + T_1)$. So, we have

$$\text{Var}(S|Y^{(i)}, i = 1, 2, \dots T_0 + T_1) = \sum_i \text{Var}(g(w^{(i)}) - \bar{g}|Y^{(i)}) \leq E((g(w^{(i)}) - \bar{g}))^2 \leq T_1 g_o^2.$$

So we have, (using the inequality $(A + B)^2 \leq 11A^2 + 1.1B^2$)

$$\begin{aligned} E(S^2) = E(((S - S') + S')^2) &\leq 11E((S - S')^2) + 1.1E((S')^2) \\ &\leq 1.1(1 - 300\alpha(u, v)\log\theta) T_1 A_1 \left(1 + \frac{\theta^{T_0}}{\pi_*}\right) \frac{2(1+e^{-1})}{(-\log\theta)} g_0^2, \end{aligned} \tag{10}$$

where we have used a lower bound on A_1 which we derive now. First note that

$$A_1 = \frac{\text{vol}(\tau(F(w)))}{e^{.8}\sum_L G} |I(u, v)|.$$

We have $\sum_L G \leq e^{.2} \int_T G$, where T is the union of cubes in L.

In a sequence of lemmas below, we show that $\int_T G \leq e^{4.4}\text{vol}(QQ)$. So, using Lemma 3,

$$A_1 \geq e^{-4.4} \frac{\text{vol}(\cup_{I(u,v)} \tau(F(w)))}{\text{vol}(QQ)} = e^{-4.4} \frac{\text{vol}(\cup_{I(u,v)} F(w))}{\text{vol}(P(r, c))} \geq \frac{e^{-4.4}}{\alpha(u, v)}.$$

We will now use equations 9 and 10 to argue the conclusion of the Theorem. To this end, let

$$E\left(\left(\frac{1}{|W|}\sum_{w\in W} g(w)-\bar{g}\right)^2 \ \mid\ |W|=s\right)=\gamma(s)$$

and let $\Pr(|W|=s)=\nu(s)$. Then equation 9 gives us

$$\sum_s \nu(s)(s-T_1A_1)^2 \le T_1A_1\left(1+\frac{\theta^{T_0}}{\pi_*}\right)\frac{2(1+e^{-1})}{(-\log\theta)}$$

which implies

$$\sum_s \nu(s)\gamma(s)(s-T_1A_1)^2 \le T_1A_1\left(1+\frac{\theta^{T_0}}{\pi_*}\right)\frac{2(1+e^{-1})}{(-\log\theta)}g_0^2$$

Now 10 gives

$$\sum_s \nu(s)\gamma(s)s^2 \le 1.1(1-300\alpha\log\theta)T_1A_1\left(1+\frac{\theta^{T_0}}{\pi_*}\right)\frac{2(1+e^{-1})}{(-\log\theta)}g_0^2.$$

Using the inequality $2(s^2+(s-T_1A_1)^2)\ge T_1^2A_1^2$, we get

$$\sum_s \nu(s)\gamma(s) \le (11.5-1806\alpha\log\theta)\frac{1}{T_1A_1}\left(1+\frac{\theta^{T_0}}{\pi_*}\right)\frac{1}{-\log\theta}g_0^2,$$

which gives us the theorem using proposition 3.

□

Lemma 4 *For any real number t (positive, negative or zero), let*

$$K(t)=P'(r,c)\cap\{x: x_{mn}=t\}\quad \textit{and} \qquad v(t)=\mathit{Vol}_{N-1}(K(t)).$$

For $t_1\le t_2<$ Min (r_m,c_n), we have

$$\frac{v(t_1)}{v(t_2)}\le\left(\frac{r_m-t_1}{r_m-t_2}\right)^{n-2}\left(\frac{c_n-t_1}{c_n-t_2}\right)^{m-2}.$$

Proof : For a real $n+m-4$ vector $\lambda = \lambda_{1n}, \lambda_{2n} \ldots \lambda_{m-2,n}, \quad \lambda_{m1}, \lambda_{m2}, \ldots \lambda_{m,n-2}$ (we assume our λ vector is indexed as above), define $K(t|\lambda)$ as the set of tables x satisfying

$$x_{mj} = \lambda_{mj} \text{ for } j = 1, 2, \ldots n-2; x_{m,n-1} = r_m - t - \sum_{j=1}^{n-2} \lambda_{mj};$$

$$x_{in} = \lambda_{in} \text{ for } i = 1, 2, \ldots m-2; x_{m-1,n} = c_n - t - \sum_{i=1}^{m-2} \lambda_{in}.$$

$K(t|\lambda)$ is the set of tables in $K(t)$ with their last row and column dictated by λ. Let us denote $r_m - t - \sum_{j=1}^{n-2} \lambda_{mj}$ by $\lambda_{m,n-1}(t)$ and $c_n - t - \sum_{i=1}^{m-2} \lambda_{in}$ by $\lambda_{n,m-1}(t)$. Define $\Lambda(t)$ to be the set of nonnegative vectors λ satisfying :

$$\sum_{j=1}^{n-2} \lambda_{mj} + t \le r_m; \sum_{i=1}^{m-2} \lambda_{in} + t \le c_n;$$

$$\lambda_{mj} \le c_j \text{ for } j = 1, 2, \ldots n-2; \lambda_{m,n-1}(t) \le c_{n-1}; \lambda_{in} \le r_i \text{ for } i = 1, 2, \ldots m-2; \lambda_{m-1,n} \le r_{m-1}.$$

Then, we have that $K(t|\lambda)$ is nonempty iff λ belongs to $\Lambda(t)$. In general, $K(t|\lambda)$ is an $(m-2)(n-2)$ dimensional set and by the volume of $K(t|\lambda)$, we mean its $(m-2)(n-2)$ dimensional volume. We have

$$\text{Vol}(K(t|\lambda)) = \phi \text{Vol}(P(r(t,\lambda), c(t,\lambda))), \text{ where,}$$

$$r(t,\lambda)_i = r_i - \lambda_{in} \text{ for } i = 1, 2, \ldots m-2; \quad c(t,\lambda) = c_j - \lambda_{mj} \text{ for } j = 1, 2, \ldots n-2;$$

$$r(t,\lambda)_{m-1} = r_{m-1} - \lambda_{m-1,n}(t); \quad c(t,\lambda)_{n-1} = c_{n-1} - \lambda_{m,n-1}(t).$$

Consider the linear transformation τ given by

$$(\tau(\lambda))_{mj} = \frac{r_m - t_2}{r_m - t_1} \lambda_{mj} \text{ for } j = 1, 2 \ldots n-2;$$

$$(\tau(\lambda))_{in} = \frac{c_n - t_2}{c_n - t_1} \lambda_{in} \text{ for } i = 1, 2 \ldots m-2.$$

It is easy to see that τ is a 1-1 map of $\Lambda(t_1)$ into $\Lambda(t_2)$.

$$\text{Vol}(K(t_1)) = \int_{\lambda\in\Lambda(t_1)} \text{Vol}(K(t_1|\lambda))d\lambda = \int_{\alpha\in\Lambda(t_2)} \text{Vol}(K(t_1|\tau^{-1}\alpha))|\det(\tau^{-1})|d\alpha$$

$$= \left(\frac{r_m - t_1}{r_m - t_2}\right)^{n-2} \left(\frac{c_n - t_1}{c_n - t_2}\right)^{m-2} \int_\alpha \phi\text{Vol}(P(r(t_1,\tau^{-1}\alpha), c(t_1,\tau^{-1}\alpha)))d\alpha.$$

It is easy to check that $r(t_1,\tau^{-1}\alpha) \leq r(t_2,\alpha)$ and $c(t_1,\tau^{-1}\alpha) \leq c(t_2,\alpha)$. This implies that the integrand in the last integral is bounded above by $\phi\text{Vol}(P(r(t_2,\alpha), c(t_2,\alpha)))$. This function of course integrates to $\text{Vol}(K(t_2))$ completing the proof.

$\square$

Lemma 5 *For any $t_1 < 0$, we have $v(t_1)$ (defined in lemma 4) satisfies*

$$v(t_1) \leq \frac{e^2}{\rho_{mn}} \left(\frac{r_m - t_1}{r_m}\right)^{n-2} \left(\frac{c_n - t_1}{c_n}\right)^{m-2} \mathit{Vol}_N(QQ(r,c))$$

where $\rho_{mn} =$ Min $(\frac{r_m}{n-1}, \frac{c_m}{m-1})$.

Proof For t_2 in the range $[0\ \ \rho_{mn}]$, we have

$$\text{Vol}(K(t_2)) \geq \left(\frac{r_m - t_2}{r_m - t_1}\right)^{n-2} \left(\frac{c_n - t_2}{c_n - t_1}\right)^{m-2} \text{Vol}(K(t_1))$$

$$\geq \left(\frac{r_m}{r_m - t_1}\right)^{n-2} \left(\frac{c_n}{c_n - t_1}\right)^{m-2} e^{-2}\text{vol}(K(t_1)).$$

Integrating this over this range of t_2, we get the lemma.

Lemma 6 *:*

$$\int_{QQ'(r,c)} G(y)dy \leq e^2 \int_{QQ(r,c)} G(y)dy \leq e^2 \mathit{vol}(QQ).$$

Proof : From the last lemma, we have for any $t_1 < 0$,

$$v(t_1) \leq e^{\frac{-(n-2)t_1}{r_m} + \frac{-(m-2)t_1}{c_n}} \frac{e^2}{\rho_{mn}} \text{Vol}(QQ).$$

Now the lemma follows by integration.

□

Lemma 7 *Let T be the union of $C(y)$ over all lcc 's y. Then we have*

$$\int_T G(y)dy \leq e^{3.6} \int_{QQ(r,c)} G(y)dy.$$

Proof : We have :

$$T \subseteq \{y : \frac{1}{N}\sum_j \rho_{ij}y_{ij} \leq r_i + \frac{.4}{N}\sum_j \rho_{ij} \forall i \quad \frac{1}{N}\sum_i \rho_{ij}y_{ij} \leq c_j + \frac{.4}{N}\sum_i \rho_{ij} \forall j \quad y_{ij} \geq 0\}. \tag{11}$$

This implies that $T \subseteq (1 + \frac{.4}{N})QQ'$. Note also that for any y in QQ', we have $G((1 + \frac{.4}{N})y) \leq G(y)e^{0.4M\sum_{ij} \min(\sum_i r_i, \sum_j c_j)/(N^2)} \leq G(y)e^{1.6}$. Thus we have the lemma using the last lemma.

□

Proposition 3

$$\pi_*^{-1} \leq e^{5N} N^{2N} e^4.$$

Proof The number of states is at most N^{2N}. It is easy to check that the ratio of the maximum value of G to the minimum over lcc 's is at most e^{5N+4} completing the proof.

□

Algorithm to estimate the number of contingency tables

Using the sampling algorithm, we will be able to estimate the number of contingency tables in time polynomial in the data and $\alpha(u,v)$. We will only sketch the method here.

We first estimate by sampling from $P(r,c)$ the following ratio :

$$\frac{\text{vol}(P(r,c))}{\text{vol}(\cup_{w\in I(u,v)}F(w))}.$$

Next, we define a sequence of contingency polytopes $P_1, P_2, \ldots$, each obtained from the previous one by increasing r_m and c_j $(j = 1\cdots n)$ by $\left\lfloor \min\left(\frac{1}{m}c_n, \frac{1}{n}r_m\right)\right\rfloor$ until we have $c_j \geq \sum_{i=1}^{m-1} r_i$ for $j = 1\cdots n$. Then $|I(r,c)| = \prod_{i=1}^{m-1}\binom{r_i+n-1}{n-1}$ and the volume of $P(r,c)$ is read off as the coefficient of β^N in the easily calculated polynomial $|I(\beta\times r, \beta\times c)|$ $(= \prod_{i=1}^{m-1}\left(r_i^{n-1}/((n-1)!)\right))$.

6 Bound on the spectral gap

We refer the reader to Diaconis and Strook [6] or [15] for a discussion of the eigenvalues of a reversible Markov Chain. The second largest absolute value of such a chain gives us a bound on the "time to converge" to the steady state as described in these references. Here, we use Theorem 2 of [15] to bound the second largest absolute value of our chain. Note that though their theorem does not explicitly state so, the λ_2 in that Theorem is the second largest absolute value of any eigenvalue of the chain. This section is a technical evaluation of the various quantities needed to plug into the expression for λ_2 in [15]; we do not redefine these quantities here.

We first calculate the diameter of T (the largest Euclidean distance between two points in T). To do so, we let

$$I_1 = \{(ij) : \rho_{ij} = \frac{r_i}{n-1}\} \qquad I_2 = \{(ij) : \rho_{ij} < \frac{r_i}{n-1}\}.$$

Then, for any $y \in T$, (using the fact that $(y - .4\underline{1}) \in QQ'$,)

$$\forall i, \sum_{\{j:(ij)\in I_1\}} y_{ij} \leq N(n-1)(1+\frac{.4}{N}) \qquad \text{and}$$

$$\forall j, \sum_{\{i:(ij)\in I_2\}} y_{ij} \leq N(m-1)(1+\frac{.4}{N}).$$

So,

$$\sum_{ij} y_{ij}^2 = \sum_{I_1} y_{ij}^2 + \sum_{I_2} y_{ij}^2 \leq N^2(1+\frac{.4}{N})^2[(n-1)^2(m-1)+(m-1)^2(n-1)].$$

So the diameter $d = d(T)$ of T satisfies

$$d(T) \leq \sqrt{2}N^{1.5}\sqrt{(n+m-2)}(1+\frac{.4}{N}).$$

For every unit (length) vector $u \in \mathbf{R}^N$, with $u_{ij} \geq 0$, let $l(u)$ be the ray $\{y = \lambda u : \lambda \geq 0\}$ from the origin along u. Note that the ray intersects T in a line segment (since $y \in T$ iff $(0.4)\lfloor 2.5\ y \rfloor \in QQ'$ where floor denotes componentwise floor.) . Let $R = R(u)$ be the length of the segment $l(u) \cap QQ'(r, c)$ and $R_1 = R_1(u)$ be the length of the segment $l(u) \cap T$. Then there exists an $i_o \in \{1, 2, \ldots m-1\}$ such that

$$\sum_j R u_{i_o j} \frac{\rho_{i_o j}}{N} = r_{i_o}$$

or a $j_o \in \{1, 2, \ldots n-1\}$ such that $\sum_i R u_{ij_o} \frac{\rho_{ij_o}}{N} = c_{j_o}$. Assume without loss of generality the first option. Since $R_1 u$ belongs to T, the vector $R_1 u - \underline{0.4}$ belongs to QQ'. So, we also have,

$$\sum_j R_1 u_{i_o j} \frac{\rho_{i_o j}}{N} \leq r_{i_o} + (0.4)\sum_j \frac{\rho_{i_o j}}{N}.$$

Using the fact that $\sum_j \rho_{ij} \leq r_i$, we get that

$$R_1 \leq R(1 + \frac{.4}{N}).$$

The above implies that $R_1 \leq 2R$ as required by (7) of [[15]].

Also, Ru belongs to QQ' implies

$$R\sum_j u_{ij}\frac{\rho_{ij}}{N} \leq r_i \forall i \qquad R\sum_i u_{ij}\frac{\rho_{ij}}{N} \leq c_j \forall j$$

which implies by [2]

$$(R_1 - R)\sum_{i,j} u_{ij}\frac{\rho_{ij}}{N} \leq \frac{R_1 - R}{R}\sum_i r_i \leq \frac{0.4 r_m}{n-1}.$$

Similarly $(R_1 - R)\sum_{i,j} u_{ij}\frac{\rho_{ij}}{N}$ is also at most $0.4c_n/m-1$. Thus $e^{M(R_1-R)\sum_{i,j} u_{ij}\frac{\rho_{ij}}{N}}$ is at most $e^{1.6}$. This implies that we may take κ_1 of (8) of [[15]] to be $25/3$.

Also, $R_1 - R \leq .4R/N$. So κ_2 of [[15]] is given by (since δ, the step size is 0.4 here)

$$\kappa_2 \leq \frac{50}{3}(\frac{R}{N} + \sqrt{N}) = (50/3)(\sqrt{N} + o(1)). \tag{12}$$

Also,

$$\kappa_0 = \frac{1}{.8}N^3(n + m - 2)(1 + o(1)). \tag{13}$$

We now want to bound $(1 + \alpha)$ of [15]. To this end, we need to prove upper and lower bounds on

$$\frac{\int_{z \in C(y)} G(z)\, dz}{G(y)\delta^N}$$

where y is an lcc. Since G is a convex function, it is clear that 1 is a lower bound on this ratio. To get an upper bound, we use Heoffding's bounds [18] on the probability that the sum $\sum_{ij} \rho_{ij}[z_{ij} - y_{ij}]$ deviates from its mean of 0 and (after some complicated integration) arrive at an upper bound of $1.05e^{8/N^{1.5}}$.

Plugging all this into the formula for λ_2^{-1} in Theorem 2 of [[15]], we get :

Theorem 3

$$(1-\theta)^{-1} \leq [1.2e^{24/N^{1.5}}]\frac{N}{.4}[\frac{N^3}{.8}(n+m-2)(1+\frac{1.5}{N})(28/3)$$

$$+25\sqrt{N}\sqrt{n+m-2}(1+\frac{.8}{3N})+2N^{3/2}\sqrt{n+m-2}] \leq (35+o(1))N^4(n+m-2),$$

For $N \geq 10, n+m \geq 11$*, a calculation shows that the* $(35+o(1))$ *may be replaced by 36.*

References

[1] D. ALDOUS, "Some inequalities for reversible Markov chains", *Journal of the London Mathematical Society*, **(2)25**, pp. 564-576, 1982.

[2] D. APPLEGATE, AND R. KANNAN, "Sampling and integration of near log-concave functions", *23rd ACM Symposium on the Theory of Computing*, 1991.

[3] A. I. BARVINOK, "A polynomial time algorithm for counting integral points in polyhedra when the dimension is fixed", *Proceedings of the Symposium on the Foundations of Computer Science*, 1993.

[4] P. DIACONIS, AND B. EFRON, "Testing for independence in a two-way table", *Annals of Statistics*, **13**, pp 845-913, 1985.

[5] P. DIACONIS, AND A. GANGOLLI, "Rectangular arrays with fixed margins", To appear in the *Proceedings of the workshop on Markov Chains*, Institute of Mathematics and its Applications, University of Minnesota, (1994).

[6] P. DIACONIS, AND D. STROOCK, "Geometric bound for eigenvalues of Markov chains," *Annals of Applied Probability 1,* pp. 36-61, 1991.

[7] P. DIACONIS, AND B. STURMFELS, “Algebraic algorithms for sampling from conditional distributions”, Technical Report, Department of Statistics, Stanford University (1993)

[8] M. E. DYER, AND A. M. FRIEZE, “On the complexity of computing the volume of a polyhedron”, *SIAM J. Comput.*, **17** (1988), 27–37.

[9] M. E. DYER, AND A. M. FRIEZE, “Computing the volume of convex bodies: a case where randomness provably helps”, *Proceedings of Symposia in Applied Mathematics*, Vol 44, 1991.

[10] M. E. DYER, A. M. FRIEZE, AND R. KANNAN, “A Random Polynomial-Time Algorithm for Approximating the Volume of Convex Bodies”, *Journal of the Association for Computing Machinery*, **Vol 38 No.1** (January 1991), pp.1-17.

[11] A. SCHRIJVER, *Theory of linear and integer programming*, John Wiley, Chichester, 1986.

[12] R. P. STANLEY, *Enumerative combinatorics*, Vol. I, Wadsworth and Brooks/Cole, Monterey, CA, 1986.

[13] M. E. DYER, A. FRIEZE, R. KANNAN, A. KAPOOR, L. PERKOVIC, AND U. VAZIRANI, “A mildly exponential time algorithm for approximating the number of solutions to a multidimensional knapsack problem”, *Combinatorics, Probability and Computing*, (to appear).

[14] J. A. FILL, “Eigenvalue bounds on the convergence to stationarity for nonreversible Markov Chains, with an application to the exclusion process”, *The Annals of Applied Probability*, Vol. 1, No. 1, pp 62-87, 1991.

[15] A. FRIEZE, R. KANNAN AND N. POLSON, "Sampling from log-concave distributions", *The Annals of Applied Probability*, Vol. 4, No. 3 (1994) pp 812-837.

[16] A. R. GANGOLLI, *Convergence Bounds for Markov Chains and Applications to Sampling,* Phd. Thesis, Stanford STAN-CS-91-1361, 1991.

[17] D. GILLMAN, "A Chernoff bound for random walks on expander graphs," *Proceedings of the 34th Annual IEEE Conference on Foundations of Computer Science*, 1993.

[18] W. HOEFFDING "Probability Inequalities For Sums of Bounded Random Variables", American Statistical Association Journal, March 1963, pp. 13-30.

[19] N. KAHALE, "Large deviation bounds on Markov Chains", DIMACS Technical Report, 94-39, 1994.

[20] L. LOVÁSZ, AND M. SIMONOVITS, "Random Walks in a Convex Body and an Improved Volume Algorithm", Revised version, March 1993.

[21] C. R. MEHTA, N. R. PATEL, "A Network Algorithm for Performing Fisher's Exact Test in $r \times c$ Contingency Tables," *Journal of the American Statistical Association 78*, pp. 427-434, 1983.

[22] R.KANNAN, L.LOVÁSZ AND M. SIMONOVITS, "Random Walks and an $O^*(n^5)$ volume algorithm for convex bodies", in preparation.

[23] E. SENETA, *Non-negative matrices and Markov Chains*, Springer-Verlag, New York, 1973.

[24] B. STURMFELS, "How to Count Integer Matrices," manuscript, December 13, 1993.

[25] L. G. VALIANT, "The complexity of computing the permanent," *Theoretical Computer Science 8,* pp. 189-201, 1979.