RANDOM WALKS, TOTALLY UNIMODULAR MATRICES, AND A RANDOMISED DUAL SIMPLEX ALGORITHM

Martin Dyer*
School of Computer Studies, University of Leeds,
Leeds, U.K.

and

Alan Frieze†
Department of Mathematics, Carnegie-Mellon University,
Pittsburgh, U.S.A.

December 8, 2001

Abstract

We discuss the application of random walks to generating a random basis of a totally unimodular matrix and to solving a linear program with such a constraint matrix. We also derive polynomial upper bounds on the combinatorial diameter of an associated polyhedron.

*Supported by NATO grant RG0088/89
†Supported by NSF grants CCR-8900112, CCR-9024935 and NATO grant RG0088/89

1 Introduction

A matrix is *totally unimodular* if all its subdeterminants are $0, \pm 1$. Totally unimodular matrices are known to be the matrix representations of *regular matroids* [21]. Let A be a totally unimodular $m \times n$ matrix with columns $a_1, a_2, \ldots, a_n$. Let $\mathcal{B}$ denote the set of *bases* of A, i.e. the set of $m \times m$ non-singular submatrices of A. These correspond to the bases of the associated regular matroid. We will, by abuse of terminology, identify A, or a basis B, with the set of columns they contain. Now, assuming A has full row rank (i.e. $\mathcal{B} \neq \emptyset$), we can define a simple random walk on $\mathcal{B}$:

NATURAL RANDOM WALK

Starting at an arbitrary basis $B_0 = B \in \mathcal{B}$, generate a random sequence $B_0, B_1, \ldots, B_t, \ldots \in \mathcal{B}$ as follows. At B_t, randomly choose columns $a \in B_t$, $a' \in A \setminus B_t$. Let $B_t' = B_t \cup \{a'\} \setminus \{a\}$. If $B_t' \in \mathcal{B}$ then $B_{t+1} = B_t'$, otherwise $B_{t+1} = B_t$.

Now the steady state distribution of this chain is uniform over $\mathcal{B}$. (This is clear from the fact that the probability transition matrix is symmetric.) The crucial question is how quickly does the chain settle down ? Does it mix rapidly–Aldous [1] ? In the particular case of our problem in which A is the node-arc incidence matrice of a (di)graph, or correspondingly the regular matroid is *graphic*, the answer is known to be affirmative–Aldous [2], Broder [6]. We will extend this result and prove

Theorem 1 *For any* $B \in \mathcal{B}$

$$|\mathbf{Pr}[B_t = B] - |\mathcal{B}|^{-1}| \leq \left(1 - \frac{1}{8m^4n^4}\right)^t.$$

$\square$

It is clear from this theorem that in time polynomial in $m, n, \ln(1/\delta)$ we can generate a basis $\hat{B}$ of A such that for any $B \in \mathcal{B}$ we have

$$|\mathbf{Pr}[\hat{B} = B] - |\mathcal{B}|^{-1}| \leq \delta,$$

i.e. $\hat{B}$ is an "almost uniformly generated" basis. We will prove the above theorem by relating the walk to one in a certain convex polytope $\zeta(A)$ associated with A, called its *zonotope*. For an arbitrary matrix A, the associated zonotope is defined by

$$\zeta(A) = \{x \in \boldsymbol{R}^m : x = Ay \text{ for some } y \in [0,1]^n\}.$$

The mixing time of Markov chains has attracted much attention lately in the Computer Science community, since the efficiency of various algorithms depends on this, e.g. Broder [5], Jerrum and Sinclair [22, 13, 14], Dyer, Frieze and Kannan [11], Karzanov and Khachyan [16], Lovász and Simonovits [17], Applegate and Kannan [3], Dyer and Frieze [10], Mihail and Winkler [19]. Theorem 1 could be used, in particular, to show how to estimate $|\mathcal{B}|$ although as we shall see, this can be done more efficiently using the Binet-Cauchy formula for the determinant of the product of two rectangular matrices.

As already observed, one can consider $|\mathcal{B}|$ to be the set of bases of the regular matroid associated with A. We note that, if one could generalise Theorem 1 to an arbitrary matroid, then one could (for example) efficiently estimate the reliability of a graph. This would be an important result.

Our second result concerns linear programs with totally unimodular matrices of coefficients. We consider the problem

$$\begin{array}{rlrcl} \mathrm{LP}(b) : & \text{minimise} & cy & & \\ & \text{subject to} & Ay & = & b \\ & & y & \geq & 0\ . \end{array}$$

Although it is already known, through the work of Tardos [24], that a strongly polynomial time algorithm exists for this problem when A is totally unimodular, we will show that random walks may be used to give a randomised version of this result. We will defer the exact statement until later. What we show, in essence, is that the above problem can be solved by a randomised dual simplex algorithm, where pivots are chosen by executing a random walk. This is discussed in Section 4. As a corollary we will be able to give a polynomial upper bound on the combinatorial diameter of the polytope $P = \{x \in \boldsymbol{R}^m : A^T x \leq c\}$ when A is totally unimodular.

2 Notation and terminolgy

Conductance

Consider a Markov chain with finite state space Ω, transition matrix $P = (p_{ij})$ and steady state distribution π_i. If $I \subseteq \Omega$ with $\pi(I) = \sum_{i \in I} \pi_i \leq \frac{1}{2}$, let

$$\Phi_I = \sum_{i \in I, j \notin I} \pi_i p_{ij} / \pi(I),$$

i.e. Φ_I is the conditional probability that the next state will be in $\bar{I}$ given that the current state is in I, assuming the steady state distribution of the chain. The conductance of the chain is then $\Phi = \min_I \Phi_I$. The following theorem is implicit in [22], and stated more directly in [17].

Theorem 2 *If $J \subseteq \Omega$ and $\pi^{(t)}(J)$ is the probability that the chain is in a state of J at time t when started in state 1, then*

$$|\pi^{(t)}(J) - \pi(J)| \leq \sqrt{\pi(J)/\pi_1}(1 - \Phi^2/2)^t.$$

□

Norms

For a vector $x \in \boldsymbol{R}^n$ we use $\|x\|$ to denote the Euclidean or L_2 norm. We will also need to use the L_1 and L_∞ norms which are written $\|x\|_1, \|x\|_\infty$ respectively.

3 Analysis of the walk

We start by proving a result on the volume $\mathrm{vol}_m(\zeta(A))$ which is attributed by Stanley [23] to McMullen. We give a proof which provides us with a decomposition of $\zeta(A)$ into parallelopipeds which is fundamental to our walks. Here A need not be totally unimodular.

Theorem 3 $vol_m(\zeta(A)) = \sum_{B \in \mathcal{B}} |\det B|$.

Proof If $B \in \mathcal{B}$ then $\zeta(B)$ is a parallelopiped with edges parallel to the columns of B and it is well known that

$$\mathrm{vol}_m(\zeta(B)) = |\det B|. \tag{1}$$

Let $c \in \boldsymbol{R}^n$ be arbitrary, $\theta > 0$ be arbitrarily small, and

$$c_\theta = (c_1 + \theta, c_2 + \theta^2, \ldots, c_n + \theta^n).$$

For each $x \in \zeta(A)$, let $P_x = \{y \in [0,1]^n : Ay = x\}$, and let $\xi(x)$ be the unique optimal solution to

$$\text{minimise } c_\theta y \quad : \quad y \in P_x. \tag{2}$$

Since θ is small, $\xi(x)$ is the (unique) lexicographical optimum to the above problem. Thus there is some basis $B_x \in \mathcal{B}$ with columns $a_i, i \in I_x$ such that $\xi_j(x) \in \{0,1\}$ for $j \notin I_x$. (If there is a choice of I_x due to degeneracy, choose the lexicographically first such.) Let now $\hat{\xi}(x)$ be defined by $\hat{\xi}_j(x) = \xi_j(x), j \notin I_x$ and $\hat{\xi}_i(x) = 0, i \in I_x$. Then

$$x \in A\hat{\xi}(x) + \zeta(B_x).$$

Observe also that if $x' \in A\hat{\xi}(x) + \zeta(B_x)$ and $0 < x'_i < 1$ for $i \in I_x$ then $I_{x'} = I_x$, $B_x = B_{x'}$ and $\xi(\hat{x'}) = \xi(\hat{x})$. This is because changing the right hand side of the linear program (2) in this way does not affect the optimality of the basis B_x.

Conversely, suppose $B \in \mathcal{B}$ with columns a_i, $i \in I$. There is a unique $\eta(B) \in \{0,1\}^n$, with $\eta_i = 0$ $(i \in I)$, such that if $x \in A\eta(B) + \zeta(B)$ and $0 < x_i < 1$ for $i \in I$, then $B_x = B$ and $\hat{\xi}(x) = \eta(B)$. Indeed, for any such x, the optimality conditions

$$c_j + \theta^j - (c_{i_1} + \theta^{i_1}, \ldots, c_{i_m} + \theta^{i_m})B^{-1}a_j \begin{array}{llll} > 0 & \Rightarrow & y_j = 0, & j \notin I \\ < 0 & \Rightarrow & y_j = 1, & j \notin I, \end{array} \tag{3}$$

where $I = \{i_1 < \ldots < i_m\}$, will ensure $B_x = B$. Since θ is arbitrarily small, the conditions (3) can of course be rewritten to be independent of θ. Summarising, the set S of parallelopipeds $\{\hat{\xi}(x) + \zeta(B_x) : x \in \zeta(A)\}$ cover $\zeta(A)$, intersect on a set of zero volume and each $B \in \mathcal{B}$ gives rise to a unique $P_B \in S$. The theorem now follows. $\square$

3.1 Digression: counting bases

Returning to the case where A is totally unimodular, we see immediately that

$$\mathrm{vol}_m(\zeta(A)) = |\mathcal{B}|.$$

Thus we could approximate $|\mathcal{B}|$ by estimating the volume of $\zeta(A)$. This is however, not the easiest way of accomplishing this. By the identity of Binet-Cauchy (see, for example, [4, p327]) we have

$$\det AA^T = \sum_I (\det B_I)^2 = |\mathcal{B}|,$$

where the above sum ranges over all m-sets $I \subseteq [n] = \{1, 2, \ldots, n\}$ and B_I is the $m \times m$ submatrix of A with columns $a_i, i \in I$. Thus $|\mathcal{B}|$ can be computed exactly by evaluating a determinant. This can be done in polynomial time, and even in NC.

This observation can be generalised somewhat. We can compute, for example, the number α_k of $k \times k$ nonsingular submatrices of A. Let A' be the $m \times (m+n)$ matrix obtained from A by adding an $m \times m$ identity matrix I_m. Then α_k is the number of bases of A' which have $(m-k)$ columns in I_m. In general suppose we have an $M \times N$ totally unimodular matrix D, $S \subseteq [N]$, and we wish to compute the numbers $\beta_k, k = 0, 1, \ldots, M$ of bases of D which have k columns with indices in S. After column rearrangement, let $D = [D_1 \mid D_2]$, where D_1 contains the columns with indices in S. Suppose z is a complex variable, and let $D_z = [\sqrt{z} D_1 \mid D_2]$. (The choice of branch for the square-root is unimportant.) The Binet-Cauchy theorem implies that

$$\det D_z D_z^T = \det(z D_1 D_1^T + D_2 D_2^T) = \sum_{k=0}^{M} \beta_k z^k, \tag{4}$$

and we can compute the β_k by evaluating the coefficients of the above polynomial by interpolation, using only rational values of z.

More generally still, suppose for any fixed r, we partition the columns of D into sets $S_1, S_2, \ldots, S_r$ and ask for the number of bases with k_1 columns in S_1, ..., k_r columns in S_r, where $k_1 + \cdots + k_r = M$. Call this number $\beta(k_1, \ldots, k_r)$. Letting $D = [D_1 \mid \cdots \mid D_r]$ and $D_z = [\sqrt{z_1} D_1 \mid \cdots \mid \sqrt{z_r} D_r]$, gives

$$\begin{aligned} \det D_z D_z^T &= \det(z_1 D_1 D_1^T + \cdots + z_r D_r D_r^T) \\ &= \sum_{k_1+\ldots+k_r=M} \beta(k_1, \ldots, k_r) z_1^{k_1} \cdots z_r^{k_r}. \end{aligned} \tag{5}$$

The right hand side is a homogeneous polynomial in $z_1, \ldots, z_r$ whose coefficients can again be determined by interpolation.

We note in passing that equation (4) gives strong information about the numbers β_k. It follows easily, by simultaneous diagonalisation of the pair of positive semidefinite matrices $D_1 D_1^T$ (see [4]), $D_2 D_2^T$, that the polynomial $\sum_{k=0}^{M} \beta_k z^k$ has only nonpositive real roots. This answers a question of Stanley [23], and implies (as he observes) that $\gamma_k = \binom{n}{k}^{-1} \beta_k$ is a log-concave sequence, i.e. $\gamma_k^2 \geq \gamma_{k-1}\gamma_{k+1}$ $(k = 1, 2, \ldots, M-1)$. (See [4, p53] for a proof of the log-concavity of the sequence $\{\gamma_k\}$ in this situation.)

It follows further that the coefficients $\beta(k_1, \ldots, k_r)$ are $\prod_{i=1}^{r} k_i!$ times the *mixed discriminants* of the quadratic forms $x^T(D_i D_i^T)x$ $(i = 1, \ldots, r)$, where $x \in \boldsymbol{R}^n$. See [7, p169] for definitions and properties. It is now easy to derive Stanley's theorem (Theorem 2.1 of [23]) on the log-concavity properties of the $\beta(k_1, \ldots, k_r)$ from the theory of mixed discriminants, in a similar way to which Stanley derives it from the theory of *mixed volumes.* Mixed volumes and discriminants have many close relationships with enumeration. Approximation of mixed volumes is considered in [9].

3.2 Analysis of the walk continued

We will consider a random walk on $\mathcal{B}$ which is slightly different from the one described in Theorem 1. We call it the *restricted* random walk. Take $c = 0$ in Theorem 3 and consider the graph $\Gamma = (\mathcal{B}, \mathcal{E})$ in which B, B' are adjacent if $P_B, P_{B'}$ share an $(m-1)$-dimensional face. If B now has degree d_B in Γ we add $(4m - d_B)$ loops at B. We consider the usual random walk on Γ.

We start at an arbitrary basis B_0, which we can find in $O(m^2 n)$ time. We then re-arrange the columns of A so that those of B_0 are last and then we have $P_{B_0} = \zeta(B_0)$. In a general step of the walk, at basis B_t, we randomly choose one of the $4m$ edges of Γ that are incident with B_t and traverse it. If the edge is not a loop, then moving to the associated neighbouring base requires a dual simplex pivot and this can be carried out in $O(mn)$ time. (We select the column to leave the basis, whether its associated variable is to become non-basic at 0 or 1, and then find the unique column to replace it, if one exists. Furthermore there is no need to keep an explicit value for θ since the replacing column is independent of θ whenever θ is sufficiently small.) For $B \in \mathcal{B}$ let $\pi_B^{(t)} = \mathbf{Pr}(B_t = B)$. We are interested in the distance between the distribution $\pi^{(t)}$ and the (uniform) steady state distribution. This can be bounded in terms of the conductance Φ of the associated Markov chain. Theorem 2 implies that for our random walk

$$|\pi_B^{(t)} - |\mathcal{B}|^{-1}| \leq \left(1 - \frac{\Phi^2}{2}\right)^t \tag{6}$$

where, for $S \subseteq \mathcal{B}$,

$$\Phi = \min_{|S| \leq \frac{1}{2}|\mathcal{B}|} \Phi_S.$$

Here, letting $e(S : \bar{S})$ denote the number of edges between S and $\bar{S}$ in Γ, it follows easily that

$$\Phi_S = \frac{e(S : \bar{S})}{4m|S|}. \tag{7}$$

To put a lower bound on Φ, fix S and let $R = \bigcup_{B \in S} P_B$, $W = \mathrm{vol}_{m-1}(\partial R \setminus \partial\zeta(A))$, where ∂ denotes boundary. Now, by a theorem of Lovász and Simonovits [17], with a small improvement from Dyer and Frieze [10],

$$\begin{aligned} W &\geq \frac{2\,\mathrm{vol}_m(R)}{\mathrm{diam}(\zeta)} \\ &\geq \frac{2\,|S|}{n\sqrt{m}}, \end{aligned} \tag{8}$$

where we have used the fact that, since A has entries $0, \pm 1$,

$$\mathrm{diam}(\zeta) \leq \sum_{i=1}^{n} \|a_i\| \leq n\sqrt{m},$$

On the other hand

$$W \leq \sqrt{m}\, e(S : \bar{S}), \tag{9}$$

since the area of each facet of each P_B is at most $\sqrt{m}$. To see this suppose that $a_1, a_2, \ldots, a_m$ are linearly independent. Let b be the normal to the hyperplane H through the origin generated by $a_1, a_2, \ldots, a_{m-1}$. Assume, after relabelling coordinates if necessary, that $b_1 \neq 0$ and scale so that $b_1 = 1$. It follows from Cramer's rule applied to the equations $b_1 = 1, b \cdot a_i = 0, 1 \leq i \leq m-1$ that $b_i \in \{0, \pm 1\}, 1 \leq i \leq m-1$. Now the perpendicular distance h from the point a_m to H is $|b \cdot a_m| / \|b\|$. But $b \cdot a_m$ is a non-zero integer and $\|b\| \leq \sqrt{m}$. The area upper bound follows since h^{-1} is the area of the face of the parallelopiped generated by $a_1, a_2, \ldots, a_{m-1}$. Applying (9) in (8) gives

$$\Phi \geq \frac{1}{2m^2 n}. \tag{10}$$

To obtain Theorem 1 we need only argue that the conductance Φ' of the natural random walk is at least Φ/n. But this follows easily from the definition of conductance. Fix a set $S \subseteq \mathcal{B}$. Let $B_{NW}^{(t)}$ and $B_{RW}^{(t)}$ refer to the t'th of the natural and restricted random walks respectively, assuming they are both started in their steady state, i.e. uniformly on $\mathcal{B}$. Let

$$\pi_S = \mathbf{Pr}(B_{NW}^{(0)} \in S) = \mathbf{Pr}(B_{RW}^{(0)} \in S) = |S|/|\mathcal{B}|.$$

Then, in an obvious notation,

$$\begin{aligned}
\pi_S \Phi'_S &= \sum_{B \in S, B' \in \bar{S}} \mathbf{Pr}(B_{NW}^{(0)} = B \text{ and } B_{NW}^{(1)} = B') \\
&\geq \sum_{B \in S, B' \in \bar{S}} \frac{1}{n-m} \mathbf{Pr}(B_{RW}^{(0)} = B \text{ and } B_{RW}^{(1)} = B') \\
&= \frac{1}{n-m} \pi_S \Phi_S.
\end{aligned}$$

where the inequality follows from the fact that at any basis B the natural random walk has probability $\frac{1}{n-m}$ of making the same move as the restricted random walk. This clearly proves $\Phi' > \Phi/n$ and Theorem 1 follows.

4 Linear Programming

In this section we consider the effectiveness of a random walk in solving the linear program LP(b) of Section 1. Our randomised algorithm works on the assumption that LP(b) is feasible, in which case it solves it with high probability. The assumption of feasibility is no restriction, since we can use, for example, the "big-M" method [21, p136] if we suspect the problem may be infeasible.

We first describe an algorithm METROPLEX which (almost) solves the problem $\mathrm{LP}'(b)$ in which $y \geq 0$ is replaced by $0 \leq y_j \leq 1$, $j = 1, 2, \ldots, n$. We will then use this algorithm as a subroutine to solve $\mathrm{LP}(b)$.

Assume that we are given $0 < d, \delta < 1$. Our result will be

Theorem 4 *With probability at least* $(1 - \delta)$*, METROPLEX computes* $\hat{y}$ *such that, if* $\hat{b} = A\hat{y}$*,*

1. $\|\hat{b} - b\| \leq d$
2. $\hat{y}$ *solves* $LP(\hat{b})$

Furthermore METROPLEX runs in $O(m^{11.5} n^4 d^{-3} \ln^3(m/(d\delta))$ *time.*

METROPLEX uses a random walk to choose dual simplex pivots in a manner to be described shortly. We start conceptually with the decomposition of $\zeta(A)$ into parallelopipeds induced by c as given in Theorem 3. Let

$$\begin{aligned} M &= \lceil 8((m+1)\ln(160m^3/d) + \ln(4/\delta))/d \rceil, \\ N &= \lceil 5Mm^{3/2} \rceil, \\ T &= \lceil 9m^4 n^2 N^2 (Mn\sqrt{m} + m\log(15Mn\sqrt{m}) + 2\log(2/\delta)) \rceil. \end{aligned}$$

Each parallelopiped P_B is divided into N^m sub-parallelopipeds (which we will call *cells*) of equal size in an obvious way. Denote these by $P_{B,i}, i =$

$1, 2, \ldots, N^m$ and let $\sigma_{B,i}$ denote the centre of $P_{B,i}$. Also for $x \in \zeta(A)$ let $\phi(x) = e^{-M\|x-b\|}$ and $\psi(P_{B,i}) = \phi(\sigma_{B,i})$.

We do a random walk on the cells as follows. We start with an arbitrary basis B and choose a cell in P_B, any will do. Suppose now that $P = P_{B,i}$ and $P' = P_{B',i'}$ share a common facet. When the walk is at P the (transition) probability $\lambda(P, P')$ that the next state is P' is given by

$$\lambda(P, P') = \frac{1}{4m} \min \left\{ 1, \frac{\psi(P')}{\psi(P)} \right\}.$$

This is achieved by choosing a neighbouring cell at random (as in our first random walk) and moving to it with probability $\psi(P')/\psi(P)$. This is an example of the Metropolis algorithm [18]. It is straightforward to check that the steady state probabilities $\pi(P)$ are proportional to $\psi(P)$. We run the walk for T steps and then with high probability the current cell P is "close" to the cell which contains b. If in fact $P \subseteq P_B$ where $b \in P_B$ then B will be an optimal basis for $\mathrm{LP}'(b)$ and this will be easily recognisable and we will have solved $\mathrm{LP}'(b)$. In general we cannot conclude this although, when d is small and b is chosen randomly from $\zeta(A)$, this is likely to be the case.

Before analysing the mixing time of this walk we remark that the algorithm does have the flavour of a randomised dual simplex algorithm. Moving from cell to cell in the same parallelopiped requires adjusting the value of one basic variable and moving from parallelopiped to parallelopiped requires a dual simplex pivot as before. The random walk between pivots is used to determine which column should leave the basis.

We now discuss the conductance of the walk. Let Ω denote the set of all cells. Let $S \subseteq \Omega$ satisfy $\pi(S) \leq \frac{1}{2}$ and let $\bar{S} = \Omega \setminus S$. Let

$$(S : \bar{S}) = \{(P, P') : P \in S, P' \in \bar{S} \text{ and } P, P' \text{ share a facet}\}.$$

$$\begin{aligned} \Phi_S &= \frac{\sum_{(P,P') \in (S:\bar{S})} \pi(P) \lambda(P, P')}{\pi(S)} \\ &= \frac{\sum_{(P,P') \in (S:\bar{S})} \min\{\psi(P), \psi(P')\}}{4m \sum_{P \in S} \psi(P)} \end{aligned}$$

Observe next that if x, x' lie in the same cell P then

$$\begin{aligned} \phi(x)/\phi(x') &\leq \exp\{M \text{diam}(P)\} \\ &\leq \exp\left\{\frac{Mm^{3/2}}{N}\right\} \\ &\leq e^{1/5} \end{aligned}$$

Thus, if $R = \bigcup\{P : P \in S\}$ and $W = \partial R \setminus \partial \zeta(A)$, then

$$N^{-m} \sum_{P \in S} \psi(P) \leq e^{1/5} \int_R \phi(x) dx$$

and

$$N^{1-m} \sqrt{m} \sum_{(P,P') \in (S:\bar{S})} \min\{\psi(P), \psi(P')\} \geq e^{-1/5} \int_W \phi(x) dx.$$

(The volume of each cell is N^{-m} and each of its facets has area at most $N^{1-m}\sqrt{m}$.) It follows that

$$\begin{aligned} \Phi_S &\geq \frac{e^{-2/5}}{4Nm^{3/2}} \frac{\int_W \phi(x) dx}{\int_R \phi(x) dx} \\ &\geq \frac{e^{-2/5}}{4Nm^{3/2}} \cdot \frac{2}{\text{diam } \zeta(A)} \\ &\geq \frac{1}{3m^2 nN}. \end{aligned}$$

The second inequality is a sharpening of an inequality of Applegate and Kannan [3] from Dyer and Frieze [10]. This inequality is a generalisation of the Lovász-Simonovits isoperimetric inequality used in Theorem 1 and relies on the fact that $\phi(x)$ is log-concave. Thus the conductance Φ satisfies

$$\Phi \geq \frac{1}{3m^2nN}.$$

We need an upper bound on the total probability of cells at distance greater than d from b. Now each cell $P_{B,i}$ contains a ball of radius $\rho = (2\sqrt{m}N)^{-1}$ centred at $\sigma_{B,i}$. Hence the number of cells whose centres lie in a ball of radius r is at most $((r+\rho)/\rho)^m$. Hence if Ω_d denotes the set of cells with centre at distance at least d from b then

$$\begin{aligned}
\psi(\Omega_d) &\leq \sum_{k=1}^{\infty}\left(\frac{d+k\rho}{\rho}\right)^m \exp\{-M(d+(k-1)\rho)\} \\
&= e^{-M(d-\rho)}\rho^{-m}\Big(\sum_{k\rho\leq d}(d+k\rho)^m e^{-Mk\rho} + \sum_{k\rho>d}(d+k\rho)^m e^{-Mk\rho}\Big) \\
&\leq e^{-M(d-\rho)}\rho^{-m}\left(2^m d^m + 10m^2 2^m d^m e^{-Md}\right).
\end{aligned}$$

For the first sum we replace $(d+k\rho)$ by $2d$. For the second sum we replace $(d+k\rho)$ by $2k\rho$ and then observe that the ratio of successive terms is

$$(1+1/k)^m e^{-M\rho} \leq \exp\left\{\frac{m}{k} - M\rho\right\} \leq \exp\left\{\frac{m\rho}{d} - M\rho\right\} \leq e^{-7M\rho/8} \leq 1-\frac{1}{40m^2},$$

since $M > 8m/d$, and $M\rho \geq 1/(6m^2)$. Now, using this latter inequality again, together with $M \geq 4$, $\rho \leq d/2$,

$$\psi(\Omega_d) \leq e^{-Md/2}(40Mm^2)^{m+1}. \tag{11}$$

To bound this, let $\lambda = Md/(2(m+1))$. Then

$$\begin{aligned}
\psi(\Omega_d) &\leq (80\lambda e^{-\lambda} m^2(m+1)/d)^{m+1}. \\
&\leq (160 e^{-\lambda/2} m^3/d)^{m+1} \\
&\leq \frac{\delta}{4},
\end{aligned}$$

provided $\lambda \geq 2(\log(160m^3/d) + \frac{1}{m+1}\log(4/\delta))$. Since our value of M is large enough to ensure this, we have the bound. Moreover, since $\phi(b) = 1$, we see that if b lies in a cell P then $\psi(P) \geq \frac{1}{2}$. Thus

$$\pi(\Omega_d) \leq 2\psi(\Omega_d) \leq \delta/2.$$

We also need a lower bound on the steady state probability of the initial cell. If this is P_0, then using the bound on the diameter of $\zeta(A)$,

$$\psi(P_0) \geq \exp\{-Mn\sqrt{m}\}.$$

Also, since $\psi(P) \leq 1$ for all cells $P \in \Omega$,

$$\psi(\Omega) \leq \binom{n}{m} N^m \leq (enN/m)^m \leq (15Mn\sqrt{m})^m.$$

Thus

$$1/\pi(P_0) \leq (15Mn\sqrt{m})^m \exp\{Mn\sqrt{m}\} = C_0,$$

say. We now use our conductance bound and Theorem 2. This implies that after T steps the probability we are in Ω_d is at most

$$\begin{aligned} \left(\frac{\delta}{2} + \sqrt{C_0}\left(1 - \frac{\Phi^2}{2}\right)^T\right) &\leq \frac{\delta}{2} + \sqrt{C_0}\exp\left\{-\frac{T}{18m^4n^2N^2}\right\} \\ &\leq \delta, \end{aligned}$$

on substituting the relevant values. The time estimate for METROPLEX is $O(Tmn)$ and Theorem 4 is proven.

We now return to the solution of LP(b). Let x^* be the (lexicographically first) optimum solution. Since it is a basic solution, and A is totally unimodular, we have

$$\|x^*\|_\infty \leq \beta = \|b\|_1.$$

Now change variables to $y = x/2\beta$ so that $\|y^*\| \leq \frac{1}{2}$ and define $b' = b/(2\beta)$. We now run METROPLEX on the problem LP$'(b')$ with $d = \frac{1}{6m^2n}$ and $\delta = \frac{1}{2m}$. Suppose it is successful and produces $\hat{y}$ and $\hat{b} = A\hat{y}$ where $\|\hat{b}\|_1 \geq \frac{1}{3}$. It follows from a theorem of Cook, Gerards, Schrijver and Tardos (see Schrijver [21, p126]) that $\|\hat{y} - y^*\|_\infty \leq nd = \frac{1}{6m^2}$ where $y^* = x^*/(2\beta)$ solves LP$'(b')$. Clearly $\hat{y}_j = 0$ for non-basic j (else $\|\hat{y} - y^*\|_\infty \geq \frac{1}{2}$.) Also, since A is a 0,$\pm$1 matrix and $\hat{y}$ is a basic solution, the largest component $\hat{y}_\ell$ of $\hat{y}$ is at least $\frac{1}{m}\|\hat{b}\|_\infty \geq \frac{1}{3m^2}$. Hence

$$y_\ell^* \geq \hat{y}_\ell - \frac{1}{6m^2} \geq \frac{1}{6m^2}.$$

But this implies that x_ℓ^* is basic. Knowing this we can eliminate x_ℓ and one row from the problem LP(b). Hence after m successful iterations the problem will be solved. We thus have the following

Theorem 5 *With probability at least $\frac{1}{2}$ the above algorithm solves LP(b). The running time is*
$O(m^{17.5}n^7\ln^3(mn))$ and so the algorithm is strongly polynomial. □

Of course repeated applications of the above algorithm will make the failure probability as small as we like.

Remark

The above analysis can be applied to matrices A whose entries are $0, \pm 1, \pm 2$ and in which the sum of the absolute values in each column is at most 2. Call this a G-matrix and observe that the property is preserved under Gaussian elimination after removing the pivot row and column i.e. after removing the "discovered" basic variable. The crucial property for the success of METROPLEX is that the ratio of face surface area to volume is polynomially bounded for each cell. With a polynomial ratio of area to volume we can use (weighted) surface area and volumes as approximations in estimating conductance.

Assume A is a G-matrix and B is a basis matrix. Observe that, if B_1 is the matrix formed by deleting column 1 from B, then the (m-1)-dimensional volume of $\zeta(B)$ spanned by the last m-1 columns of B satisfies

$$\begin{aligned} S^2 &= \det\, B_1^T B_1 \\ &= \sum_{i=1}^{m} (\det\, B_{i,1})^2, \end{aligned}$$

where $B_{i,1}$ is the matrix obtained by deleting row i and column 1. We will have (almost) justified our remark if we can prove, say, that

$$|\det\, B_{1,1}| \leq 2|\det\, B|. \tag{12}$$

Let the entries of B be denoted $b_{i,j}$. If column 1 of B has a unique non-zero α then $|\det\, B| = |\alpha||\det\, B_{1,1}|$. So assume w.l.o.g. that $b_{1,1} = b_{2,1} = 1$. A square matrix can be re-ordered to yield a decomposition into minimal (indecomposable) diagonal blocks. We can assume w.l.o.g. that B is indecomposable. An $m \times m$ non-singular G-matrix is associated with an m-vertex unicyclic connected graph Γ (the columns define the edges, ± 2's yielding loops). It follows that $|\det\, B| = 1$ or 2. Now, perhaps after some re-arrangement, the block decomposition of $B_{1,1}$ consists of a number of ± 1's plus a single non-trivial block which is a G-matrix (2 or more such blocks would mean that Γ has 2 or more cycles). Thus $|\det\, B_{1,1}| = 1$ or 2 and (12) follows.

The above analysis shows that the conductance is sufficiently large. The remainder of the proof can easily be justified once we observe that B^{-1} has entries in $\{0, \pm\frac{1}{2}, \pm 1, \pm 2\}$.

5 Diameter of a polyhedron

In this section we give a polynomial bound on the combinatorial diameter γ_Q of the polyhedron

$$Q = \{x \in \boldsymbol{R}^m : A^T x \leq c\}$$

where A is a totally unimodular $m \times n$ matrix. (By combinatorial diameter we mean the diameter of the graph induced by the vertices and edges of Q.)

We can assume that Q is non-degenerate. If not then a change of c to c_θ will make it non-degenerate and the combinatorial diameter will not decrease.

Take two vertices v_1, v_2 of Q. For $i = 1, 2$ the support of v_i decomposes A into B_i and N_i and correspondingly c into c_B^i, c_N^i such that

$$B_i^T v_i = c_B^i, \quad N_i^T v_i < c_N^i, \quad i = 1, 2.$$

Let e denote the m-dimensional vector of all 1's and $\lambda_i = \|B_i e\|_1$ $(i = 1, 2)$. Note $\lambda_i \geq 1$. We then let

$$b_i = \frac{1}{20\lambda_i\sqrt{m}} B_i e.$$

Clearly $b_i \in \zeta(A)$. Note that $\|b_i\| \leq \|b_i\|_1 = 1/(20\sqrt{m})$. Thus v_i is the (unique) optimum solution to

$$\text{maximise } b_i^T x \text{ subject to } x \in Q.$$

By duality B_i is also the optimum basis matrix for the problem LP(b_i) of Section 2.

We discuss applying a modification of the random walk of the previous section starting at b_1 with *target* b_2. The modification will be that we will only walk on cells Ω' for which the centre

$$\sigma \in D = \{b : \|b - b_2\| \leq \frac{9}{10\sqrt{m}}\}.$$

Note that the cell containing b_1 belongs to Ω'. By doing this we ensure that (in the notation of Theorem 3) $\hat{\xi}(b) = 0$ for each cell centre b. In fact $\hat{\xi}(b) \neq 0$ implies $\|b\| \geq 1/\sqrt{m}$. To see this observe that now b is outside any of the parallelopipeds which contain the origin and we have shown (see the argument following (9)) that the distance from the origin to any facet of any of these parallelopipeds is at least $1/\sqrt{m}$.) Thus the upper bounds $y_j \leq 1$ are inactive in the walk and the bases met are dual feasible for LP(b), for some b, and so feasible for the polyhedron Q.

Once we show that the conductance of this walk is bounded below by $1/p(m,n)$ for some polynomial p we will be almost done. For then we will have shown that we can get from B_1 to B_2 (with positive probability) by a walk involving a polynomial number of pivots such that all intermediate bases are feasible for Q. We will then be able to prove

Theorem 6

$$\gamma_Q = O(m^{16}n^3(\ln(mn))^3).$$

□

(Our current estimate for the polynomial must be far from the truth.)

Let us first change the sizes of M, T to

$$M = \lceil 100m^{7/2}n\ln(e^{10}m^6n^4)\rceil,$$

$$T = \lceil 10m^{3/2}MN^2 \rceil.$$

N remains the same function of M and we promise to run METROPLEX with

$$d = \frac{1}{30m^{3/2}n} \text{ and } \delta = \frac{1}{2m}.$$

In order to prove the theorem we use the notion of μ-conductance introduced by Lovász and Simonovits [17]. We use the notation of Section 2 here. For $0 \le x \le 1$ we let

$$h_t(x) = \max\{\pi^{(t)}\omega - x : \omega \in [0,1]^{|\Omega'|}, \pi\omega = x\} \tag{13}$$

where now $\pi^{(t)}, \pi$ are treated as vectors of length $|\Omega'|$. Thus h_t is a continuous version of the variational distance between $\pi^{(t)}$ and π. For $0 \le \mu \le 1/2$ we let the μ-conductance of the chain be

$$\Phi_\mu = \min_{\mu \le \pi(I) \le 1/2} \left\{ \frac{\sum_{i \in I, j \notin I} \pi_i p_{ij}}{\pi(I) - \mu} \right\}.$$

Lovász and Simonovits proved the following generalisation of Theorem 2:

Theorem 7 *Let* C=max$\{h_0(x) : x \in [0,\mu] \cup [1-\mu, 1]\}$. *Then*

$$h_t(x) \le C + \exp\{-\tfrac{1}{2}\Phi_\mu^2 t\}/\sqrt{\pi_0}.$$

□

We now proceed to bound C and Φ_μ where we let

$$\mu = e^{-M/(4\sqrt{m})} \tag{14}$$

which from (11) (with $d = 8/(9\sqrt{m})$) is an upper bound to the limiting probability of the walk being in a cell which meets the boundary of the ball D.

5.1 Bounds for π_0 and C

Suppose now that b_1 lies in a cell P_0 and that $\pi_0 = \pi(P_0)$. Then

$$\begin{aligned}\pi_0 &= \frac{\psi(P_0)}{\sum_{P\in\Omega'}\psi(P)} \\ &\geq \frac{e^{-M/(9\sqrt{m})}}{\binom{n}{m}N^m} \\ &\geq e^{-M/(8\sqrt{m})}. \end{aligned} \tag{15}$$

Suppose first that $0 \leq x \leq \mu$. Then from (13) and (15) we have

$$\begin{aligned}h_0(x) &= \frac{x}{\pi_0} - x \\ &\leq \mu e^{M/(8\sqrt{m})} \\ &= e^{-M/(8\sqrt{m})}.\end{aligned}$$

If $1-\mu \leq x \leq 1$ then (13) implies that $h_0(x) \leq \mu$ and so

$$C \leq e^{-M/(8\sqrt{m})}. \tag{16}$$

Lower bound for Φ_μ

Suppose now that $S \subseteq \Omega'$ and $\mu \leq \pi(S) \leq 1/2$. Let $\hat{S} = \{P \in S : P \subseteq D\}$. Next let $R = D \cap (\bigcup_{P\in S} P)$ and $W = \partial R \setminus \partial D$. Then

$$\begin{aligned}\Phi_S &= \frac{\sum_{(P,P')\in(S:\bar{S})}\pi(P)\lambda(P,P')}{\pi(S)-\mu} \\ &\geq \frac{\sum_{(P,P')\in(S:\bar{S})}\pi(P)\lambda(P,P')}{\pi(\hat{S})}\end{aligned}$$

since $\pi(S) \leq \pi(\hat{S}) + \mu$. Applying the reasoning of the previous section to D we obtain

$$\begin{aligned}\Phi_S &\geq \frac{e^{-2/5}}{4Nm^{3/2}} \frac{\int_W \phi(x)dx}{\int_R \phi(x)dx} \\ &\geq \frac{e^{-2/5}}{4Nm^{3/2}} \cdot \frac{2}{\text{diam }(D)} \\ &\geq \frac{1}{6mN}.\end{aligned}$$

and so

$$\Phi_\mu \geq \frac{1}{6mN}. \tag{17}$$

Applying Theorem 7, (11), (16) and (17) we see that after T steps of METROPLEX the probability we are at a distance greater than d from b_2 is at most $C + 1/(2m) < 1$. The final part of the proof that we can identify a basic variable goes through as we have scaled d by $1/(10\sqrt{m})$. (We previously solved LP(b') with $\|b'\|_1 = \frac{1}{2}$ but now $\|b_2\|_1 = 1/(20\sqrt{m})$.) Note that identifying a basic variable is equivalent to identifying a facet of Q containing v_2. The remainder of the path from v_1 to v_2 will be restricted to this facet. In all we have to identify m such facets. The total number of pivots required is therefore $O(Tm)$ and Theorem 6 follows.

6 Concluding remarks

Our results on random generation extend those of Aldous and Broder for trees. They also differ in one important respect, in that we have showed that the "conductance" approach succeeds for the most natural random walk on these objects. The challenge of generalising these results to arbitrary matroids seems to us most likely to be achieved this way. (Though our proofs give no clue as to how this might be done.)

The time estimate for our linear programming algorithm is clearly rather large. We have not attempted to make its time bound as tight as possible, since our result is merely intended to demonstrate the existence of a polynomial time "simplex" algorithm for this class of problems. There are, of course, worst case strongly polynomial algorithms for these problems [24], but none resembles the simplex method except in very special cases. Note that since total unimodularity is preserved under duality, our algorithm may also be regarded as a primal simplex method in which cost-increasing pivots are allowed with low probability. We believe our result on totally unimodular linear programs gives the most general problem class for which a strongly polynomial time variant of the simplex method is known to exist. We include all network problems, for example, where the usual variants of the primal simplex method are known not to be even polynomial [26, 8]. The best (exponential) bound here on the number of pivots which is independent of the size of the numbers is due to Tarjan [25], who also gives a polynomial "simplex" algorithm allowing cost-increasing pivots, but with only a weakly polynomial bound on the number of pivots. It must also be observed that Tarjan's algorithms use much of the sophisticated machinery of non-simplex network flow techniques. Thus his methods depart from the spirit of the simplex method in a way which ours do not. The obvious challenge is to generalise our results to linear programs in which the A matrix has bounded entries or, more ambitiously, to arbitrary linear programs. Another issue is to what extent our linear programming algorithm can be de-randomised.

We have given a bound on the combinatorial diameter of polyhedra defined by totally unimodular constraint systems. Thus these polyhedra satisfy the so-called *polynomial diameter conjecture* [15] which is a weakening of the famous Hirsch conjecture. Again, as far as we know, this is one of only two rich classes of polyhedra for which the polynomial diameter conjecture is currently known to be true, the other being polytopes with 0-1 vertices only, Naddef [20].

We should finally point out that Feder and Mihail [12] have extended Theorem 1 to a wider class of matroids, called *balanced matroids.* Unfortunately, their proof does not seem to help in extending Theorems 5 or 6.

References

[1] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, Séminaire de Probabilités XVIII, 1981–82, Springer Lecture Note in Mathematics 986, pp. 243–297.

[2] D. Aldous, *The random walk constructon of uniform spanning trees and uniform labelled trees*, SIAM J. of Disc. Math. **3** (1990) 450–465.

[3] D. Applegate and R. Kannan, *Sampling and integration of near log-concave functions* Proc. 23rd ACM Symposium on Theory of Computing, 1991, pp. 156-163.

[4] J. W. Archbold, *Algebra*, Pitman, London, 1958.

[5] A. Z. Broder, *How hard is it to marry at random ? (On the approximation of the permanent)* Proc. 18th ACM Symposium on Theory of Computing, 1986, pp. 50-58.

[6] A. Z. Broder, *Generating random spanning trees*, Proc. 30th Annual Symposium on Foundations of Computer Science, 1989, pp. 442-447.

[7] Y. D. Burago and V. A. Zalgaller, Geometric inequalities, Springer-Verlag, Berlin, 1980.

[8] W. H. Cunningham, *Theoretical properties of the network simplex method*, Mathematics of Operations Research **4**, 196–208.

[9] M. E. Dyer, *Approximation of mixed volumes*, in preparation.

[10] M. E. Dyer and A. M. Frieze, *Computing the volume of convex bodies: a case where randomness provably helps*, Proceedings of AMS Short Course on Probability and Combinatorics (B. Bollobás, Ed.), to appear.

[11] M. E. Dyer, A. M. Frieze and R. Kannan, *A random polynomial time algorithm for approximating the volume of convex bodies*, Journal of the A. C. M. **38** (1991), 1–17.

[12] T.Feder and M.Mihail, *Balanced matroids*, Proc. of the 24'th Annual ACM Symposium on Theory of Computing, 1992, pp26–38.

[13] M. R. Jerrum and A. J. Sinclair, *Approximating the permanent*, SIAM Journal on Computing **18** (1989) 1149-1178.

[14] M. R. Jerrum and A. J. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, Department of Computer Science, Edinburgh University, November 1989.

[15] G. Kalai, *The diameter of graphs of convex polytopes and f-vector theory*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, to apppear.

[16] A. Karzanov and L. G. Khachyan, *On the conductance of order Markov chains*, Technical Report DCS TR 268, Rutgers University, 1990.

[17] L. Lovász and M. Simonovits, *The mixing rate of Markov chains, an isoperimetric inequality and computing the volume*, Proc. 31st Annual Symposium on Foundations of Computer Science, 1990, pp. 364–355.

[18] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *Equation of state calculation by fast computing machines*, J. Chemical Physics **21** (1953) 1087–1091.

[19] M. Mihail and P. Winkler, *On the number of Euler orientations of a graph* Proc. Third Annual ACM-SIAM Symposium on Discrete Algorithms, 1992, pp138–145.

[20] D.Naddef, *The Hirsch conjecture is true for 0-1 polytopes* , Mathematical Programming 45 (1989) 109-110.

[21] A. Schrijver, *The theory of linear and integer programming*, John Wiley, Chichester, 1986.

[22] A. J. Sinclair and M. R. Jerrum, *Approximate counting, uniform generation and rapidly mixing Markov chains*, Information and Computation **82** (1989) 93–133.

[23] R. M. Stanley, *Two combinatorial applications of the Aleksandrov-Fenchel inequalities*, J. Combinatorial Theory A **31** (1981) 56–65.

[24] E. Tardos, *A strongly polynomial algorithm to solve combinatorial linear programs*, Operations Research **34** (1986) 250–256.

[25] R. E. Tarjan, *Efficiency of the primal network simplex method algorithm for the minimum-cost circulation problem*, Mathematics of Operations Research **16** (1991), 272–291.

[26] N. Zadeh, *A bad network for the simplex method and other minimum cost flow algorithms*, Mathematical Programming **5**, 255–266.