Application of Convex Sampling to Optimization and Contingency Table Generation/Counting

John Mount

May 1995

CMU-CS-95-152

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Thesis Committee:

Ravi Kannan, Chair
Alan Frieze
Garry Miller
Sridhar Tayur, GSIA

This research was partially supported by National Science Foundation grant CCR 9208597.

Abstract

The optimization problem of minimizing a linear objective function over a general convex body given only by a weak membership oracle is a central problem in convex optimization. Traditional methods require the (expensive) construction of separating relations (cuts) or gradient information (which is often expensive). We demonstrate how a sampling procedure can be used as the central routine in a randomized polynomial time algorithm for approximately minimizing a linear objective function over an up-monotone convex set presented by a membership oracle. The sampling procedure is a Markov chain that uses only local membership tests. We further demonstrate a direct application of this technique to an important stochastic optimization problem called "component commonality."

A second application of the above sampling scheme is a statistical hypothesis testing procedure involving *contingency tables*. The test involves counting contingency tables (matrices of non-negative integers) obeying given row and column constraints which is known to be $\sharp P$ hard. Under fairly natural assumptions the Markov technique yields an effective procedure for approximating, to any desired accuracy, the necessary counts. We also develop a powerful exact counting scheme, for fixed dimension, and use it to validate the random technique.

Acknowledgements

I would like to thank everybody who helped me with my thesis work. In particular: my parents, for their faith and encouragement; Nina Zumel, for her patience and support; Ravi Kannan, for guidance and sharing his intuition; and Peter Stout, for use of his **WAX** system .

Finally I would like to thank Ahab F. Sluggo of Transarc for always being available with vivid demonstrations of the need for fault tolerance in any large or distributed computation.

Contents

Chapter 1

Introduction

This thesis applies samplers based on rapidly mixing Markov chains to both convex optimization and counting contingency tables. A sampler, S, is a randomized algorithm (or random variable) that, given a description of a feasible set and probability measure μ on the feasible set, generates a feasible element such that for any non-negligible μ-measurable set V, $Pr[S \in V] \approx \mu(V)$. This thesis is primarily concerned with samplers that generate points from convex bodies in $\mathbf{R}^n$ which are presented by membership oracles.[1] [23, 21, 5, 40]

The first part of this thesis is joint work with Ravi Kannan and Sridhar Tayur and is a randomized polynomial time algorithm to approximately minimize a linear function c over an up-monotone convex set, K, (i.e., a convex set with the property that if x belongs to the set and y is component wise greater than or equal to x, then y belongs to the set also) in the positive orthant given by a membership oracle.[39] By "membership oracle" we mean a black box procedure that can tell us if $x \in K$. It does not supply any additional structural information. It is also assumed that an initial point $x^f \in K$ is known. This model is motivated by a stochastic optimization problem called *component commonality* [37] where membership can be tested by performing a numerical integration, but separating hyperplanes or even gradient directions are expensive to obtain.

Our approach is to optimize using only local information. We present a non-uniform sampler for $\mathbf{R}^n$ that obeys a probability distribution that grows geometrically in the direction of the objective function and falls off geometrically in distance away from the convex body. The central technique is to design the

[1] Actually the intersection of a very fine lattice in $\mathbf{R}^n$ with a convex body.

distribution to be smooth enough that the sampler can converge in polynomial time.

The second application of this thesis is joint work with Martin Dyer and Ravi Kannan and is a near-uniform sampler and an approximate counting scheme for two way contingency tables that simultaneously obey row and column constraints. A two way contingency table is a matrix of non-negative integers. These tables are often used to represent how many individuals have pairs of attributes (i.e. rows are indexed by one attribute, columns by the other) in a population.[2] Diaconis and Efron [16] have proposed a significance test and a model fit that can be performed with the aid of a near-uniform sampler for tables that obey given row and column constraints.

We show how, given certain assumptions, simple geometry can be used to convert a near-uniform convex body sampler to a near-uniform contingency table sampler. That is: given certain easily met conditions on the row and column totals our continuous techniques can solve this discrete problem. An approximate counting scheme is designed using this sampler.

We also develop, with Bernd Sturmfels, an effective exact counting scheme for fixed dimension contingency tables. This scheme helped validate the random results and allowed us to analyze the asymptotic accuracy of some estimation schemes in the literature.

1.1 Convex optimization.

In principle any convex optimization problem can by solved in polynomial time by the Ellipsoid Algorithm. One can also see Vaidya[53] for a state of the art general convex programming algorithm. However, these approaches all require a separation oracle. A technique described by Yudin and Nemirovskii can be used to convert a membership oracle to a separation oracle [33], but this is quite expensive. Previous efforts to solve component commonality have used non-linear programming methods, but these require that gradients be calculated, which can be as expensive as computing separating cuts.

We solve the following more general problem: optimize a linear function over an up-monotone set presented by a membership oracle. A set $K \subseteq \mathbf{R}^n$ is called up-monotone if $x \in K$ and $y \geq x \rightarrow y \in K$. This property holds for linear programs of the form $Ax \geq b$ when the matrix A is non-negative. Many supply and covering problems have this form. We do not require that K be bounded but require, with out loss of generality, that our objective function, c, be strictly positive and K be contained in the non-negative orthant of $\mathbf{R}^n$. Thus for any d we have the set $K \bigcap \{x \in \mathbf{R}^n \mid c \cdot x \leq d\}$ is bounded.

The component commonality problem is a stochastic or chance constraint optimization problem. Good sources for this problem are [54] and [37]. The informal description is as follows. Every month a factory receives orders for its m products. The total orders for a month are considered a vector in $\mathbf{R}^m$ (we will not deal with the integral nature of this problem at this time) distributed according to some log-concave density.[2] We assume that we can evaluate $F(d)$ at any point $d \in \mathbf{R}^m$. This is enough to allow us to numericly integrate F over various regions. We also have a $n \times m$ non-negative matrix A in which we interpret $A_{a,b}$ as the amount of component a used in the production of one unit of product b. Our goal is to find a vector $x \in \mathbf{R}^n$ that is a stocking level of components so we have $Ad \leq x$ with probability at least α when d is drawn from the order distribution. That is if we stocked our warehouse according to the plan x we would be able to fill all our orders with probability at least α. It should be clean that component commonality is a special case of up-monotone optimization.

1.2 Contingency tables.

Diaconis and Efron [16] have developed a method of modeling distributions that allows one to make some qualitative and quantitative statements about an unknown distribution. This technique is potentially much more useful than the usual task of proving the unknown distribution is not various test distributions. Their scheme requires several significance tests which come down to counting contingency tables. Unfortunately the counting problem is difficult ($\sharp P$ hard) even for $2 \times n$ tables.

An example (adapted from [17]) is in order (this is example should not be construed as being the only application).

The table 1.1 is from Snee [49]. A uniform generator of contingency tables could be used to implement many different statistical tests, we give an example here chosen for simplicity (rather than power or generality). Suppose a census

[2] A density, F, is log-concave if $\log(F)$ is concave. A density, G, is concave if $G(\lambda x + (1-\lambda)y) \geq \lambda G(x) + (1-\lambda)G(y)$ for all x, y and $0 \leq \lambda \leq 1$. So F is log-concave if and only if $F(\lambda x + (1-\lambda)y) \geq F(x)^{\lambda} F(y)^{1-\lambda}$ with x, y, λ as above. Any positive concave function is log-concave.

Eye Color	Hair Color				Total
	Black	Brunette	Red	Blond	
Brown	68	119	26	7	220
Blue	20	84	17	94	215
Hazel	15	54	14	10	93
Green	5	29	14	16	64
Total	108	286	71	127	592

Table 1.1: Eye v.s. Hair Color, observed.

Eye Color	Hair Color				Total
	Black	Brunette	Red	Blond	
Brown	40	106	27	47	220
Blue	39	104	26	46	215
Hazel	17	45	11	20	93
Green	12	31	7	14	64
Total	108	286	71	127	592

Table 1.2: Eye v.s. Hair Color, nearly independent.

were performed where 50 different statistics (hair color, eye color, income level, ...) each divided into discrete categories (Black, Brunette, Red, Blond; Brown, Blue, Hazel, Green; 0 to 6999 krona, 7000 to 13999 krona ...) were collected on 592 people. A plausible task for a statistician would be to identify pairs of traits that support some meaningful relationship. This could be the first step in an epidemiological study. One technique would be to examine the $\binom{50}{2} = 1225$ two way tables. Table 1.1 could be one such pairing. The statistician does not wish to look at 1225 two way tables- so an automatic technique of identifying interesting ones is required.

Let X be m by n a table and let $r_i = \sum_{j=1}^{n} X_{i,j}$ and $c_j = \sum_{i=1}^{m} X_{i,j}$. X would be considered uninteresting if the variables were independent, or equivalently $X_{i,j} \approx r_i c_j/592$. A nearly independent table would look like table 1.2.

Comparing these two tables one might conclude that people with brown eyes and blond hair occur less often in the observed table than in the nearly independent table. The question is not if we have discovered a relationship between hair and eye colors in the table at hand, but if the relationship observed in this table is a likely due to sampling error. One could compute the χ^2 statistic of the observed table (which is a measure of departure from independence) and get a value of 138.29. The question then becomes is 138.29 a typical or atypical amount of deviation from independence? One method put forward by Diaconis and Efron to determine this is to compute the significance level of the observed table with respect to the uniform distribution. This is defined as the ratio of the number of integer matrices matching our row and column totals that have a χ^2 statistic of at least 138.29 to the total number of integer matrices matching our row and column totals. That would be to say "of all the possible tables with the same disjoint distributions (ie. with the same r_i and c_j) of hair and eye colors what proportion have $\chi^2 \leq 138.29$".

For the table in question asymptotic and volume techniques gave "about 10%"[16] and later randomized techniques gave 16.3% (with no confidence interval)[20]. My current randomized simulation work shows that the true value is between 15.3 and 15.7 with a confidence level of over 99.7%.[3] In any case, the result is that about $1/6$ of all possible tables with the same row and column sums look more independent than the observed one. So the relationship between eye color and hair color observed in the table is not very strong and would not be a good candidate for further investigation.[4] The two way tables could be ranked according to how atypical they are (the significance of their χ^2) and the human statistician could then be presented the top few dozen such tables.[5]

With a counting technique (and some Bayesian priors) much more sophisticated analysis can be performed.

[3] This should illustrate the difficulty encountered analyzing tables as small as 4×4.

[4] It is important to note that we are not saying that the relationship doesn't exist or is statistically insignificant. It is also important to remember that when the number of variables is commensurate with the sample population that one would expect many statistically significant but meaningless cross correlations.

[5] They could then discover a relationship like "unusually few people who juggle knives drink every day" indicating either few people partake in both or the combination is fatal.

1.3 Basic Markov chains.

Here we will define the Markov chain terminology we are using in this thesis. We will consider a Markov chain to be a finite directed graph $G = (V, E)$ and a transition function $P : V \times V \to \mathbf{R}$ such that

$$\begin{array}{ll} P(x,y) \geq 0 & \forall x, y \in V \\ P(x,y) > 0 & \forall (x,y) \in E \\ P(x,y) = 0 & \forall (x,y) \notin E \\ \sum_{y \in V} P(x,y) = 1 & \forall x \in V \end{array}.$$

Note that this definition requires $(x, x) \in E$ if $\sum_{y \in V\ y \neq x} P(x, y) < 1$. We interpret $P(x, y)$ as the probability that the Markov chain will be in state y at time $t + 1$ given that it is in state x at time t (or $P[X_{t+1} = y|X_t = x]$). We will also consider our Markov chain as a linear system where X_t will be a vector in $\mathbf{R}^V$ and by an abuse of notation we take P to be the $|V| \times |V|$ matrix such that $P_{x,y} = P(x, y)$. The idea is that if X_t is a vector such that $X_t \geq 0$ and $1 \cdot X_t = 1$ we can interpret $(X_t)_v$ as $P[X_t = v]$, the probability that our Markov chain is in state v at time t. The Markov chain then progresses under the simple relation

$$X_{t+1} = X_t P.$$

As one would expect we are immediately concerned about the eigenvalues and eigenvectors of these systems.

We will not treat Markov chains in their full generality, but discuss only the "nice" chains used in this thesis. For a much more general treatment one should refer to [12]. "Nice" involves some technical definitions from the theory of Markov chains, but we will define all the terms used in this section. The chains we will use will be chosen to be irreducible, aperiodic and time-reversible. Each of these terms has an interpretation in terms of stochastic processes, graph theory and linear operators. To motivate the definitions we will point out all three interpretations where appropriate.

A finite Markov chain is called *irreducible* if $P[X_t = y \text{ infinitely often}] = 1$ for all all $y \in V$. This means that the chain has no transient states (states that can not be reached from other sets of states). If fact it is exactly equivalent to the underlying graph G being strongly connected (for every $a, b \in V$ there exists a directed path of edges in E from a to b). This condition is stronger than saying that the matrix P is an irreducible matrix but is equivalent to $P^{|V|} > 0$.

An irreducible Markov chain is stationary or *aperiodic* if $\forall x, y \in V \quad \lim_{t\to\infty} P[X_t = y|X_0 = x]$ exists. This means that for any k X_t and X_{t+k} are distributed identically as $t \to \infty$. It is easy to see that for an irreducible aperiodic Markov chain the above limits must independent of x and it makes sense to write $\lim_{t\to\infty} P[X_t = y]$. An irreducible Markov chain is aperiodic if and only if the least common divisor of the lengths of all cycles in G (counting self loops as cycles of length 1) is equal to 1. The linear operator interpretation of aperiodic is that all eigenvalues of the operator are real (and ≥ -1). If the chain is irreducible and aperiodic then the eigenvalue 1 occurs with multiplicity exactly one and the unique eigenvector π such that $1 \cdot \pi = 1$ corresponding to this eigenvalue is called the stationary distribution. We then have $\forall y \in V \ \pi_y = \lim_{t\to\infty} P[X_t = y]$ and $\pi_y > 0$ for all $y \in V$.

An irreducible aperiodic Markov chain is *time reversible* if $\forall (x, y) \in E \quad \pi(x)P(x, y) = \pi(y)P(y, x)$. Time reversibility has a number of interpretations. The most obvious is that for every edge $(x, y) \in E$ we have $(y, x) \in E$ and $\lim_{t\to\infty} P[X_t = x \wedge X_{t+1} = y] = \lim_{t\to\infty} P[X_t = y \wedge X_{t+1} = x]$, or each edge direction is equally likely in the limit. In terms of linear operators if a Markov chain is irreducible, aperiodic and time reversible then the matrix DPD^{-1} is symmetric where D is the $|V| \times |V|$ diagonal matrix such that $D_{x,x} = \sqrt{\pi(x)}$. Thus P is in fact similar to a diagonal matrix.[6]

From now on we will assume our Markov chain is irreducible, aperiodic and time reversible. We have for such chains if q is any vector that obeys the time reversal equation $\forall x, y \in V \quad P(x, y)q_x = P(y, x)q_y$ then $q = \lambda\pi$ for some scalar λ. This property is very useful in designing Markov chains with specific stationary distributions.

If G is a strongly connected symmetric graph ($(x, y) \in E \leftrightarrow (y, x) \in E$) there is a very powerful general method of designing a Markov chain by choosing the transition probabilities according to the "Metropolis filter." Suppose we have a positive function F and we wish that the stationary distribution π of our Markov

[6] It is interesting to notice that time reversibility taken alone does not have such a simple interpretation as the 3 state Markov chain with transition matrix

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

is not similar to any symmetric or diagonal matrix.

chain has $\pi_x = F(x)/(\sum_{y \in V} F(y))\ \forall x \in V$. For $x \in V$ let $N(x)$ be the set of all $y \in V,\ y \neq x,\ (x, y) \in E$. Let Υ be $\max_{x \in V} |N(x)|$, the largest degree of any vertex in G and define

$$P(x,y) = \begin{cases} \frac{1}{2\Upsilon}\min(1, \frac{F(y)}{F(x)}) & y \in N(x) \\ 1 - \sum_{z \in N(x)} P(x,z) & x = y \\ 0 & \text{otherwise} \end{cases} . \tag{1.1}$$

The Markov chain with transition matrix P is irreducible, aperiodic and time-reversible. The irreduciblity come directly from G. Aperiodicity because $P(x,x) > 0$ for all x. The chain is time-revisable because if $x \neq y$ and $P(x,y) \neq 0$ then it time-reverses with F:

$$\begin{aligned} P(x,y)F(x) = \tfrac{1}{2n}\min(1, \tfrac{F(y)}{F(x)})F(x) = \\ \tfrac{1}{2n}\min(F(x), F(y)) = \\ \tfrac{1}{2n}\min(1, \tfrac{F(x)}{F(y)})F(y) = P(y,x)F(y). \end{aligned}$$

Furthermore this is enough to show that the unique stationary distribution π is such that $\pi_x = F(x)/(\sum_{y \in V} F(y))\ \forall x \in V$. It is very important that F enters into the transition probability equation as a ratio- so one never actually has to pre-compute the normalizing factor $\sum_{y \in V} F(y)$ as this factor is often the count or volume that one needs the Markov chain to estimate.

We want our Markov chain to be *rapidly mixing*.[3] To explain this property we pick a measure of similarity between distributions on V. Let p_A and p_B be two probability distributions on V. The distance from p_A to p_B can be defined a number of ways:

$\sqrt[k]{\sum_{y \in V} \lvert p_A(y) - p_B(y)\rvert^k}$	l_k distance
$\max_{y \in V} \lvert p_A(y) - p_B(y)\rvert$	l_∞ distance
$\sum_{y \in V} \frac{(p_A(x) - p_B(x))^2}{p_A(x)}$	chi-sq distance with respect to p_A.

We use l_1 or *variational* distance because it is the distance used in many of the papers in the literature and has a slightly simpler interpretation. Let $d(p_A, p_B)$ denote one of the above distance measures. Let p_t be the distribution at time time ($(p_t)_y \stackrel{\text{def}}{=} P[X_t = y]$) it is well known that for any Markov chain as above that there is a constant $\chi < 1$ such that $d(\pi, p_t) \leq \chi^t f(\pi, p_0)$ where f is some function of π and p_0. For the chi-sq distance we can take $f(\pi, p_0) = \sum_{y \in V} \frac{(\pi(x) - p_0(x))^2}{\pi(x)}$ (which is again the chi-sq distance, simplifying subsequent analysis).[27] A Markov chain is rapidly mixing if $1/(1-\chi)$ is smaller than some polynomial in parameters we care about.

We will be looking at chains whose states are lattices in $\mathbf{R}^n$ intersected with bodies of diameter d whose transitions correspond "King's moves." In this case rapidly mixing will mean that $1/(1-\chi)$ is bounded above by a polynomial in n and d. It would be nice if rapidly mixing had a simple definition such as "polynomial in $\log(|V|)$" but it is easy to show that Markov chain on the points in $x \in Z^n_{\geq 0}$ with $x_i < d$ has d^n states and has $1/(1-\chi) \in O(nd^2)$ which is not polynomial in $\log(|V|) = n \log d$. We wish to consider this chain rapidly mixing, so we are forced to abandon any such simple definition of rapidly mixing. It is important to notice that the graph-diameter of such a Markov chain is $O(nd)$ thus our chain is rapidly mixing in the sense it mixes in time polynomial in the diameter of the chain (instead of the much weaker property of mixing in time polynomial in the number of states).

1.4 Some geometric Markov chains.

Here we quickly outline some of the typical Markov chains used to generate points from convex regions. All the Markov chains in this section are designed by picking a finite set of points from inside a convex region as the set of states. The underlying graph $G = (V, E)$ of the Markov chain is then chosen as above and such that for each state it is easy to generate a neighbor with probability $P(x, y)$.

1.4.1 King's moves

The idea of "King's moves" (used in both the optimization problem and for contingency tables) is to have the set of states of the Markov chain be some affine translate of the standard integer lattice. The edges in this Markov chain are the states that differ by one unit in exactly one coordinate. These are not the moves that the King makes in chess (which would allow a unit difference in any number of coordinates) but for lack of a better name have been called "King's moves". In this scheme each state in the Markov chain is identified with the cube consisting of points closest to the state. States roughly correspond to volume elements and edges correspond to facets of cubes (or area elements). This correspondence of states and transitions to simple geometric objects is of central importance in analyzing the behavior of these Markov chains.

1.4.2 Ball moves

Ball moves, and the related "Normal moves", also have a strong geometric interpretation. In ball moves each point in $\mathbf{R}^n$ is a state in the Markov chain (actually one is forced to discretize space on a very fine grid, but for all intents and purposes one deals with a Markov chain on $\mathbf{R}^n$) and transitions are vectors drawn uniformly from a ball (or in the normal case vectors drawn from $\mathbf{R}^n$ using the normal density function). A transition corresponds to adding a vector to the current state to get a new state. This is very much like a discrete time Brownian process. In this treatment one rarely discusses individual states but instead reasons about, measurable, collections of states. The probability of moving from a given state into a set of states is then the proportion of the measure of the step set (ball or normal) that is inside the destination set.

1.4.3 Gibbs sampler

The Gibbs sampler is a classic method used in statistics. What one does is either run through coordinate directions in a cyclic order or pick them at random. When one has coordinate direction one then generates a point on the line passing through the current point in the given coordinate direction. The idea behind this method is to mix perfectly with respect to one variable at a time. The hope is that in relatively few runs through the variables that the process will mix. Gibbs has the advantage that the moves, which seem quite powerful, are often quite easy to implement. This method has not been well characterized as a Markov chain. Most proofs of mixing time prove only that Gibbs is not much worse than King's moves. These moves were also called "Rooks moves" in Applegate's thesis.

1.4.4 Hit and run

Hit and run is the name of an important class of stochastic techniques.[55, 56, 48, 9] Hit and run is used in non-linear programing and it is of interest to this thesis that it was recently shown to be rapidly mixing by Martin Dyer and Leen Stougie.[25] Then method is natural but has the weakness that no one has identified a simple geometric quantity that predicts the mixing rate like isoperimetry does for King's or ball moves. In Hit and Run one picks a direction r uniformly from the surface of the unit ball and then in one step generates a point distributed according to F (the density function we are trying to achieve) restricted to the ray emanating in the direction of the ray. It is interesting to note that this chain is time reversible without any application of the Metropolis filter, just by the fact that a ray r and $-r$ are generated with equal probability.

1.5 Counting/computing volume.

If one can count lattice points one can compute volumes and the converse is often true. The following is taken from Gruber and Lekkerkerker [34].

A lattice polytope is a polytope with all integral vertices. For a lattice polytope P and natural number k we have

$$\sharp(Z^n \cap kP) = \sum_{i=0}^{n} L_i(P)k^i$$

where the L_i are rational numbers depending only on P and n. Let f be any function from the set of all lattice polytopes. We say that f is *unimodular invariant* if for any lattice polytope P and unimodular integral transformation U and integral vector u:

$$f(UP + u) = f(P).$$

We say that f is *additive* if for all lattice polytopes P_1, P_2 such that $P_1 \cup P_2$ is again a lattice polytope then

$$f(P_1 \cup P_2) + f(P_1 \cap P_2) = f(P_1) + f(P_2).$$

A theorem of Betke's shows if f is unimodular invariant and additive (volume is such a function) then $f(P) = \sum_{i=0}^{n} \lambda_i L_i(P)$ for all P where L_i are as above and λ_i are scalars depending only on f (independent of P). Thus the ability to count lattice points is in some sense universal for a large class of invariants of integral polytopes.

The volume of a convex body differs from the number of lattice points in the convex body by an amount proportional to the number of lattice points whose Vornoi cells are not completely covered by the body or its compliment. Often, as is the case with contingency tables, the body contains a reasonable sized copy of a n-cube and therefore the volume is a good approximation of the number of lattice points in the body. This allows one to generate a lattice point from a distribution arbitrarily close to uniform (by a rejection sampling technique) and then generate estimates for the number of lattice points with any desired level of accuracy.

Chapter 2

Up Monotone Optimization

2.1 Membership model.

We are interested in minimizing a linear function, c, over a convex body $K \subset \mathbf{R}^n$ presented by a membership oracle. This is the most general form of a linear convex optimization problem. Under the "oracle" model [33] the convex body K is presented as a tape containing (n, x, r, R) where n is dimension of space we are working in, $x \in \mathbf{Q}^n$ is a rational point in K, $r > 0$ is such that $B_r(x) \subseteq K$ and R is such that $K \subseteq B_R(x)$, where $B_r(x)$ denotes the ball of radius r centered at x. It is also assumed that one can determine if $y \in K$ for any $y \in \mathbf{R}^n$. In this model the only optimization algorithms known to run in polynomial time are relatives of the classic ellipsoid method.

A large problem with these method for convex optimization are that separating inequalities are required to run the algorithm. Given a point $x \in \mathbf{R}^n,\ x \notin K$ a separating inequality for K and x is a vector a and number b such that $a \cdot x > b$ and $\forall y \in K\ a \cdot y \leq b$. The fact that any two closed disjoint convex sets can be separated by such a linear relation is one of the central properties of convexity. Separating relations can be derived from a membership oracle [33, 43] but the reduction is complicated and expensive.

2.2 Algorithm outline and time bounds.

Using the membership oracle, we approximately minimize a linear function over an up-monotone convex feasible set in the positive orthant as follows. We may assume a suitable upper bound on the variables so we can enclose this feasible region in a rectangle (in n dimensions). At the heart of our approach is a positive real-valued logarithmically concave function F on the rectangle with the following properties: (1) the integral of F over a region consisting of near-optimal solutions is at least a constant fraction of the integral of F over the whole feasible set, and (2) the integral of F over the feasible set is at least a constant fraction of the integral of F over the entire rectangle. Thus, if we pick a "random sample" from the rectangle with probability density proportional to F (we refer to this throughout as "sampling according to F"), we would get a near-optimal solution with constant probability; this probability can be boosted by repeated sampling. Our algorithm, then, is simply a choice of F (determined by two parameters α, a damping factor favoring feasible points, and β, a bias favoring points with better objective values) and a method to obtain a sample according to F. We show that a certain biased random walk (on the uniform grid of size δ, to be determined later), starting from a feasible solution (x^f), is indeed able to pick a random sample from the feasible set with probability (approximately) proportional to F. While it is relatively easy to argue that in the steady state, this random walk picks a sample with density proportional to F, it is nontrivial to show that this steady state is approached in a polynomial number of steps. To accomplish this central result, we draw on recently developed results in the theory of rapidly mixing Markov Chains as well as on random walks in convex sets [23] , [5]. The latter paper gives a technique for sampling from log-concave distributions which we use here, although, we have tried to make this description self-contained by giving as many details as possible. Our random walk can be executed with only local knowledge of F as well as a membership (not a separation) oracle for the feasible set.

Given an instance of the problem (a membership oracle for K and objective function $c > 0$), $\epsilon > 0$ (relative error), $\kappa > 0$ (failure probability) and an initial feasible point x^f, the algorithm succeeds with probability at least $1 - \kappa$ in finding a $q^{alg} \in \mathbf{R}^n$ which is feasible and such that $c \cdot q^{alg} \leq (1+\epsilon)(c \cdot q^{opt})$. Rather than come within ϵ of the optimal in one long random walk, we develop an adaptive algorithm which improves the feasible solution in stages (by iteratively refining the gap between a known feasible solution and a probabilistic point-wise lower bound lower bound L on optimal cost). This "staging" of the algorithm decreases the run-time's dependence on ϵ.

Each stage begins with a feasible solution x^f and a probabilistic "lower bound" L where if there is a feasible x with $c \cdot x < L$, then the algorithm has failed. (We will of course ensure that the probability of failure is low.) We refer to $c \cdot x^f - L$

as the "gap" (at the beginning of the stage). At the end of the stage, we have a new lower bound and a new feasible solution; we ensure that the gap at the end of a stage is at most 1/2 of the gap at the beginning. We stop when the gap is at most ϵ times the value of the cost lower bound. The algorithm is described in detail in figure 2.4 and justified in section 2.7, but we give here a short verbal description.

Each stage proceeds as follows: the feasible set is enclosed in a rectangular solid. We devise a log-concave function F on the rectangle with the two properties described above. [The function F has two components: a "penalty" (called the gauge function in what follows) for going out of the feasible set which increases as we become "more infeasible" and a bias (drift) which favors low objective function value. In some vague sense, this is similar to Lagrangian relaxation with both feasibility and optimality represented by one function.]

We then discretize by dividing the rectangle into small cubes. We perform a random walk on the cubes with transition probabilities depending on F. It will be easy to see that the steady state probabilities of this random walk will be proportional to F. We will also show fast convergence to the steady state, so that after a polynomial number of steps, we are "close" to the steady state probabilities.

After doing the random walk for this number of steps, one of the following two scenarios occurs:

(i) We have found a feasible solution whose value cuts down the gap by a factor of at least 1/2. In this case, we replace our old feasible solution by this and go to the next stage.

(ii) Otherwise, we have (probabilistic) proof of a greater lower bound and we go to the next stage with this new lower bound (again cutting the gap down by a factor of 1/2).

Although F has been devised accurately to have the desired properties, several errors are introduced in the sampling procedure. There are errors due to discretizing into small cubes, due to the inexact computation of the gauge function and due to the fact that the lower bounds are only probabilistic. The management of these errors is the main focus of section 2.5.

Our main result is two bounds on the running time of the algorithm. [The running time is bounded above by the minimum of the two.]

If ν is the ratio of the value of the given initial feasible solution to the optimal value, ϵ is the required relative error, $1-\kappa$ the required success probability and n, the dimension of the up-monotone convex feasible set K, the first bound is

$$O\left(\frac{n^7\left(\log\left(\frac{n}{\epsilon}\right)\right)^2\log\left(\frac{\log(\frac{1}{\epsilon})}{\kappa}\right)\log\left(\frac{\nu}{\epsilon}\right)}{\epsilon^2}\right).$$

If in addition, we are given $x^l > 0$ such that $\forall y \in K$, we have $x^l \leq y$ and an x^u such that there is an optimal solution z with $z \leq x^u$ (so we can replace K by $K \cap \{x : x \leq x^u\}$), then we also have a bound

$$O\left(n^5\left(\frac{\left\|x^u - x^l\right\|_\infty}{\epsilon \min_i x_i^l}\right)^2\left(\ln\left(\frac{n\left\|x^u - x^l\right\|_\infty}{\epsilon \min_i x_i^l}\right)\right)^2 \ln(\frac{1}{\kappa})\right).$$

The rest of the chapter is organized as follows. Section 2.3 notes that the CC problem falls into the framework of up-monotone convex sets and has some general remarks. Section 2.4 constructs the appropriate log concave and gauge functions that are to be used in each stage of the algorithm. Section 2.5 describes the random walk to be performed at any stage of the adaptive algorithm, and contains the analysis of the errors introduced due to discretization, gauge function approximation and walking for finite number of steps. In section 2.6, we find a bound on the spectral gap of the Markov chain, which allows us to use results on rapid mixing to provide a bound on the number of steps required in any stage. In section 2.7, we prove the adaptive algorithm's correctness and run time, and present two variants and prove their run times. Proofs for certain lemmas have been deferred until the end of the chapter.

2.3 The component commonality problem.

The problem was described in detail in the Introduction. If U denotes the matrix of the u_{ij} 's there, let $y = y(d) = Ud$. Under the assumption that $h(\cdot)$ is log-concave, it is easy to see that y has a log-concave density. Let D denote the density of the y. Let μ_D denote the corresponding measure. (So for any measurable set S, $\mu_D(S) = \int_S D$.) Then the feasible set K of stock-levels can be expressed as

$$K = \{x \in \mathbf{R}^n : \mu_D(\text{dom}(x)) \geq \gamma\}$$

where $\text{dom}(x)$ is $\{y : y \leq x\}$. It is easy to show that K is convex ([54]). It is also clearly up-monotone.

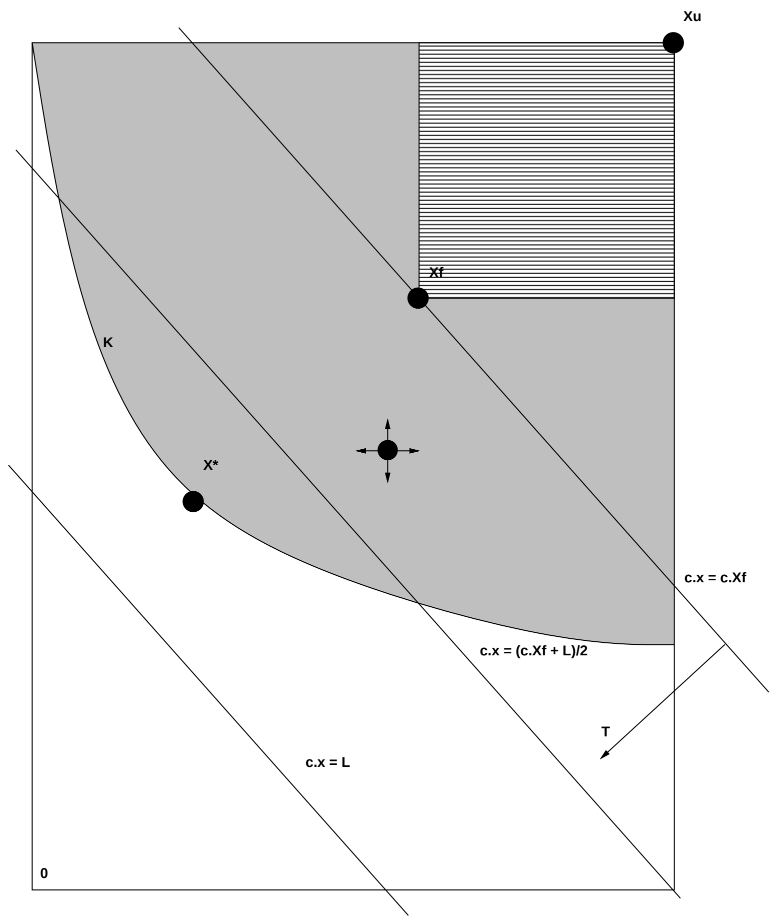


Figure 2.1: The basic geometry.

2.3.1 Why component commonality is convex.

It is clear that component commonality, as described in the introduction, is an up-monotone problem contained in the positive orthant. It is also easy in practice to find an initial feasible stocking. So if the feasible set is convex then it fits into our monotone optimization scheme. Let K is the feasible region of a component commonality problem its convexity follows from a theorem of Prékopa [54] but it is worthwhile to see how this is proven directly from the Brunn-Minkowski theorem.[10] The Brunn-Minkowski theorem is: if Vol is the standard volume function and A, B are two compact convex sets in $\mathbf{R}^n$ then the function

$$f : [0, 1] \to \mathbf{R} : \quad f(\lambda) = \text{Vol}(\lambda A + (1 - \lambda)B)^{1/n}$$

is concave.[10] If our density F was the indicator function of a convex set B (1 if $x \in B$, 0 otherwise) then the component commonality problem would be convex because then the probability that a stocking x is sufficient would be exactly $\text{Vol}(\{d \in \mathbf{R}^{m+} \mid Ad \leq x\}) / \text{Vol}(B)$ and we would have for any two feasible stockings x, y and $\lambda \in [0, 1]$

$$\begin{aligned}
&\text{Vol}(\{d \in \mathbf{R}^{m+} \mid Ad \leq \lambda x + (1-\lambda)y\})^{1/m} \\
&\quad \geq \text{Vol}(\lambda \{d \in \mathbf{R}^{m+} \mid Ad \leq x\} + (1-\lambda)\{d \in \mathbf{R}^{m+} \mid Ad \leq y\})^{1/m} \\
&\quad \geq \lambda \text{Vol}(\{d \in \mathbf{R}^{m+} \mid Ad \leq x\})^{1/m} + (1-\lambda) \text{Vol}(\{d \in \mathbf{R}^{m+} \mid Ad \leq y\})^{1/m} \\
&\quad \geq \min\left(\text{Vol}(\{d \in \mathbf{R}^{m+} \mid Ad \leq x\})^{1/m}, \text{Vol}(\{d \in \mathbf{R}^{m+} \mid Ad \leq y\})^{1/m}\right).
\end{aligned}$$

Thus $\lambda x + (1 - \lambda)y$ is feasible and the feasible region must be convex (since x, y were arbitrary). The first inequality is because the Minkowski sum of $\{d \in \mathbf{R}^{m+} \mid Ad \leq x\}$ and $\{d \in \mathbf{R}^{m+} \mid Ad \leq y\}$ is contained in the set $\{x \in \mathbf{R}^{m+} \mid Ad \leq \lambda x + (1-\lambda)y\}$. The second inequality is the Brunn-Minkowski theorem and the third is an obvious fact about concave functions.

To extend the argument to more general F we define:

$$\Delta_a(x) = \{d \in \mathbf{R}^{m+} \mid Ad \leq x, \ \ F(d) \geq a\},$$

the set of points with density $\geq a$ that do not exceed the stocking vector x. Again by simple properties of the Minkowski sum we have for $\lambda \in [0, 1]$

$$\Delta_a(\lambda x + (1-\lambda)y) \supseteq \lambda \Delta_a(x) + (1-\lambda)\Delta_a(y).$$

And, by Brunn-Minkowski,

$$\left(\int_{\left(\lambda\Delta_a(x)+(1-\lambda)\Delta_a(y)\right)} 1\right)^{\frac{1}{m}} \geq$$

$$\lambda\left(\int_{\Delta_a(x)} 1\right)^{\frac{1}{m}} + (1-\lambda)\left(\int_{\Delta_a(y)} 1\right)^{\frac{1}{m}}$$

so

$$\int_{\Delta_a\left(\lambda x+(1-\lambda)y\right)} 1 \geq \min\left(\int_{\Delta_a(x)} 1, \int_{\Delta_a(y)} 1\right)$$

Using that for any z

$$\int_{\{d \mid Ad \leq z\}} F(d) = \int_{a=0}^{\infty}\int_{\Delta_a(z)} 1\,\mathrm{d}\,a$$

we can integrate over levels of the density F and get

$$\int_{\Delta_0\left(x+(1-\lambda)y\right)} G \geq \min\left(\int_{\Delta_0(x)} G, \int_{\Delta_0(y)} G\right)$$

thus if $x, y \in K$ then $(\lambda x + (1-\lambda)y) \in K$ and K is convex.

2.3.2 Hardness of discrete version of component commonality

For m points $y^1, y^2, \ldots y^m$ in $\mathbf{R}^n$ and x also in $\mathbf{R}^n$ let $\mathrm{dom}(x)$ be the set $\{i \mid i \in 1\cdots m,\ x \geq y^i\}$.

Theorem 1 *Given m points $y^1, y^2, \ldots y^m$ in $\mathbf{R}^n$, a fraction γ, a positive vector c and a real number ζ, deciding if there exist an $x \in \mathbf{R}^n$ such that $x \geq y^i$ is satisfied for at least γm different i 's and $c \cdot x \leq \zeta$ is NP complete.*

Proof sketch: Let $G = (V, E)$ be an undirected graph with vertices $V = \{1, 2\cdots n\}$ and edges E. Given an integer k, the clique problem is: "does G have a complete induced subgraph on some k vertices?" For each $e \in E$ let $y^e \in \mathbf{R}^n$ be the vector such that

$$y_i^e = \begin{cases} 1 & \text{vertex } i \text{ is an endpoint of edge } e \\ 0 & \text{otherwise} \end{cases}$$

Let $c = \vec{1}$, $\gamma = (k-1)k/(2|E|)$ and $\zeta = k$.

Let x be a solution to the above problem. WLOG assume x is a 0-1 vector and define $G_x = (V_x, E_x)$ to be the induced subgraph of G where $V_x = \{i \in V \mid x_i \geq 1\}$. We then have $|G_x| = c \cdot x$ and $|E_x| = |\operatorname{dom}(x)|$. So we see that x such that $c \cdot x \leq k$ and $|\operatorname{dom}(x)| \geq k(k-1)/2$ correspond precisely to induced subgraphs of G with $\leq k$ vertices and $\geq k(k-1)/2$ edges: k-cliques. $\square$

This result compliments the recent result of Martin Dyer and Leen Stougie[25] which shows that two stage stochastic programming with simple recourse is $\sharp P$ complete.

2.3.3 General remarks

While we do assume that a membership oracle is available for K, we do not assume that a separation oracle is available. By a theorem of Yudin and Nemirovskiĭ it is known that membership and separation are polynomial time equivalent for convex programming problems like this one [33]. But the conversion has a large exponent, and the approach here is more efficient.

Instead of computing the integral each time the membership oracle is called, we could instead pick m samples $y^1, y^2, \ldots y^m$ according to D at the outset and then for each query x to the membership oracle for K, say yes to the query if and only if for at least γm of the y^i, we have $x \geq y^i$. It is easy to show by standard techniques that for m large enough, this suffices as a an approximate test of membership in K. We do not go into the details here. Also, in the actual situation, either the y^i may be available from past data or if one hypothesizes a particular log-concave density D, we may draw samples according to the density using the techniques of [5].

One may be tempted to solve the discrete problem: given m points $y^1, y^2, \ldots y^m$ in $\mathbf{R}^n$, a fraction γ, a positive vector c and a real number ζ, does there exist an $x \in \mathbf{R}^n$ such that $x \geq y^i$ is satisfied for at least γm different i 's and we also have $c \cdot x \leq \zeta$? In this generality, we show (in the appendix) that the problem is NP-hard. So one needs to exploit the special nature of K, namely its convexity.

For an NLP approach to the CC problem, see [37] and references provided there.

2.4 The function F: bias and damping.

Let K be an up-monotone convex set contained in the positive orthant of $\mathbf{R}^n$, x^f a known feasible point and $c > 0$ the cost vector in the linear objective function. For real numbers $L \geq 0, T > 0$, we define a log-concave distribution $B_{(L,T)}$ on $\mathbf{R}^n$ such that if $L \leq \inf\{c \cdot y \mid y \in K\}$ then $\mu_{B_{(L,T)}}(\{y \in K \mid c \cdot y \leq \inf\{c \cdot y \mid y \in K\} + T\})/\mu_{B_{(L,T)}}(\mathbf{R}^n) \geq \frac{1}{4}$ and show how to use the membership oracle for K to approximately sample from this distribution in an efficient manner.

Let x^l be such that $\forall y \in K,\ x^l \leq y$ and let x^u be such that $\forall y \in K,\ \exists z \in K$ such that $c \cdot z \leq c \cdot y$ and $z \leq x^u$. Note that by the up-monotone property, $x^u \in K$. If x^l and x^u are not explicitly given we can take $x^l = 0$ and x^u such that $x_i^u = (1/c_i)\sum_{i=1}^n c_i x_i^f$. Let $K_L = K \cap \{y \mid c \cdot y \geq L\}$ (we will later further restrict our attention to a "rectangle" that is roughly $\left\{x \;\middle|\; x^l \leq x \leq x^u\right\}$).

Definition 1 *For any real $L \geq 0$, $T > 0$ let "the tip" be the set $x \in K_L$ such that $c \cdot x \leq \inf\{c \cdot y \mid y \in K_L\} + T$.*

Let $\psi_{(K_L,x^u)}(x)$ denote the infimum of all real positive numbers λ such that $x^u + \frac{x-x^u}{\lambda} \in K_L$. This is the dilation of K_L about x^u needed to contain x and is called the *gauge function* associated with K_L. When the context is clear we will suppress the subscripts and use ψ instead of $\psi_{(K_L,x^u)}$.

F will be of the form

$$F(x) = e^{-\alpha \max(\psi_{(K_L,x^u)}(x)-1,0)} e^{-\beta c \cdot x} \tag{2.1}$$

where β and α are positive reals to be determined later. We take $F(x)$ as the unnormalized density function for $B_{(L,T)}$, a log-concave distribution on $\{x \in \mathbf{R}^n \mid x \leq x^u\}$. Note that $e^{-\beta c \cdot x}$ is the bias (in favoring better objective values), and $e^{-\alpha \max(\psi_{(K_L,x^u)}(x)-1,0)}$ is the function to damp this bias in regions that are infeasible.

The selection of α and β is done in two stages: (1) we first find β such that at least half of the probability mass in the body (according to the density $B_{(L,T)}$) is in "the tip", and then (2) find α such that at least half the mass in the entire rectangle is in the body.

2.4.1 Getting in the tip

For x in K_L, the distribution F is a function of the $c \cdot x$ only (since $\psi(x) \leq 1$).

Lemma 1 *If $L \geq 0$,$T > 0$ and $\beta \geq \frac{n}{T}$, we have*

$$\int_{\text{“the tip”}} F \geq (1/2) \int_{K_L} F.$$

Proof: Let $c^* = \inf \{c \cdot y \mid y \in K_L\}$, $z \in K_L$ such that $c \cdot z = c^*$ and A^* be the intersection of the hyperplane $c \cdot x = c^* + T$ and K. Clearly, the convexity of K_L implies that $\int_{K_L \cap \{x: c \cdot x \leq c^* + T\}} F$ is not increased if we replace $K_L \cap \{x : c \cdot x \leq c^* + T\}$ by the convex hull of z and A^*. Also, replacing $K_L \cap \{x : c \cdot x \geq c^* + T\}$ by the truncated cone formed by intersecting $\{x : c \cdot x \geq c^* + T\}$ with the minimal pointed cone with vertex z containing A^* cannot decrease $\int F$ over this set. So it suffices to prove the lemma with K_L equal to the infinite cone. Then the ratio of integrals in the lemma is

$$\frac{\int_{c^*}^{c^*+T} \left(\frac{\lambda - c^*}{T}\right)^{n-1} \text{area}(A^*) e^{-\beta\lambda} \,\mathrm{d}\,\lambda}{\int_{c^*}^{\infty} \left(\frac{\lambda - c^*}{T}\right)^{n-1} \text{area}(A^*) e^{-\beta\lambda} \,\mathrm{d}\,\lambda} = \frac{\int_{c^*}^{c^*+T} (\lambda - c^*)^{n-1} e^{-\beta\lambda} \,\mathrm{d}\,\lambda}{\int_{c^*}^{\infty} (\lambda - c^*)^{n-1} e^{-\beta\lambda} \,\mathrm{d}\,\lambda}.$$

We change variables (and consult standard integral tables) to get:

$$\frac{\int_0^T \lambda^{n-1} e^{-\beta\lambda} \,\mathrm{d}\,\lambda}{\int_0^\infty \lambda^{n-1} e^{-\beta\lambda} \,\mathrm{d}\,\lambda} = 1 - e^{-\beta T} \sum_{k=0}^{n-1} \frac{(\beta T)^k}{k!}.$$

By picking $\beta = \frac{n}{T}$ the ratio is $1 - e^{-n} \sum_{k=0}^{n-1} \frac{n^k}{k!}$, which is $\geq \frac{1}{2}$ for $n \geq 1$, so any

$$\beta \geq \frac{n}{T}, \tag{2.2}$$

will do. □

2.4.2 Staying in the body.

We now show that at least $\frac{1}{2}$ of the mass is in the body K_L, which would imply that at least $\frac{1}{4}$ of the mass of $B_{(L,T)}$ is in the tip (the near optimal feasible region), for a suitable choice of α.

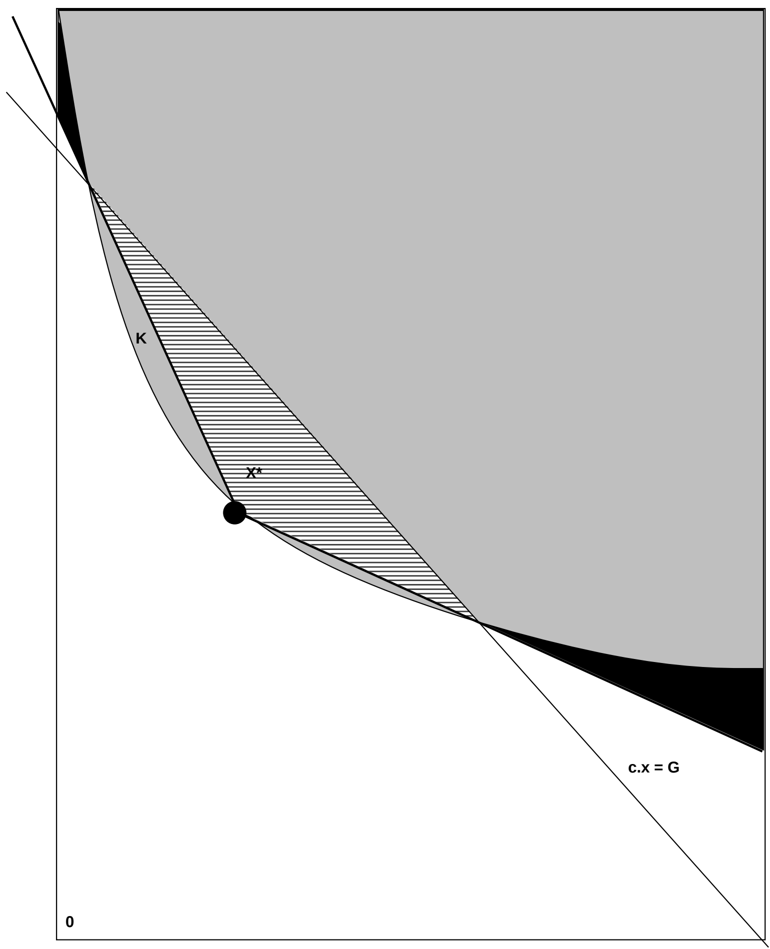


Figure 2.2: The body versus a cone.

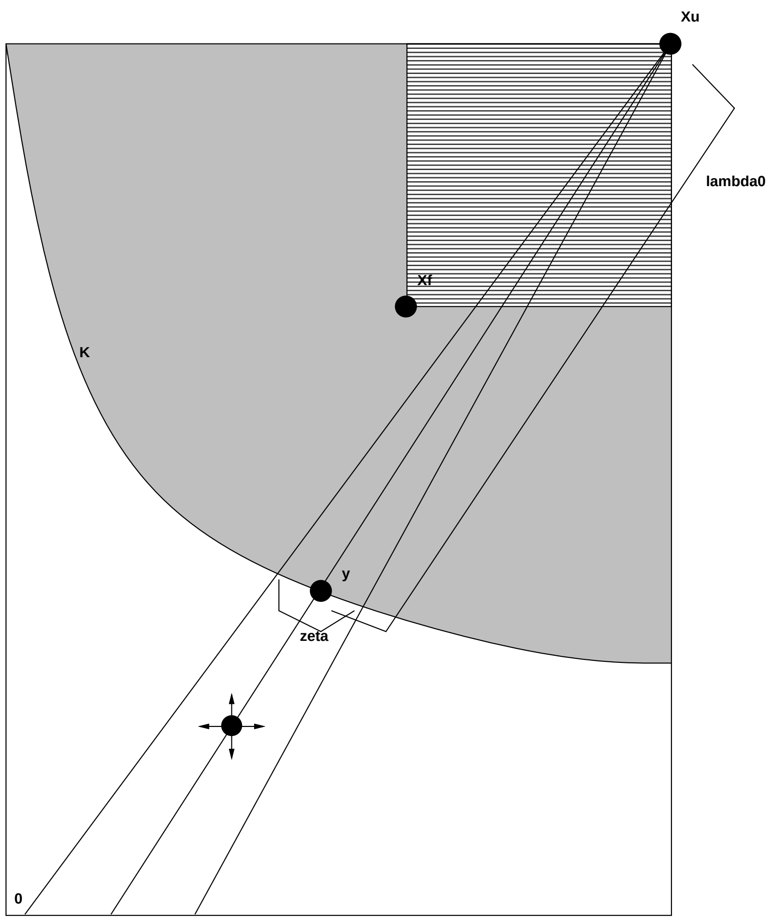


Figure 2.3: An example “thin” cone.

Let ∂K_L be the set of y in the boundary of K_L. For any $\zeta > 0$ all of K_L can be covered by a collection C of disjoint cones such that for each cone $r \in C$, we have $\|x-y\|_2 \leq \zeta$ for $x, y \in r \cap \partial K_L$.

Now, if less than half of the mass according to $B_{(L,T)}$ is in K_L then there must be a cone $r \in C$ such that less than half the mass according to $B_{(L,T)}$ restricted to r is in $r \cap K_L$. By the arbitrary nature of ζ we see the same must hold for some "infinitely" thin cone. Chose such a cone and let y be the point at which the cone intersects ∂K and set $\lambda_0 = \|x^u - y\|_2$.

Lemma 2 *Suppose*

$$\alpha \geq \max\Big(3\beta(c \cdot x^u - L) + 3n - 5, n(e^2+1) + 1\Big) \tag{2.3}$$

Then the mass of any infinitesimal cone outside of K_L can be shown to be no more than the mass of the same cone inside K_L , thus yielding a ratio of feasible to total of at least $\frac{1}{2}$.

Proof: For the proof of this lemma only, it will be convenient to multiply masses by $e^{\beta c \cdot x^u}$; so for any set S, we mean by the mass of S, the quantity $e^{\beta c\cdot x^u}\int_S F$. The mass of the cone outside K_L is given by:

$$\begin{aligned}
& \int_{\lambda_0}^{\infty} \lambda^{n-1} e^{\beta(c\cdot x^u - c\cdot y)\frac{\lambda}{\lambda_0} - \alpha(\frac{\lambda}{\lambda_0}-1)}\,\mathrm{d}\,\lambda \\
= \quad & \lambda_0^n \int_1^{\infty} t^{n-1} e^{\beta(c\cdot x^u - c\cdot y)t - \alpha(t-1)}\,\mathrm{d}\,t \\
\leq \quad & \lambda_0^n \int_1^{\infty} e^{(t-1)(n-1)} e^{\beta(c\cdot x^u - c\cdot y)t - \alpha(t-1)}\,\mathrm{d}\,t \\
= \quad & \lambda_0^n e^{\beta(c\cdot x^u - c\cdot y)}/(\alpha - \beta(c\cdot x^u - c\cdot y) - (n-1)).
\end{aligned}$$

For the mass of the ray inside K_L we will break into two cases depending if $\beta(c \cdot x^u - c \cdot y)$ is ≥ 2 or not. The mass of the ray inside K_L is at least

$$\begin{aligned}
& \int_0^{\lambda_0} \lambda^{n-1} e^{\beta\frac{\lambda}{\lambda_0}(c\cdot x^u - c\cdot y)}\,\mathrm{d}\,\lambda \\
= \quad & \lambda_0^n \int_0^1 t^{n-1} e^{\beta t(c\cdot x^u - c\cdot y)}\,\mathrm{d}\,t.
\end{aligned}$$

We now consider the two cases.

Case 1: $\beta(c \cdot x^u - c \cdot y) \geq 2$: An integration by parts gives the mass of the ray inside K_L equals

$$\begin{aligned} & \frac{\lambda_0^n}{\beta(c \cdot x^u - c \cdot y)}(e^{\beta(c \cdot x^u - c \cdot y)} - (n-1)\int_0^1 t^{n-2} e^{\beta t(c \cdot x^u - c \cdot y)} \,\mathrm{d}\, t) \\ \geq \; & \frac{\lambda_0^n}{\beta(c \cdot x^u - c \cdot y)}(e^{\beta(c \cdot x^u - c \cdot y)} - (n-1)\int_0^1 e^{(t-1)(n-2)} e^{\beta t(c \cdot x^u - c \cdot y)} \,\mathrm{d}\, t) \\ \geq \; & \frac{\lambda_0^n e^{\beta(c \cdot x^u - c \cdot y)}}{\beta(c \cdot x^u - c \cdot y)}(1 - \frac{(n-1)}{\beta(c \cdot x^u - c \cdot y) + n - 2}). \end{aligned}$$

The ratio of mass inside K_L to outside is at least

$$\frac{\beta(c \cdot x^u - c \cdot y) - 1}{\beta(c \cdot x^u - c \cdot y) + n - 2} \frac{\alpha - \beta(c \cdot x^u - c \cdot y) - n + 1}{\beta(c \cdot x^u - c \cdot y)}.$$

Since $\beta(c \cdot x^u - c \cdot y) \geq 2$ and $\alpha \geq 3\beta(c \cdot x^u - c \cdot y) + 3n - 5$, the ratio is at least 1.

Case 2: $\beta(c \cdot x^u - c \cdot y) < 2$**:** The mass of the ray inside K_L is at least

$$\lambda_0^n \int_0^1 t^{n-1} \,\mathrm{d}\, t \min\left(1, e^{\beta(c \cdot x^u - c \cdot y)}\right)$$

yielding a ratio of

$$\frac{\alpha - \beta(c \cdot x^u - c \cdot y) - (n-1) \min\left(1, e^{\beta(c \cdot x^u - c \cdot y)}\right)}{n e^{\beta(c \cdot x^u - c \cdot y)}}$$

which, by our assumption on β, is at least

$$\frac{\alpha - 2 - (n-1)}{n e^2}$$

and since $\alpha \geq n(e^2 + 1) + 1$ this is at least 1.

□

Summarizing, we have the following:

Theorem 2 *For any $L \geq 0$, $T > 0$ and F as in (2.1) with α, β satisfying*

$$\begin{aligned} \beta &\geq \frac{n}{T} \\ \alpha &\geq \max\left(3\beta(c \cdot x^u - L) + 3n - 5, n(e^2+1) + 1\right) \end{aligned}$$

then

$$\int_{\text{“the tip”}} F \geq (1/4) \int_{\mathbf{R}^n} F.$$

2.5 Sampling Procedure.

We now show how to approximately sample according to F ($= B_{(L,T)}$). First, we discretize the rectangle $x^l \leq x \leq x^u$. Next, we find an approximation for F in regions not in K_L. Third, we devise the transition matrix of a Markov chain which realizes the desired random walk.

Let E^i denote the unit vector directed in the ith coordinate direction. Let $C_\delta(p)$ denote the cube of side 2δ centered at p.

We will divide the rectangle $x^l \leq x \leq x^u$ into small cubes of side 2δ where δ will be specified later.

$$U \stackrel{\text{def}}{=} \left\{ x \in \mathbf{R}^n \;\middle|\; \frac{x - x^f}{2\delta} \in Z^n \,,\; C_\delta(x) \cap \left\{ y \in \mathbf{R}^n \;\middle|\; x^l \leq y \leq x^u \right\} \neq \emptyset \right\} \tag{2.4}$$

$$J \stackrel{\text{def}}{=} \bigcup_{x \in U} C_\delta(x) \tag{2.5}$$

We will take a random walk on the graph whose vertices are the set U of centers of cubes. Also, notice that even though there may be $x \in U$ such that $x \not\geq x^l$ we do have $x + \delta\vec{1} \geq x^l$ for all $x \in U$. It is important to notice that the l_∞ diameter of J is no more that $\left\| x^u - x^l \right\|_\infty + 4\delta$.

Many of the lemmas require that δ not be too large with respect to x^u, x^f, α and β. To formalize this we will can δ “fine” if we have

$$\delta \leq \min\left(\frac{\min_i (x_i^u - x_i^f)}{7\alpha}, \frac{1}{7\sqrt{n}\beta c_i}\right), \tag{2.6}$$

and we will often invoke the following result.

Proposition 1 *If δ is “fine” and α, β meet the conditions of Theorem 2, then $\frac{x_i^u - x_i^f}{\delta} \geq 7((e^2+1)n+1)$ for all i.*

2.5.1 Discretization and Approximate Sampling using Membership Oracle.

Errors due to two sources need to be analyzed here. One source of error is because we discretize the region, and approximate the integral of $F(x)$ over a small rectangle by $F(p)*$(volume of the rectangle), where p is the center of the rectangle. The second source of error is in computing the gauge function for points outside the feasible region. This is because in practice we may only be able to calculate $\hat{F}(p)$, an approximation for $F(p)$ (using the membership oracle and bisection methods).

In this section (and in section 2.6) we will need for every $p \in U$ that the integral of F over any rectangle C centered at p and approximately contained in $C_\delta(p)$ is well approximated by $F(p)Vol(C)$.[1]

More precisely: For every $p \in J$ we need to determine a lower bound on ρ such that for some continuous monotone decreasing function ξ such that $\xi(0) = 0$ and all $\eta \geq 0$ in some open neighborhood of 0: if C is any rectangular region with center p contained in $C_{\delta+\eta}(p) \cap J$ then

$$\rho(1-o(1))\left(\frac{1}{Vol(C)}\int_C F\right) \leq F(p) \leq (\rho(1-o(1)))^{-1}\left(\frac{1}{Vol(C)}\int_C F\right). \quad (2.7)$$

We will also need a lower bound on σ such that

$$\sigma F(x) \leq F(x + \lambda E^i) \quad (2.8)$$

for all $x \in K$, $i \in \{1 \cdots n\}$ and all $|\lambda| \leq 2\delta$.

The lemma below summarizes the errors induced here. The proof is in section 2.8.

Lemma 3 *(2.7) and (2.8) hold with*

$$\sigma \geq e^{-2\alpha\delta/(\min_i(x_i^u - x_i^f))}e^{-2\beta\delta\|c\|_\infty} \quad (2.9)$$

$$\rho \geq \frac{e^{-\alpha\delta/(\min_i(x_i^u - x_i^f))}}{1 + \sqrt{2\pi}\beta\delta\,\|c\|_2 \operatorname{erf}\left(\frac{\beta\delta\|c\|_2}{\sqrt{2}}\right) e^{\beta^2\delta^2\|c\|_2^2/2}} \quad (2.10)$$

[1]The approximate containment, characterized by a parameter η, is technical point used only to facilitate the proof of Lemma 4.

Let $\hat{F}(x)$ be an approximation for $F(x)$ calculated using only x and the membership oracle. In the feasible region, $\hat{F} = F$. In the infeasible region, $\hat{F}$ may not equal F because the dilation cannot be computed exactly. Note that for our analysis this must be a deterministic approximation and not one obtained by sampling; to be clear, we always calculate the same value for $\hat{F}(x)$. This is calculated to a relative accuracy of $1 \pm \frac{1}{11}$ (i.e. $\frac{10}{11}F(x) \leq \hat{F}(x) \leq \frac{12}{11}F(x)$.)

To calculate $\hat{F}(x)$ to a relative accuracy of $1/11$, it is sufficient to calculate the gauge function to an absolute error of $\pm\frac{\ln(12/11)}{\alpha}$. To achieve this, it is sufficient to calculate the distance of the point in K on the line segment from x to x^u that is farthest from x^u to an absolute accuracy of $\frac{\ln(12/11)(\min_i(x_i^u - x_i^f))^2}{2\left\|x^u - x^f\right\|_2 \alpha}$. This can be done very quickly by bisection search using our membership oracle.

2.5.2 The Markov Chain.

For each $x \in U$ let $N(x)$ be the "neighborhood" of x which is the set of all vertices in U that differ from x in exactly one coordinate by $\pm 2\delta$. The transition probabilities $P(x, y)$ will be:

$$P(x,y) = \begin{cases} \frac{1}{2n}\min(1, \frac{\hat{F}(y)}{\hat{F}(x)}) & y \in N(x) \\ 1 - \sum_{z \in N(x)} P(x,z) & x = y \\ 0 & y \notin N(x) \end{cases} \tag{2.11}$$

where $\hat{F}$ is a deterministic estimate for F. It is easy to see that $P(x, y)$ induces a time reversible irreducible aperiodic Markov chain with steady state probabilities $\pi(\cdot)$ proportional to $\hat{F}(\cdot)$. We will show, in the next section, that after a sufficient number of steps, we are fairly close to the steady state. Let π be the unique steady state distribution for our chain (approximately $B_{(L,T)}$). But first we show that in the steady state there is a reasonable chance of observing states corresponding to cubes covering the near optimal feasible portions of K_L.

Theorem 3 *Let $\pi(\cdot)$ be the steady state probabilities of the above Markov chain, then if α, β satisfy the conditions of Theorem 2 and δ is "fine" and "the tip" is contained in J then*

$$\sum_{x \in U, C_\delta(x) \cap \text{"the tip"} \neq \emptyset} \pi(x) \geq \frac{1}{6}.$$

Proof: Accounting for the errors due to discretization and in gauge function computation, $\pi(x)$ ($= D\hat{F}(x)$) satisfies

$$D(1-\frac{1}{11})\rho\left(\frac{1}{(2\delta)^n}\int_{C_\delta(x)}F\right) \leq \pi(x) \leq D(1+\frac{1}{11})\rho^{-1}\left(\frac{1}{(2\delta)^n}\int_{C_\delta(x)}F\right). \tag{2.12}$$

where $D = \left(\sum_{x\in U}\hat{F}\right)^{-1}$.

Now, since the "tip" probability is at least $\frac{1}{4}$ (by construction of F), we have the steady state probability of the cubes covering the tip is:

$$\sum_{x\in U, C_\delta(x)\cap\text{"the tip"}\neq\emptyset} \pi(x) \geq \frac{1\times(1-\frac{1}{11})\rho}{3\times(1+\frac{1}{11})\rho^{-1}+1\times(1-\frac{1}{11})\rho}.$$

δ is fine, so:

$$\frac{\alpha\delta}{\min_i(x_i^u - x_i^f)} \leq \frac{1}{7} \tag{2.13}$$

and

$$\beta\delta\|c\|_2 \leq \frac{1}{7}, \tag{2.14}$$

and the theorem follows from inequality (2.10). □

2.6 Spectral Gap of the Markov Chain.

In this section, we determine how many steps are necessary for the Markov chain to "mix". We need to find a relationship between the number of steps walked and how close we are to the steady state. For this we need a central result of Sinclair and Jerrum [47]. We also need several well-know facts that are collected in [19]. Let X be the set of states in our Markov chain. For $x, y \in X$, let $P^t(x,y)$ be the probability that starting in state x we are in state y after t steps. As before, let π be the unique steady state distribution for our chain (approximately $B_{(L,T)}$). Recall that $P(x,y)$ induces a time reversible irreducible aperiodic Markov chain; however, it is not strongly aperiodic [47]. This is because we do not insist that we have $P(x,x) \geq \frac{1}{2}$ for all x. Because of this, we not only need an upper bound for the second largest eigenvalue but also need a lower bound on the smallest eigenvalue [19].

We will use Proposition 3 from [19] which says:

$$\sum_{y \in X} |P^t(x,y) - \pi(y)| \leq \sqrt{\frac{1-\pi(x)}{\pi(x)}}(\chi^*)^t$$

.

where, $\chi^* = \max(\chi_1, |\chi_m|)$, where χ_1 is the second largest eigenvalue and χ_m is the smallest eigenvalue of P, the matrix of transition probabilities.

We find an upper bound on χ_1 by appealing to a result of Sinclair and Jerrum's [47], which is quoted also as Proposition 6 of [19], namely, $\chi_1 \leq 1 - \frac{\hat{\phi}^2}{2}$, where $\hat{\phi}$ is a lower bound on the "conductance" (defined in section 2.6.1) of the Markov chain. We find a lower bound on the conductance in section 2.6.1 below. A lower bound on χ_m is obtained by appealing to Proposition 2 of [19], described in detail in section 2.6.2.

We also prove that (see sections 2.6.1-2.6.2) $|\chi_m| \leq 1 - \frac{\hat{\phi}^2}{2}$. Thus, we have:

$$\sum_{y \in X} |P^t(x,y) - \pi(y)| \leq \frac{1}{\sqrt{\pi(x)}}(1 - \frac{\hat{\phi}^2}{2})^t.$$

What we wish to do is find t so that $\sum_{y \in X} |P^t(x,y) - \pi(y)|$ is under $1/12$. Recall that by Theorem 3, the states of our Markov chain corresponding to cubes that cover the "tip" have mass at least $1/6$. Thus, we will have a chance of at least $1/6 - 1/12 = 1/12$ that a random walk of t steps will end in one of the states corresponding to a cube covering the tip (i.e. close to the optimum). For this, it suffices to have

$$t \geq \ln\left(\frac{\sqrt{\pi(x)}}{12}\right) / \ln\left(1 - \frac{\hat{\phi}^2}{2}\right) \geq \ln\left(\frac{12}{\sqrt{\pi(x)}}\right) / \left(\frac{\hat{\phi}^2}{2}\right) = \frac{2\ln(12) - \ln(\pi(x))}{\hat{\phi}^2}. \tag{2.15}$$

We now wish to prove a lower bound on $\pi(x^f)$, our starting point, so that we can apply the above inequality.

We assume that the walk is started deterministicly at x^f. Since we know that at least half of the mass of $B_{(L,T)}$ is in the body and the highest possible stationary probability of a cube intersecting K_L is at most $e^{\beta(c \cdot (x^f + 2\delta\vec{1}) - L)}\pi(x^f)$ and there are at most $\prod_{i=1}^{n} \left\lceil \frac{x_i^u - x_i^l}{2\delta} + 1 \right\rceil$ states in K_L (or even U), we know

$$\frac{\rho 10/11}{\rho^{-1} 12/11}\frac{1}{2} \leq \pi(x^f) e^{\beta(c \cdot (x^f + 2\delta\vec{1}) - L)} \prod_{i=1}^{n} \left\lceil \frac{x_i^u - x_i^l}{2\delta} + 1 \right\rceil$$

which yields

$$\pi(x^f) \geq \frac{\rho^2 5 e^{\beta(L - c\cdot(x^f + 2\delta\vec{1}))}}{12 \prod_{i=1}^n \left\lceil \frac{x_i^u - x_i^l}{2\delta} + 1 \right\rceil}$$

but we will just use the easier form:

$$\frac{5\rho^2 e^{\beta(L - c\cdot(x^f + 2\delta\vec{1}))}}{12 \left\lceil \frac{\left\| x^u - x^l \right\|_\infty}{2\delta} + 1 \right\rceil^n}. \tag{2.16}$$

Plugging in, we get

Theorem 4 *If the conditions of Theorem 3 are met then running the above Markov chain for at least*

$$\frac{2\ln(12) + \ln\left(\frac{12}{5\rho^2}\right) + \beta(c\cdot(x^f + 2\delta\vec{1}) - L) + n\ln\left(\left\lceil \frac{\left\| x^u - x^l \right\|_\infty}{2\delta} + 1 \right\rceil\right)}{\hat{\phi}^2} \tag{2.17}$$

steps is sufficient to ensure with probability at least $\frac{1}{12}$, the chain will stop at a state x such that $x + \delta\vec{1}$ is feasible and within cost $T + 2\delta \left\| c \right\|_1$ of the optimal point.

Although at the end of Section 2.4, we had shown that the steady state probability of being in the tip is at least $\frac{1}{6}$, now we only guarantee the probability that the sample is in the tip at the end of t steps is (at least) $\frac{1}{12}$. Thus, the three sources of errors–discretization, approximation of the gauge function, walking for t steps–reduce the tip probability from $\frac{1}{4}$ to $\frac{1}{12}$.

2.6.1 Conductance.

Take aribtrary $V \subseteq U$ and $\bar{V} = U \setminus V$. We define the conductance of V by

$$\phi_V = \frac{\sum_{x\in V, y\in \bar{V}\cap N(x)} \pi(x)P(x,y)}{\min(\pi(V), \pi(\bar{V}))} = \frac{\sum_{x\in V, y\in \bar{V}\cap N(x)} \min(\hat{F}(x), \hat{F}(y))}{2n \min(\hat{F}(V), \hat{F}(\bar{V}))}. \tag{2.18}$$

The conductance of the chain defined by

$$\phi = \min_{V\subseteq U} \phi_V. \tag{2.19}$$

We use an "isoperimetric inequality" to find a lower bound for ϕ. An isoperimetric inequality was proved in [23]; a simpler proof of a stronger inequality was given in [40]. The inequality was generalized to the case of log-concave functions in [5]. We use here a version of this from [21]. If $\text{dist}(x, y) = ||x - y||$ where $||\cdot||$ is an arbitrary norm on $\mathbf{R}^n$ and $\text{diam}(K) = \max_{x,y \in K} \text{dist}(x, y)$ we have the following theorem from [21]:

Theorem 5 *Let $J \subseteq \mathbf{R}^n$ be a convex body and F a log-concave function defined on $\text{int}\, J$ and μ the induced measure. Let $S_1, S_2 \subseteq J$, and $t \le \text{dist}(S_1, S_2)$ and $d \ge \text{diam}(J)$. If $B = J \setminus (S_1 \cup S_2)$, then*

$$\min(\mu(S_1), \mu(S_2)) \le \frac{1}{2}(d/t)\mu(B). \tag{2.20}$$

We will use $\text{dist}(x, y) = \|x - y\|_\infty$.

Lemma 4 *If δ is fine then $\phi \ge \frac{\delta}{3n\|x^u - x^l\|_\infty} = \hat{\phi}$ (say).*

Proof: For any $V \subseteq U$ and $\bar{V} = U \setminus V$, and

$$\begin{aligned} V_\delta &= \bigcup_{x \in V} C_\delta(x) \\ \bar{V}_\delta &= \bigcup_{x \in \bar{V}} C_\delta(x) \end{aligned}$$

Let η be a small positive real that will tend to zero. Let B_δ be the $\eta/2$ neighborhood of $V_\delta \cap \bar{V}_\delta$ and $B_\delta(x, y)$ be the $\eta/2$ neighborhood of $C_\delta(x) \cap C_\delta(y)$. Let S_1, S_2 and B be $V_\delta \setminus B_\delta, \bar{V}_\delta \setminus B_\delta$ and $B_\delta \cap J$ respectively.

From inequalities (2.7), (2.8), (2.9), (2.10) and (2.12), it is clear that

$$\begin{aligned} \sum_{x \in V, y \in \bar{V} \cap N(x)} \min(\hat{F}(x), \hat{F}(y)) &\ge \frac{10}{11} \sum_{x \in V, y \in \bar{V} \cap N(x)} \min(F(x), F(y)) \\ &\ge \frac{10}{11}\sigma \sum_{x \in V, y \in \bar{V} \cap N(x)} F\left(\frac{x+y}{2}\right) \\ &\ge \frac{10}{11}\sigma \sum_{x \in V, y \in \bar{V} \cap N(x)} \rho \frac{1}{\eta(2\delta + \eta)^{n-1}} \int_{B_\delta(x,y)} F \end{aligned}$$

$$
\begin{aligned}
&\geq \frac{10}{11}\frac{\sigma\rho}{\eta(2\delta+\eta)^{n-1}}\int_{B_\delta} F \\
&\geq \frac{10}{11}\frac{\sigma\rho}{\eta(2\delta+\eta)^{n-1}}\int_{B} F \\
\text{similarly} \quad \min(\hat{F}(V),\hat{F}(\bar{V})) &\leq \frac{12}{11}\frac{1}{\sigma\rho(2\delta)^n}\min\left(\int_{V_\delta} F, \int_{\bar{V}_\delta} F\right) \\
&\leq \frac{12}{11}\frac{1}{\sigma\rho(2\delta-\eta)^n}\min\left(\int_{S_1} F, \int_{S_2} F\right)
\end{aligned}
$$

From the isoperimetric inequality,

$$
\frac{\int_B F}{\min(\int_{S_1} F, \int_{S_2} F)} \geq \frac{2\eta}{\|x^u - x^l\|_\infty + 4\delta}.
$$

Combining this with the inequalities above and taking the limit $\eta \to 0$ we get:

$$
\phi \geq \frac{2\frac{10}{11}\sigma^2\rho^2\delta}{\frac{12}{11}n(\|x^u - x^l\|_\infty + 4\delta)}. \tag{2.21}
$$

Since δ is fine (from inequalities (2.10),(2.9))

$$
\phi \geq \frac{\delta}{3n\,\|x^u - x^l\|_\infty} = \hat{\phi}. \tag{2.22}
$$

□

2.6.2 Comparison of χ_1 and $|\chi_m|$

Here we find a lower bound for χ_m, by the *canonical odd path* argument outlined in Proposition 2 of [19].

Let

$$
\Delta = \frac{\delta}{8n\,\|x^u - x^l\|_\infty}. \tag{2.23}
$$

For each state x let ω_x be the smallest non-negative integer such that $P(x + 2\omega_x\delta E^1, x + 2\omega_x\delta E^1) \geq \Delta$ (this is always possible since $P(a,a) \geq \frac{1}{2n} \geq \Delta$ on the border of our bounding region). Let σ_x be the $2\omega_x + 1$ step path of from x to x given by:

$$
\underline{x} \mapsto \underline{x+2\delta E^1} \mapsto \underline{x+4\delta E^1} \mapsto \cdots \overbrace{\underline{x+2\omega_x\delta E^1} \mapsto \underline{x+2\omega_x\delta E^1}}^{\text{“self loop”}} \mapsto \cdots \underline{x+4\delta E^1} \mapsto \underline{x+2\delta E^1} \mapsto \underline{x}
$$

We will call σ_x "the canonical odd path for x".

Proposition 6 of [19] states for any selection of canonical odd paths we have $\chi_m \geq -1 + \frac{2}{\iota}$, where

$$\iota \stackrel{\text{def}}{=} \max_{(a,b)} \sum_{\sigma_x \ni (a,b)} \|\sigma_x\|_P \, \pi(x), \tag{2.24}$$

$$\text{where} \quad \|\sigma_x\|_P \stackrel{\text{def}}{=} \sum_{(a,b)\in\sigma_x} \frac{1}{\pi(a)P(a,b)}. \tag{2.25}$$

To prove the bound for χ^*, we will show that $|\chi_m|$ is less than the upper bound for χ_1. From the discussion above, it is sufficient to show the following.

Lemma 5 $\iota \leq \frac{4}{\phi^2}$

Proof: By our choice of paths, we have for any $i \leq \omega_x - 1$:

$$\begin{array}{rcccl} \frac{1}{2n} - \Delta & \leq & P(x + i2\delta E^1, x + (i+1)2\delta E^1) & \leq & \frac{1}{2n} \\ & & P(x + (i+1)2\delta E^1, x + i2\delta E^1) & \leq & \frac{1}{2n} \end{array}$$

and by time reversibility

$$\pi(x+(i+1)2\delta E^1) = \frac{P(x + i2\delta E^1, x + (i+1)2\delta E^1)}{P(x + (i+1)2\delta E^1, x + i2\delta E^1)}\pi(x+i2\delta E^1) \geq \frac{\frac{1}{2n} - \Delta}{\frac{1}{2n}}\pi(x+i2\delta E^1).$$

For any x we have,

$$\begin{aligned} \|\sigma_x\|_P &\leq 2\sum_{i=1}^{\omega_x} \frac{1}{\pi(x)(1-2\Delta n)^i(\frac{1}{2n} - \Delta)} + \frac{1}{\pi(x)(1-2\Delta n)^{\omega_x}\Delta} \\ &\leq 2 \sum_{i=1}^{\left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil} \frac{1}{\pi(x)(1-2\Delta n)^i(\frac{1}{2n} - \Delta)} + \frac{1}{\pi(x)\Delta(1-2\Delta n)^{\left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil}} \\ &\leq 2\left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil \frac{2n}{\pi(x)\left(1 - \left\lceil \frac{x_1^u - x_1^l}{2\delta} + 2 \right\rceil 2\Delta n\right)} + \frac{1}{\pi(x)\Delta\left(1 - \left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil 2\Delta n\right)}. \end{aligned}$$

Using Proposition 1 it is easy to show that

$$\left\lceil \frac{x_1^u - x_1^l}{2\delta} + 2 \right\rceil 2\Delta n \leq \frac{1}{3}.$$

Continuing,

$$\|\sigma_x\|_P \leq \left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil \frac{6n}{\pi(x)} + \frac{3}{2\pi(x)\Delta}.$$

Because each edge can be used by at most $\left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil$ canonical odd paths, we have

$$\iota \leq \left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil \left(\left\lceil \frac{x_1^u - x_1^l}{2\delta} + 1 \right\rceil 6n + \frac{3}{2\Delta} \right). \tag{2.26}$$

The result now follows from Lemma 4, inequality (2.26) and Proposition 1. □

2.7 Description and Analysis of the Algorithm.

Suppose we are given a feasible point x^f, a relative accuracy goal (ϵ) and a desired upper bound on the probability that the algorithm fails (κ). Let ν be the ratio of $c \cdot x^f$ to $c \cdot x^{opt}$. We assume that without loss of generality, the problem has been rescaled so $c_i = 1$ for all i ($x_i \to c_i x_i$, $c_i \to \frac{c_i}{c_i} = 1$).

2.7.1 In the worst case

Here, we assume that $n \geq 2$ and $\epsilon \leq 1$. We present **AlgorithmA**.
The analysis of **AlgorithmA** is fairly straight forward.

- It is easy to see that x^u such that $x_i^u = \sum_j x_j^f$ must dominate any optimal point.

- Each time we draw a sample according to Theorem (4) we have at least a chance of $\frac{1}{12}$ that it is feasible and within a cost of $T + 2n\delta$ of the optimum. We have picked T and δ such that $T + 2n\delta < \frac{\sum_j x_j^f - L}{2}$.

- When the **repeat** loop terminates either

  – $\sum_i x_i^{\text{best}} \leq \frac{\sum_j x_j^f + L}{2}$, or

Input: x^f, ϵ, κ, c and a membership oracle for K.

rescale problem so $c = \vec{1}$

$x^l \leftarrow 0$ (point-wise lower bound)

$L \leftarrow 0$ (probalistic cost lower bound)

$x^{\text{best}} \leftarrow x^f$ (the best feasible solution observed)

while $\sum_i x_i^f - L > \epsilon L$

$x_i^u \leftarrow 2\sum_j x_j^f$, $i = 1 \cdots n$ (point-wise upper bound)

$T \leftarrow \frac{\sum_j x_j^f - L}{3}$ ($\frac{1}{3}$ of the current gap)

$\beta \leftarrow \frac{n}{T}$ (objective function "bias")

$\alpha \leftarrow \frac{7n^2 \sum_i x_i^f}{T}$ (gauge function gain)

$\delta \leftarrow \frac{T}{49n^2}$ (step size)

repeat $\left\lceil \log_{\frac{12}{11}} \left(\frac{\lceil \log_2(\frac{7}{\epsilon}) \rceil + 1}{\kappa} \right) \right\rceil$ times or until $\sum_i x_i^{\text{best}} \leq \frac{\sum_j x_j^f + L}{2}$

run the random walk of section 2.5 with the above parameters for the number of steps prescribed in Theorem 4, let x be the stopping point of the walk.

if $x + \delta\vec{1}$ feasible and $\sum_i (x_i + \delta) < \sum_i x_i^{\text{best}}$

then $x^{\text{best}} \leftarrow x + \delta\vec{1}$

endrepeat

if $\sum_i x_i^{\text{best}} > \frac{\sum_j x_j^f + L}{2}$

then $L \leftarrow \sum_i x_i^{\text{best}} - T - 2n\delta$

$x^f \leftarrow x^{\text{best}}$

endwhile

return x^f

Figure 2.4: **AlgorithmA**.

  - enough samples have been drawn such that with confidence at least $(1-\kappa)^{\frac{1}{\lceil \log_2\left(\frac{7}{\epsilon}\right)\rceil+1}}$ one of them was feasible and within distance $T+2n\delta$ of the optimum.

  Either way, the distance from the new x^f to the new L is no more than half of the old distance.

- Thus, the outer **while** loop will run no more than $\lceil \log_2\left(\frac{\nu}{\epsilon}\right)\rceil$ times.

- Furthermore, we see the first time the algorithm alters the lower bound (it must establish a lower bound to halt), we have $L \geq \frac{\sum_i x_i^f}{2} - T - 2n\delta \geq \frac{1}{7}\sum_i x_i^f$.

  - Therefore, the outer while loop will cycle no more than $\left\lceil \log_2\left(\frac{7}{\epsilon}\right)\right\rceil$ more times.
  - Thus the lower bound can be altered at most $\left\lceil \log_2\left(\frac{7}{\epsilon}\right)\right\rceil + 1$ times.
  - Since each lower bound alteration is correct with chance at least $(1-\kappa)^{\frac{1}{\lceil \log_2\left(\frac{7}{\epsilon}\right)\rceil+1}}$ we see that they all are correct with odds at least $1-\kappa$.

Thus the algorithm fails with chance less than κ.

We must check that α, β and δ were picked correctly.

- β clearly satisfies inequality (2.2).

- Assuming that $n \geq 2$, we see that α satisfies inequality (2.3).

- It is easy to see that inequality (2.6) is satisfied.

- Invoking Lemma (4), $\hat{\phi} \geq \frac{\epsilon}{6174n^3}$.

- So $t = \frac{(6174)^2 n^6\left(7+3.2n+n\ln\left(\frac{1173n^2}{2\epsilon}+2\right)\right)}{\epsilon^2}$ steps are enough to draw a sample. Thus, each sample can be drawn in $O\left(\frac{n^7\log\left(\frac{n}{\epsilon}\right)}{\epsilon^2}\right)$ steps.

- Each step requires at most $O\left(log\left(\frac{n}{\epsilon}\right)\right)$ membership queries to compute the gauge function.

So the total number of membership queries is

$$O\left(\frac{n^7\left(\log\left(\frac{n}{\epsilon}\right)\right)^2 \log\left(\frac{\log\left(\frac{1}{\epsilon}\right)}{\kappa}\right)\log\left(\frac{\nu}{\epsilon}\right)}{\epsilon^2}\right), \tag{2.27}$$

which, if we ignore lesser log factors, can be thought of as $O^{\sim}\left(\frac{n^7\log(\nu)\log\left(\frac{1}{\kappa}\right)}{\epsilon^2}\right)$.

It is quite natural to wonder if an algorithm with this poor a runtime can possibly be an improvement on the ellipsoid algorithm. In [33] the version of the Yudin- Nemirovskiiĭ separation from membership algorithm requires $O^{\sim}(n^6)$ membership tests to build a single approximate separator and then the shallow cut ellipsoid algorithm may need $O^{\sim}(n^4)$ separators to optimize. It must be emphasized that this reference in *only* concerned with proving the problem is in P, so it is unlikely that $O^{\sim}(n^{10})$ membership queries is the best known. Finally it should be emphasized that the ellipsoid algorithm doesn't have a super logarithmic dependence on ϵ.

Remark: 1 *With a new result of Frieze, Kannan and Polson that the algorithm outlined here can have it's dependence on n brought down to n^6 by performing the sampling walk in a bounding sphere instead of a bounding box and estimating χ^* without introducing the idea of conductance. Other methods such as using newer chains available from Lovasz and Simonovits, or Kannan, Lovasz and Simonovits allow running times of $O(n^5)$ and possibly $O(n^3)$. None of these improvements are capable of replacing the dependence on ϵ with one on* $\log(\epsilon)$

2.7.2 With advantageous bounds.

Here we analyze the situation where good bounds x^l, L and x^u are known and attempt to lower the dependence of the runtime on n. To do this meaningfully, all dot products (and norms other than infinity) must be removed from the expressions as they hide n's. In this subsection, we work out a bound on run time that explicitly shows all of the powers of n.

We still assume that the problem has been rescaled so $c_i = 1$ for all i ($x_i \to c_i x_i$, $c_i \to \frac{c_i}{c_i} = 1$) and that

$$\frac{\min_i(x_i^u - x_i^f)}{\|x^u\|_\infty - \min_i x_i^l} \geq \frac{1}{2} \tag{2.28}$$

$$\text{and}\quad \min_i(x_i^u) + \min_i(x_i^l) \quad\geq\quad 2\left\|x^f\right\|_\infty. \tag{2.29}$$

This is easy to ensure by replacing x_i^u with $\left\|x^f\right\|_\infty + \|x^u\|_\infty$ and has the geometric interpretation of making the problem "well rounded".

We notice that if $\epsilon > \frac{\|x^f\|_\infty}{\min_i x_i^l} - 1$ then x^f is already a solution of the desired accuracy. This and inequality (2.29) imply

$$\frac{\|x^u\|_\infty - \min_i x_i^l}{\epsilon \min_i x_i^l} \geq 2. \tag{2.30}$$

Rather than the adaptive approach, first consider a sampling algorithm that comes within ϵ of optimal in one long walk:

- We set $L = \left\|x^l\right\|_1$ and $T = \epsilon\left\|x^l\right\|_1$.
- We will chose β to be at least $\frac{11}{10}\frac{n}{T}$ instead of as stated in inequality (2.2). The $\frac{11}{10}$ is required to guarantee that we get within T of the optimum instead of the $T + 2\delta\|c\|_1$ we could expect because of discretization. To guarantee this, we must show that $\frac{\epsilon}{10}\min_i x_i^l \geq 2n\delta$.
  - So we set
$$\beta = \frac{11}{10\epsilon \min_i x_i^l} \tag{2.31}$$
$$\text{and}\quad \alpha = \frac{5n(\|x^u\|_\infty - \min_i x_i^l)}{\epsilon \min_i x_i^l}. \tag{2.32}$$
  - By inequality (2.30), this satisfies inequality (2.3).
  - Now setting
$$\delta = \frac{\epsilon \min_i x_i^l}{70n} \tag{2.33}$$
    satisfies inequality (2.6), by inequality (2.28).
  - Clearly, we have $2n\delta \leq \frac{1}{10}\epsilon \min_i x_i^l$.

  So the $2\delta\|c\|_1$ factor has been dealt with.

Now, by Lemma (4) and Theorem (4), we have

$$\begin{aligned} \hat{\phi} &= \frac{\epsilon \min_i x_i^l}{210n^2 \left\| x^u - x^l \right\|_\infty} && (2.34) \\ t &= \left\lceil \left(7 + \frac{11n \left\| x^f - x^l \right\|_\infty}{10\epsilon \min_i x_i^l} + n \ln \left(\left\lceil \frac{70n \left\| x^u - x^l \right\|_\infty}{2\epsilon \min_i x_i^l} + 1 \right\rceil \right) \right) \left(\frac{210n^2 \left\| x^u - x^l \right\|_\infty}{\epsilon \min_i x_i^l} \right)^2 \right\rceil && (2.35) \end{aligned}$$

Which, if we take the middle term of the sum to be dominant, is

$$O\left(n^5 \left(\frac{\left\| x^u - x^l \right\|_\infty}{\epsilon \min_i x_i^l} \right)^3 \right) \tag{2.36}$$

steps.

We will call this algorithm **AlgorithmB**. **AlgorithmB** can then be repeated $\left\lceil \frac{\ln(\kappa)}{\ln(11/12)} \right\rceil$ times to amplify the chance of success to at least $1 - \kappa$.

The dependence on $\frac{\left\| x^u - x^l \right\|_\infty}{\epsilon \min_i x_i^l}$ can be improved by designing a new algorithm (**AlgorithmC**) that runs **AlgorithmB** in stages like we did for **AlgorithmA**.

The analysis is as before and the run time comes out to

$$O\left(n^5 \left(\frac{\left\| x^u - x^l \right\|_\infty}{\epsilon \min_i x_i^l} \right)^2 \left(\ln \left(\frac{n \left\| x^u - x^l \right\|_\infty}{\epsilon \min_i x_i^l} \right) \right)^2 \ln(\frac{1}{\kappa}) \right) \tag{2.37}$$

membership queries.

2.8 Errors Due to Discretization

For every $p \in J$ we need to determine a lower bound on ρ such that for some continuous monotone decreasing function ξ such that $\xi(0) = 0$ and all $\eta \geq 0$ in some open neighborhood of 0: if C is any rectangular region with center p contained in $C_{\delta+\eta}(p) \cap J$ then

$$\rho(1 - \xi(\eta)) \left(\frac{1}{Vol(C)} \int_C F \right) \leq F(p) \leq (\rho(1 - \xi(\eta)))^{-1} \left(\frac{1}{Vol(C)} \int_C F \right). \tag{2.38}$$

We will also need a lower bound on σ such that

$$\sigma F(x) \leq F(x + \lambda E^i) \tag{2.39}$$

for all $x \in K$, $i \in \{1 \cdots n\}$ and all $|\lambda| \leq 2\delta$.

For vectors a and b let ab be the vector $(ab)_i = a_i b_i$. To facilitate the analysis we break F into its two constituent parts f and g where $f(x) = e^{-\alpha(\max(\psi_{(K_L,x^u)}(x)-1,0))}$, $g(x) = e^{-\beta c \cdot x}$, and $F(x) = f(x)g(x)$.

σ is easy to deal with when we apply the well known fact that a gauge function based on K can fall no faster than one based on a convex subset of K (containing x^u). To be precise we apply Corollary 52 from [5] to get:

$$|\psi_{(K_L,x^u)}(x) - \psi_{(K_L,x^u)}(y)| \leq \frac{\|x-y\|_\infty}{\min_i(x_i^u - x_i^f)}$$

which implies

$$e^{-\alpha(\delta+\eta)/(\min_i(x_i^u - x_i^f))} f(x) \leq f(z) \quad \forall z \in C \tag{2.40}$$

and inequality 2.39 is satisfied with

$$\sigma = e^{-2\alpha\delta/(\min_i(x_i^u - x_i^f))} e^{-2\beta\delta\|c\|_\infty}. \tag{2.41}$$

To get a lower bound on ρ we use the simple rule that for $f, g \geq 0$

$$\min_{x\in C}(f(x)) \int_C g \leq \int_C fg \leq \max_{x \in C}(f) \int_C g$$

and use inequality 2.40 to get the point-wise bounds on f. All that remains is to derive a lower bound ρ' such that

$$\rho' \left(\frac{1}{Vol(C)} \int_C g \right) \leq g(p) \leq \rho'^{-1} \left(\frac{1}{Vol(C)} \int_C g \right). \tag{2.42}$$

We note that g is of the form $g(x) = h(c \cdot x)$ for some non-negative convex function h in the region we are interested in and since C is symmetric about p we know that (from Jensen's inequality)

$$g(p) \leq \frac{1}{Vol(C)} \int_C g,$$

and any $\rho'^{-1} \leq 1$ satisfies the right side of inequality 2.42.

To get the left side of inequality 2.42 we define

$$\begin{aligned} \Xi_i &\stackrel{\text{def}}{=} \max\left\{\lambda \in \mathbf{R} \;\middle|\; p + \lambda E^i \in C\right\} \\ D &\stackrel{\text{def}}{=} \{x \in \mathbf{R}^n \mid |x_i - c_i p_i| \le c_i \Xi_i \; \forall i\} \end{aligned}$$

and change variables to get:

$$\frac{1}{\prod_{i=1}^n 2c_i\Xi_i} \int_D g.$$

Let $\mu(\lambda)$ be the measure of the set

$$\left\{x \in D \;\middle|\; x \cdot \vec{1} - p \cdot c \ge \lambda\right\}$$

It is easy to see μ is differentiable and $-\mu'(\lambda) \ge 0$ for $\lambda \in [0, c \cdot \Xi]$. We return to our integral (using the estimate $\mu(\lambda) - \mu(\lambda + d\lambda) = -\mu'(\lambda)d\lambda$):

$$\begin{aligned} &= \frac{1}{\prod_{i=1}^n 2c_i\Xi_i} \int_0^{c\cdot\Xi} -\mu'(\lambda) \left(e^{-\beta(c\cdot p+\lambda)} + e^{-\beta(c\cdot p-\lambda)}\right) \mathrm{d}\,\lambda \\ &= g(p) \frac{-2}{\prod_{i=1}^n 2c_i\Xi_i} \int_0^{c\cdot\Xi} \cosh(\beta\lambda)\mu'(\lambda)\,\mathrm{d}\,\lambda \\ &= g(p) \frac{-2}{\prod_{i=1}^n 2c_i\Xi_i} \left(\cosh(\beta\lambda)\mu(\lambda)\big|_{\lambda=0}^{\lambda=c\cdot\Xi} - \beta \int_0^{c\cdot\Xi} \sinh(\beta\lambda)\mu(\lambda)\,\mathrm{d}\,\lambda\right) \\ &= g(p) \frac{-2}{\prod_{i=1}^n 2c_i\Xi_i} \left(0 - \frac{\prod_{i=1}^n 2c_i\Xi_i}{2} - \beta \int_0^{c\cdot\Xi} \sinh(\beta\lambda)\mu(\lambda)\,\mathrm{d}\,\lambda\right) \\ &= g(p) \left(1 + \frac{2\beta}{\prod_{i=1}^n 2c_i\Xi_i} \int_0^{c\cdot\Xi} \sinh(\lambda)\mu(\lambda)\,\mathrm{d}\,\lambda\right) \end{aligned}$$

By Theorem 2 of [36], we have $\mu(\lambda) \le \left(\prod_{i=1}^n 2c_i\Xi_i\right) e^{-\lambda^2/(2\|c\Xi\|_2^2)}$. So continuing we have:

$$\frac{1}{Vol(C)} \int_C g \le g(p) \left(1 + 2\beta \int_0^{c\cdot\Xi} \sinh(\lambda) e^{-\lambda^2/(2\|c\Xi\|_2^2)}\,\mathrm{d}\,\lambda\right)$$

We need an upper bound for

$$2\beta \int_0^{c\cdot\Xi} \sinh(\lambda) e^{-\lambda^2/(2\|c\Xi\|_2^2)}\,\mathrm{d}\,\lambda$$

which comes out to (by standard integral tables)

$$
\begin{aligned}
& \beta \left\|c\Xi\right\|_2 \sqrt{\frac{\pi}{2}} e^{\beta^2 \|c\Xi\|_2^2/2} \left(2\,\mathrm{erf}\left(\frac{\beta \left\|c\Xi\right\|_2}{\sqrt{2}}\right) + \mathrm{erf}\left(\frac{c\cdot\Xi - \beta\left\|c\Xi\right\|_2^2}{\sqrt{2}\left\|c\Xi\right\|_2}\right) - \mathrm{erf}\left(\frac{c\cdot\Xi + \beta\left\|c\Xi\right\|_2^2}{\sqrt{2}\left\|c\Xi\right\|_2}\right)\right) \\
\leq \quad & \sqrt{2\pi}\beta\left\|c\Xi\right\|_2 \mathrm{erf}\left(\frac{\beta\left\|c\Xi\right\|_2}{\sqrt{2}}\right) e^{\beta^2\|c\Xi\|_2^2/2} \\
\leq \quad & \sqrt{2\pi}\beta(\delta + \sqrt{n}\eta)\left\|c\right\|_2 \mathrm{erf}\left(\frac{\beta(\delta+\sqrt{n}\eta)\left\|c\right\|_2}{\sqrt{2}}\right) e^{\beta^2(\delta+\sqrt{n}\eta)^2\|c\|_2^2/2}
\end{aligned}
$$

which for sufficiently small $\eta > 0$ (and $\eta = 0$) is

$$
\leq \sqrt{2\pi}\beta\delta\left\|c\right\|_2 \mathrm{erf}\left(\frac{\beta\delta\left\|c\right\|_2}{\sqrt{2}}\right) e^{\beta^2\delta^2\|c\|_2^2/2}(1+o(1))
$$

combining this with our point-wise bound on f we get

$$
\rho \geq \frac{e^{-\alpha\delta/(\min_i(x_i^u - x_i^f))}}{1 + \sqrt{2\pi}\beta\delta\left\|c\right\|_2 \mathrm{erf}\left(\frac{\beta\delta\|c\|_2}{\sqrt{2}}\right) e^{\beta^2\delta^2\|c\|_2^2/2}}
$$

2.9 Relation to Simulated Annealing

The above algorithm has some similarities with simulated annealing.[45] In simulated annealing optimization is performed by moving from feasible state to feasible state by choosing moves from small neighborhoods. The moves are selected and rejected in a manner almost identical to the Metropolis filter as applied to our function. The fact that our algorithm steps through infeasible states is not very important. This it is an artifice introduced to improve the mixing rate of the Markov chain by removing small sets from consideration. The constants in any simulated annealing problem can also be hidden in the objective function (similar to Lagrangian relaxation). The second point of similarity is that where our algorithm moves its upper bound (by finding a sufficiently better new best point) or lower bound (by building a probalistic proof) is exactly like altering the “temperature” parameter in the Boltzmann equation that selects moves in simulated annealing. In fact, though we have not found any advantage to it, we could redesign our algorithm to alter β and γ continuously to imitate simulated annealing. Unfortunately this would mean that the chain is no longer stationary (as the chain’s transition behavior would change over time). One could make crude arguments that the chains behavior over any time period would be no worse than if it had been run with the parameters at the end of the time period. Then one could appeal to a Markov chain central limit theorem to show that the "tip" would still be visited with high frequency.

It is a very interesting open problem to try and extend the theory of rapidly mixing Markov chains to the stochastic processes like simulated annealing.

2.10 Small Sets

The treatment of rapidly mixing Markov chains given here completely avoided the "small set" issue by allowing the chain to move to infeasible states and then applying a punishment factor to return the chain to the feasible region. This had the disadvantage of introducing a hard to characterize gauge style function depending on α. This forced us to make δ very small so we could apply crude point-wise bounds on this gauge function, whereas for the bias function we were able to make more sophisticated estimates of variation over regions (via the Hoeffding inequality) that would have allowed δ to be a factor of $\sqrt{n}$ larger and shaved a factor of n off the running time of the Markov chain.

The small set problem is that if we have a Markov chain on a set of points chosen from a convex body that conductance is poor near the boundary of the body. In fact if the chain uses only states that are such convex bodies it may not even be connected. An example is $K = \left\{(x,y) \in \mathbf{R}^2 \;\middle|\; y \geq 0,\ y \leq \frac{3}{2}x,\ y \geq 3x - 3\right\}$. $Z^n \cap K = \{(0,0),(1,0),(1,1),(2,3)\}$ which isn't connected by the King's moves that changes a single coordinate by ± 1. We are certainly not going to be able to get a strong conductance result for this chain as the set $\{(2,3)\}$ is inescapable. There are a number of ways to deal with the above problem. The one we used is to walk on a large region of Z^n instead of $Z^n \cap K$ and try to force the chain into K with a penalty function. This has the advantage of simplifying the structure of the chain and the disadvantage of introducing a hard to analyze penalty function.

Lovasz and Simonovits [40] dealt with this problem by introducing a notion called s-conductance. The s-conductance of a Markov chain is not the worst escape probability of any set of states in the chain but the worst escape probability of any large set of states in the chain. The authors then perform a fairly delicate analysis to show that analogues of the convergence theorems known for conductance are true for s-conductance. Nothing comes for free so some care must be made that the chain does not start in a small set. In addition to the improved mixing rate of the Markov chain described in this paper this paper is significant because it formalizes many of the geometric methods used in this field into well contained lemmas (such as "localization").

Frieze, Kannan and Polson [28] deal with the problem in a similar manner. They use a strong minimax characterization of the eigenvalues of the linear operator representing the Markov chain. In this characterization the convergence rate of the Markov chain basically the gap between the eigenvalue 1 and the next largest eigenvalue (the separation of eigenvalues from -1 is usually easy to show). By the minimax characterization the second largest eigenvalue is the eigenvalue corresponding the vector ϕ the maximizes the Rayleigh quotient $\phi^T P\phi/\phi^T\phi$ subject to $\phi^T\pi = 0$ where π is the stationary distribution. It turns out that similar convergence theorems can be proven for the vector that maximizes this quotient subject to the additional restrictions that it places no mass in any of the states near the boundary of K. The restrictions are easily expressed as linear relations and lead to an analysis that has similar consequences to the s-conductance arguments. Again the restricted result is weaker that the classical result and care must be taken that one does not start near any of the bad states. The direct handling of eigenvalues and eigenvectors, without explicit mention of conductance, are likely to yield many more results in the near future.

A last concept that is in development by Kannan, Lovasz and Simonovits is average local conductance. Local conductance, introduced by Lovasz and Simonovits, is the chance that $X_t \neq X_{t+1}$ or that the chain makes a non-trivial transition. Average local conductance is the expected value of $P[X_t \neq X_{t+1} | X_t \text{ is } \pi \text{ distributed}]$. Most current analysis of Markov chains state how many steps the chain must run for some property to hold. Average local conductance results are theorems of the form "after so many non-trivial transitions some property holds." These results are much more powerful as they not only allow one much more freedom in trying to estimate how bad states affect a Markov chain (one can show there are not many of them, or one is unlikely to visit them) but one also has the option of proving nothing about average local conductance and running chains until one observers empirically that the requisite number of non-trivial transitions have been taken.

2.11 Stopping time.

Our actual implementation, while not practical for general use, allowed us to empirically observe an expected behavior of the Markov chain. Since all our proofs of stopping time must, to be correct, be pessimistic. The chain will improve its "best" point much faster that the analysis indicates and it is worth the extra effort to continuously monitor the chain and restart it at a new best point when any significant fraction of improvement is noticed in objective value. With this modification the chain very quickly restricts to a small neighborhood of the true optimal point and then only has to run to the theoretical stopping time one set of times: to prove the lower bound. In practice almost all of the running time of this algorithm is used in proving the final lower bound.

2.12 King's move data structures.

Using the King's move has a number of implementation advantages over using a continuous move structure. First many fewer random bits are needed as we only need log(n) coin flips to decide which direction to go and some constant number of coin flips to decide whether to go or not (as we have restricted F to vary by no more than a constant). We can also update our state in constant time (change one entry in a vector). We have also found that membership in K and be checked faster than is the case for more general moves. If K is given by m linear inequalities then by incrementally updating all of the linear relations we can check membership in $O(m)$ time instead of $O(mn)$.

We have tried building a membership oracle for the component commonality problem by simulating the numerical integrations required to test membership in K by using sums over historic data. K is defined as the set of $x \in \mathbf{R}^m$ such that $\int_{y \in \mathbf{R}^n,\, Ay \leq x} F(y) \geq \gamma$ where F is the probability density function for the month's total orders. We approximate K with the non-convex body $K_s = \left\{ x \in \mathbf{R}^m \;\middle|\; \frac{1}{s} \sum_{i=1,\, Ay_i \leq x}^{s} 1 \geq \gamma \right\}$ where $y_1, \cdots, y_s$ are points drawn from $\mathbf{R}^n$ according to the density F. We are using the standard statistical trick of using the "empirical distribution" derived from previous observations to simulate the true distribution. We know by central limit arguments that $\lim_{s \to \infty} K_s = K$. We, of course, have the problem that the empirical distribution is always atomic and K_s is almost never convex. It was shown by Applegate and Kannan [5] that small departures from true convexity do not overwhelm the Markov chain technique.

Notice that by organizing multiple copies of $y_1, \cdots, y_s$ each copy sorted by a different coordinate we can compute membership in K_s in time $O(s)$ instead of $O(ms)$.

These savings are very important since the Markov chain we use the membership oracle every step and the membership oracle performs, by far, the most expensive operation each step.

Chapter 3

Contingency Tables: Exact and Approximate Counting

3.1 Background.

Consider an m by n matrix of non-negative integers. This matrix, called a contingency table is an important summary tool in statistics and arises in the following manner. Consider a finite multi-set[1] with elements (i, j) ($a \in [1 \cdots m],\ b \in [1 \cdots n]$). For each such (i, j) let $\pi_{i,j}$ denote the probability that an element drawn out of the multi-set with uniform probability is equal to (i, j). If such a multi-set represented a population of people that are identified only by hair color (Black, Brunette, Red, Blonde) and eye color (Brown, Blue, Hazel, Green) then $\pi_{\text{Black,Green}}$ would be the probability that a person picked from this population uniformly at random would have both black hair and green eyes. A possible experiment would be to draw a multi-set sample, S, of T elements out of the original multi-set (for simplicity assume we draw with replacement). We could then summarize this experiment in an m by n matrix X where

$$X_{i,j} \stackrel{\text{def}}{=} \text{the number of occurrences of } (i, j) \text{ in } S. \tag{3.1}$$

Define $\pi_{i,*} = \sum_{j=1}^{n} \pi_{i,j}$ and $\pi_{*,j} = \sum_{i=1}^{m} \pi_{i,j}$. We consider the distribution to be independent if $\pi_{i,j} = \pi_{i,*}\pi_{*,j}\ \forall i, j$. Of immediate statistical interest is testing if the original population is independent, given an observed sample X. As is often

[1] A set the allows repetitions.

the case in statistics we can not, without additional information, in the form of priors, determine if the "Null Hypothesis" $H_1 : \pi_{i,j} = \pi_{i,*}\pi_{*,j}\ \forall i,j$ is likely to be true or false. Instead we calculate a significance or how likely X's deviation from independence would be if H_1 were true. If X is too far from independent ($X_{i,j}$ too far from $X_{*,i}X_{*,j}/T$) then we may suspect that the original population was not H_1 distributed.

3.2 Statistical hypothesis testing.

The obvious method for testing if X is typical, assuming H_1, would be to use the $X_{i,j}/T$ as observed estimates for the $\pi_{i,j}$ and then to check if the independence relations approximately hold. This requires estimating $(n-1)(m-1)$ parameters from the experiment. A slightly more subtle approach is to let $r_i = X_{i,*}$ and $c_j = X_{*,j}$ and look at the, unknown, distribution restricted to Ys such that $Y_{i,*} = r_i$ and $Y_{*,j} = c_j$. The density function for any population obeying H_1 restricted to the row and column conditions is independent of all of the, unknown, $\pi_{i,j}$s. If fact $P(Y|r,c)$ (the probability of observing the table non-negative integral table Y that has row/column totals matching r and c is the so called Fisher/Yates distribution:

$$P(Y|r,c) = \frac{\frac{T!}{\prod_{i=1}^{m}\prod_{j=1}^{n}(Y_{i,j}!)}}{\left(\frac{T!}{\prod_{i=1}^{m}(r_i!)}\right)\left(\frac{T!}{\prod_{j=1}^{n}(c_j!)}\right)}. \tag{3.2}$$

Thus for two different distributions π^1, π^2 on the original population that obey the independence hypothesis H_1 we have for any two tables X, Y such that $X_{i,*} = Y_{i,*}$ and $X_{*,j} = Y_{*,j}$ we have $P_{\pi^1}(X)/P_{\pi^1}(Y) = P_{\pi^2}(X)/P_{\pi^2}(Y)$ even though $P_{\pi^1}(X)$ may be very different from $P_{\pi^2}(X)$.

Because of this observation it is considered good practice (c.f. [16]) to perform the desired significance tests in the subset of contingency tables that have $Y_{i,*} = r_i$ and $Y_{*,j} = c_j$. Given this restriction we see that measuring the deviation of X from independence is equivalent to measuring the deviation of X from the intersection of the independence surface and the subspace of tables that obey the row and column restrictions. The surface and restriction subspace intersect at exactly one point Y such that $Y_{i,j} = X_{i,*}X_{*,j}/T$ so we have again reduced the hypothesis H_1 to a single table (instead of a surface of tables) but this time without estimating $(n-1)(m-1)$ parameters. This method is not very different from the obvious one given above but does permit a sharper analysis (and tighter bounds).

It is easy to show the above density function is log-concave (replace $x!$ with $\Gamma(x+1)$ everywhere and notice that $\psi^{(1)}(x) = \frac{d^2[\ln(\Gamma(x))]}{dx^2}$ is non-negative for $x > 0$ [1] 6.4.10). Thus we know that if the density will fall off at least as fast as some exponential as distance from the mode decreases. It can be shown that in a wide variety of circumstances that the mode is very near the independent table[2] so the most important component of significance is going to be distance from the independent table.

A natural candidate for this metric is the "chi-square" distance which is defined as:

$$\chi^2(Y) = \sum_{i,j} \frac{\left(Y_{i,j} - \frac{X_{i,*}X_{*,j}}{T}\right)^2}{\frac{X_{i,*}X_{*,j}}{T}}. \tag{3.3}$$

The χ^2 of a table Y is just the square distance of Y from independent where each coordinate is rescaled by a factor of $X_{i,*}X_{*,j}/T$. A natural significance test for Y is then to measure the proportion of the Fisher-Yates distribution corresponding to tables with χ^2 higher than Y. If very few tables have χ^2 higher than Y we should know that Y is a unlikely table under the independence hypothesis H_1.

3.2.1 Fisher-Yates test

There are a lot of asymptotic methods for computing the significance (with respect to the Fisher-Yates distribution) of a contingency table [42, 30, 6, 7]. Most of these techniques rely on the central limit theorem and normal approximations to the Fisher-Yates distribution. These method are often quite far off for moderate sized tables.

Fortunately there is an obvious pseudo polynomial scheme for generating such tables from the Fisher-Yates distribution. So randomized approximation schemes are easy to implement. Consider the complete bipartite graph with vertex sets V_1, V_2 ($|V_1| = |V_2| = T$).

$$\begin{aligned} V_1 &\stackrel{\text{def}}{=} \{(i,k) \mid i = 1 \cdots m,\ k = 1 \cdots r_i\} \\ V_2 &\stackrel{\text{def}}{=} \{(j,l) \mid j = 1 \cdots n,\ l = 1 \cdots c_j\} \end{aligned}$$

[2] The mode isn't always at the independent table. Consider a tables with row sums [1 1 1 1 1 1 10] and column sums [1 1 1 1 1 1 10]. The independent table has density about 0.006 while the table Y with $Y_{i,j} = 0$ for $i, j < 7$, $Y_{i,7} = 1$ for $i < 7$, $Y_{7,j} = 1$ for $j < 7$ and $Y_{7,7} = 4$ has density about .026. These two tables differ by more than 2 in the bottom right entry.

Select a perfect matching M uniformly at random. Notice that there are $T!$ such perfect matchings (vertices are distinguishable, edges are not). Now associate with M the following table:

$$X_{i,j} = |\text{edges of the form } ((i,k),(j,l)) \text{ for any } k,l|$$

Notice X obeys the row/column sums r, c and there are exactly $\frac{(\prod_{i=1}^{m}(r_i!))\left(\prod_{j=1}^{n}(c_j!)\right)}{\prod_{i=1}^{m}\prod_{j=1}^{n}(X_{i,j}!)}$ different perfect matchings that map to this X. Thus uniform M leads to Fisher-Yates distributed X.

3.2.2 Uniform test

A major problem with the Fisher-Yates distribution is that it is very tightly concentrated about its mode. This causes to Fisher-Yates test to reject almost all tables when T is large. For example the Fisher-Yates test will reject all tables where $X_{i,j}$ differs from the independent value $X_{*,j}X_{i,*}/T$ by more than $\sqrt{X_{*,j}X_{i,*}/T}$. This guarantees rejection if any constant percentage of the classifications are in error, no matter how small, is present in the data. Diaconis and Efron [16] suggest a number of techniques to deal with this problem including a significance test based on the uniform distribution. For this test the significance is the ratio of the number of contingency tables with χ^2 greater than $\chi^2(Y)$ (and identical row and column sums) to the total number of tables with the given row and column sums = $\sharp(r,c)$. This test has the nice property that it depends on the structure of the table Y much more strongly than it depends on T. The difficulty is that computing the numerator of this ratio has been shown to be $\sharp P$ hard in general.[24] Another quantity of interest is the "likelihood ratio statistic" which is the ratio of how likely a table is under the Fisher-Yates distribution to how likely the table is under the uniform distribution ($= P(Y|r,c)/(1/\sharp(r,c))$). This quantity is useful in making qualitative statements about a table. This can be important because if one decides to reject the independence hypothesis H_1 one often still needs to characterize the table in question.

3.2.3 Some counting preliminaries

We now investigate $\sharp(r,c)$ in its own right. Good references for this problem are [18, 16, 20].

When we have $m = n$ and $r = c = k\vec{1}$ the problem reduces the the classic problem of counting the number of magic squares. The counting function turns out to be polynomial of degree $(m-1)(n-1)$ in k. The highest degree coefficient of this polynomial is the volume of the transportation polytope, a quantity of independent interest. Blakley[13] showed that $\sharp(r, c)$ is a piecewise polynomial (in r, c) of degree $(m-1)(n-1)$ and [15] reproves this result. Sturmfels strengthens these results [51, 52] and we have reworked his proofs to create an effective procedure for fixed m, n.

From [52] a generating function for $\sharp(r, c)$ is:

$$\prod_{i=1}^{m}\prod_{j=1}^{n-1}\frac{1}{1-x_iy_j} = \sum_{r_1\cdots r_m, c_1\cdots c_{n-1}\in Z, c_n=\sum_{i=1}^{m} r_i - \sum_{j=1}^{n-1} c_j;\, r,c\geq 0} \sharp(r;c)\left(\prod_{i=1}^{m} x_i^{r_i}\right)\left(\prod_{j=1}^{n-1} y_j^{c_j}\right). \tag{3.4}$$

3.3 Barvinok's algorithm.

Barvinok [8] recently published a polynomial time algorithm for counting the number of integer points in any integral polytope in fixed dimension. This result implies that counting the number of contingency tables for fixed m, n is solvable in polynomial time. Barvinok's result reduces the counting problem to solving a family of sums over the lattice points in a set of cones. This counting scheme, if applied directly, would run in time proportional to the the index, with respect to the standard lattice, of the integer lattice intersected with these cones. In general these indices (also called the cardinality of the glue group in Conway and Sloan) can be double exponential in the size of presentation.

For any simple rational cone K in $\mathbf{R}^d$ let χ_K be the indicator function of K. That is $\chi_K(x) = 1$ if $x \in K$ and 0 otherwise. Since K is a simple rational cone there is a linear independent set of integral vectors $u_1, \cdots u_k$ that generate K. Ind(K) is then the size of the quotient group $(Z^n \cap \text{Lin}(u_1, \cdots, u_k))/ < u_1, \cdots, u_k >$. Barvinok identifies a scheme which finds cones $K_1, \cdots, K_s$ and integers $\epsilon_1, \cdots, \epsilon_s$ such that $\chi_K = \sum_{i=1}^{s} \epsilon_i \chi_{K_i}$ and Ind$(K_i) \leq$ Ind$(K)^{(d-1)/d}$ where d is the, fixed, dimension of space. Thus in only a log-log number of levels of recursion the problem can be reduced to problems involving only cones with constant index. This method has not been directly applied to contingency tables, though the unimodular nature of the contingency table relations may admit a number of improvements over the general algorithm.

3.4 Piecewise polynomialality of the number of contingency tables.

Here we give a simple proof of the fact that the number of contingency table is piecewise polynomial function of row and column sums where the number of pieces is at most $(nm)^{(n+m)^2}$ for n by m contingency tables. The proof is not specialized to contingency tables, but applies to any counting problem involving totally unimodular constraints.

Take A, a totally unimodular d by w matrix of rank d such that the only solution to $Ax = 0,\ x \geq 0$ is $x = 0$. For any $v \in \mathbf{R}^d$ let P_v be the polytope defined by

$$P_v = \{x \in \mathbf{R}^w \ \mid x \geq 0 \ , \ Ax = v\} \, . \tag{3.5}$$

We call

$$\sharp(v) \stackrel{\text{def}}{=} |P_v \cap Z^w| \tag{3.6}$$

the counting function of A.

Theorem 6 *If A is as above then the counting function of A is a piecewise polynomial with no more than $d\binom{d\binom{w}{d}}{d}$ pieces and total degree equal to $w - d$.*

This is not a new result (this was known to [15] and much stronger results are known by Sturmfels [51, 52, 11]) but the presentation given here is, hopefully, more approachable and is useful in describing the algorithms. In addition to proving the counting function is piecewise polynomial we demonstrate a poly time (for fixed dimension) method of computing the decomposition into pieces and techniques for inferring the polynomials. These methods can be strengthened to deal with the non-unimodular case but run in time proportional to the indices of various sub lattices in the standard lattice. Before we prove the theorem we present some geometric lemmas.

We wish to find conditions on u, v so that the operation of Minkowski addition (point-wise adding two sets) and adding right hand sides of linear relations are identical, or

$$P_u + P_v = P_{u+v}. \tag{3.7}$$

Remark: 2 *Clearly $P_u + P_v \subseteq P_{u+v}$ but this containment can be strict, example:*

$$A \;=\; \left\{ \begin{array}{cccc} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array} \right\}$$

$$
\begin{aligned}
u &= \left\{ \begin{array}{cc} 1 & 99 \end{array} \right\}^\top \\
v &= \left\{ \begin{array}{cc} 99 & 1 \end{array} \right\}^\top
\end{aligned}
$$

$P_u + P_v \sim \left\{(x, y) \in \mathbf{R}^2 \mid x \geq 0,\ y \geq 0,\ x + y \leq 2,\ x + 2y \leq 3\right\}$ *while* $P_{u+v} \sim \left\{(x, y) \in \mathbf{R}^2 \mid x \geq 0,\ y \geq 0,\ x + 2y \leq 100\right\}$.

Let $E^i \in \mathbf{R}^w$ be the standard basis vector such that $E^i_j = \delta_{i,j}$. Let x be an arbitrary vertex of P_u. At least $w - d$ of the entries of x must be zero and there is a $\sigma \subseteq \{1, 2, \cdots w\}$ such that $|\sigma| = d$, $x_i = 0$ for $i \notin \sigma$, and if A' is A with the $w - d$ rows $E^i, i \notin \sigma$ added on then A' is rank w. For matrices A let A_σ denote the d by d matrix formed by taking (in order) the columns A_i s.t. $i \in \sigma$, for vectors x let x_σ denote the d-vector corresponding the the entries whose indices are in σ. In linear programming terms A_σ is a basis and σ is the indices of set of columns corresponding in this basis. We see that if $Ax = u$ then we must have $A_\sigma x_\sigma = u$ and $|\det(A_\sigma)| = |\det(A')|$. So A_σ is full rank and we see that $x_\sigma = A_\sigma^{-1}u$, so x is completely determined by σ (or, again in linear programming terms, the naming of the basis determines a basic solution).

Define the sets of indices corresponding to basic solutions:

$$B = \{\sigma \subseteq \{1, 2, \cdots w\} \mid |\sigma| = d,\ \det(A_\sigma) \neq 0\}. \tag{3.8}$$

And further define

$$R = \left\{x \in \mathbf{R} \;\middle|\; \exists \sigma \in B,\ x \text{ is a row of } A_\sigma^{-1}\right\} \tag{3.9}$$

as the "important" relations of the basic solutions.

For $\sigma \in B$ and $u \in \mathbf{R}^d$ let $\nu(\sigma, u)$ be the point in $\mathbf{R}^w$ such that $\nu(\sigma, u)_i = 0, i \notin \sigma$ and $(\nu(\sigma, u))_\sigma = A_\sigma^{-1}u$. We associate with each u a function $\chi_u : \mathbf{R}^w \to 0, 1$ defined such that

$$\chi_u(r) = \left\{ \begin{array}{ll} 1 & r \cdot u \geq 0 \\ 0 & \text{otherwise} \end{array} \right. . \tag{3.10}$$

It should be clear that

$$\text{vertices}(P_u) = \left\{\nu(\sigma, u) \;\middle|\; \sigma \in B,\ \ \chi_u(r) = 1 \text{ for all rows } r \text{ of } A_\sigma^{-1}\right\}.$$

Lemma 6 *If P_u and P_v are pointed and bounded and $\chi_u = \chi_v$ then $P_u + P_v = P_{u+v}$.*

Proof: If P_u or P_v is empty then they both must be empty and the theorem is true. Because P_u and P_v are both pointed and bounded we know that P_{u+v} is pointed and bounded. Then, since we only need to show $P_{u+v} \subseteq P_u + P_v$, it is sufficient to show if z is a vertex of P_{u+v} then there exists $x \in P_u$ and $y \in P_v$ such that $z = x + y$. It is easy to see that $\chi_{u+v} = \chi_u (= \chi_v)$ because $r \cdot (u + v) = r \cdot u + r \cdot v$ and sign$(r \cdot v) =$ sign$(r \cdot u)$ by assumption. So we find a $\sigma \in B$ such that $\nu(\sigma, u + v) = z$ and we know that $x = \nu(\sigma, u) \in P_u$ and $y = \nu(\sigma, v) \in P_v$. $\square$

Corollary 1 *There exists a set of full dimensional cones $C_1, C_2 \cdots C_r$ in $\mathbf{R}^d$ with the positive orthant covered by $\bigcup_{i=1}^r C_i$ such that if $u, v \in C_j$ then $P_u + P_v = P_{u+v}$ and $r \leq d\binom{d\binom{w}{d}}{d}$.*

Proof: We see that the linear relations in R divide the $\mathbf{R}^d$ space of right hand sides into pieces where Minkowski addition of polytopes and vector addition of right hand sides operate identically. $|B| \leq \binom{w}{d}$, $|R| \leq d|B| \leq d\binom{w}{d}$. It is a well known fact that k hyperplanes can split $\mathbf{R}^d$ into at most

$$\sum_{i=0}^{d} \binom{k}{i}$$

regions. We have $k \leq d\binom{w}{d}$ so we will have at most

$$\sum_{i=0}^{d} \binom{d\binom{w}{d}}{i} \leq d\binom{d\binom{w}{d}}{d}$$

full dimensional cones. $\square$

We will use Σ to denote the set of full dimensional cones (often called a “fan” when extended to include all lower dimensional cones). For m by n contingency tables we have $d = m + n - 1$ (one of the row totals is dependent on the other row and column totals) and $w = mn$. We then have the easy upper bound of $(mn)^{(m+n)^2}$ cones. We have completed the geometric preliminaries and, after a couple lemmas about polynomial arithmetic, are ready to prove the main theorem. We now demonstrate that the counting function is a low degree polynomial when restricted to any cone $C \in \Sigma$.

Lemma 7 *Suppose $\bar{p} : Q^k \to Q$ is a rational polynomial of degree s in $x_1, \cdots, x_k$, M is a $d \times k$ rational matrix of rank d and $\forall x \geq 0$ integral, $\forall y$ integral such* *that $My = 0$ and $x + y \geq 0$ we have $\bar{p}(x) = \bar{p}(x + y)$ then there is a unique rational polynomial p of degree s such that the following diagram commutes:*

$$\begin{tikzcd}
Q^k \arrow[r, "M"] \arrow[dr, "\bar{p}"'] & Q^d \sim Q^k/M \arrow[d, dashed, "p"] \\
& Q
\end{tikzcd}$$

Proof: First we prove $\forall x, y \quad My = 0 \rightarrow \bar{p}(x) = \bar{p}(x + y)$. We relax the sign conditions on x and $x + y$ and the integrality of x. Suppose $y \in Z^k$ is such that $\forall x \in Z^k \geq 0 \;\wedge\; x + y \geq 0 \quad \rightarrow \bar{p}(x) = \bar{p}(x + y)$. Pick b such $b + y \geq 0$ and define $p_1(x) = \bar{p}(x + b)$ and $p_2(x) = \bar{p}(x + b + y)$. We see that $\forall x \in Z^k \geq 0 \quad p_1(x) = p_2(x)$ which is enough to establish that they are identical polynomials. Thus we have if y is integral such that $My = 0$ then $\forall x \quad \bar{p}(x) = \bar{p}(x + y)$. Now pick any $y \in Q^k$ such that $My = 0$. Define $p_3(j) = \bar{p}(x + jy)$. p_3 is a degree s polynomial in j and we can easily find $s + 1$ values of j, $j_1, j_2 \cdots j_{d+1}$, such that $j_i y$ is integral. We have $p_3(j_i)$ is constant at these points and therefore is equal to the constant $\bar{p}(x)$ everywhere. In particular $\bar{p}(x) = p_3(0) = p_3(1) = \bar{p}(x)$.

It is clear that there is a unique function $p : Q^d \rightarrow Q$ that completes the diagram, so it remains only to show that p is a rational polynomial of degree no more than s. Pick $l_{i_1}, \cdots, l_{i_d}$ linearly independent columns of M and let $L = (l_{i_1}, \cdots, l_{i_d})$ then we see $\forall v \in Q^d \; p(v) = \bar{p}(L^{-1}v)$. $\Box$

We now state, without proof, an important result in the geometry of numbers (see [34] pp. 135-140).

Theorem 7 (McMullen) *If $P_1, P_2, \cdots P_k$ are integral polytopes in $\mathbf{R}^s$ then for non-negative $\lambda \in Z^k$ then $\sharp \left(\sum_{i=1}^{k} \lambda_i P_i\right)$ is a polynomial in $\lambda_1, \lambda_2, \cdots \lambda_k$ of total degree no more than s.*

Theorem 8 *If A is totally unimodular then for each $C \in \Sigma$ there exists p_C a polynomial of degree no more than $w - d$ in $v_1 \cdots v_d$ such that $\forall v \in C \cap Z^d$ the number of integer points in P_v is equal to $p_C(v)$.*

Proof: C is a full dimensional rational convex polyhedral cone, so by Gordon's Lemma or the Hilbert Basis Theorem [29] (pp. 11-12) we know that the semigroup of integer vectors in C is finitely generated, say with generators $c^1, \cdots, c^k \in C$. Because A is totally unimodular we know from [34] (pp. 135-140) there is a polynomial $\bar{p}_C$ of degree $w - d$ in variables $x_1 \cdots x_k$ such that if $v \in Z^d$, $x \in Z^k$ and $v = \sum_{i=1}^{k} x_i c^i$ ($x_i \geq 0$) then the number of lattice points in P_v is equal to $\bar{p}_C(x)$. If $\sum_{i=1}^{k} x_i c^i = \sum_{i=1}^{k} y_i c^i$ then we must have $\bar{p}_C(x) = \bar{p}_C(y)$ and we see that we meet the conditions of lemma 7. with the polynomial $\bar{p}_C$ and matrix $M = (c^1, \cdots, c^k)$. Thus there is a polynomial p_C of degree no more than $w - d$ in variables $v_1, v_2 \cdots v_d$ such that $p_C(v) = \sharp(P_v)$ for all $v \in C$. $\square$

For contingency tables we have A is the $(m + n - 1) \times (mn)$ matrix where

$$A_{i,j} = \begin{cases} \delta_{i-1,\lfloor (j-1)/m \rfloor} & i \leq m \\ \delta_{i-m,j-n\lfloor (j-1)/n \rfloor} & \text{otherwise} \end{cases}$$

A is totally unimodular because its rows are split into two classes ($i \leq m$ and $i > m$) such that each column has exactly one entry from each class (page 276 of [46]). The above theorems imply that for m and n fixed we can pre-compute the complete fan Σ and for each cone $C \in \Sigma$ pre-compute the enumeration polynomial p_C. This means that for fixed m and n that the counting problem for contingency tables can be solved in polynomial time.

Direct formulae for $2 \times n$ and $3 \times n$ contingency tables were given by Brad Mann [18]. While these formula are exponentially large in n it has been noticed that they have fewer terms than the a dense polynomial for a cone would have (Mann's $3 \times n$ formula has $(2^n)^2 = 4^n$ terms where the $3 \times n$ polynomial could, if dense, have $\binom{3n}{n+2} \asymp O((6.75)^n)$ terms). It should be noticed that the polynomial techniques given above can, without modification, handle contingency tables with censored entries (just leave out columns of A).

3.5 Counting tables with structural constraints.

A table with a structural constraint is a matrix X that has, in addition to the constraints mentioned in earlier sections, constraints of the form $X_{i,j} = v, X_{i,j} \leq v$ or $X_{i,j} \geq v$ for various i, j. It is easy to see (by modifying r_i and c_j) that it would be sufficient to know how to count tables with constraints of the form $X_{i,j} = 0$ efficiently. This is done by eliminating these constraints one by one using the relationship:

$$\sharp(r;c:X_{i,j}=0,\cdots)=\sharp(r;c:\cdots)-\sharp(r_1,\cdots,r_i-1,\cdots r_m;c_1,\cdots,c_j-1,\cdots c_n:\cdots) \tag{3.11}$$

(when $r_i, c_j > 0$). One could also perform directly the same cone decomposition that works for tables without structural constraints- but this would require more storage.

It is clear that the number $m \times n$ table with arbitrary structural constraints can be counted in about the same amount of time as $m \times n$ tables (if extra oracles are produced) or in time no more than 2^{mn} times the time to count one $n \times m$ table using just one counting oracle.

3.6 Assigning a manageable order to tables.

If instead of counting one wishes to explicitly enumerate contingency tables it is useful to define a total order on tables. This order is also very useful in constructing a contingency table sampling scheme from a counting oracle.

Let $T(r_1, \cdots r_m; c_1, \cdots c_n) = T(r;c)$ be the set of m by n non-negative integer matrices X such that $\forall i \quad \sum_{j=1}^{n} X_{i,j} = r_i$ and $\forall j \quad \sum_{i=1}^{m} X_{i,j} = c_j$. We assign a linear order to this set. The order we have been working with is the "lex" order defined $\forall X, Y \in T(r;c)$

$$(X > Y) \leftrightarrow \left(\begin{array}{lrrcl} \exists a,b\; 1 \le a \le m, \;\; 1 \le b \le n & & X_{a,b} & > & Y_{a,b} \\ & \wedge \quad \forall j > b & X_{a,j} & = & Y_{a,j} \\ & \wedge \quad \forall i > a \forall j & X_{i,j} & = & Y_{i,j} \end{array} \right) \tag{3.12}$$

With some care we can step through all the tables in $T(r;c)$ efficiently. We appeal to a routine called "`backfill`". When X is a m by n matrix and a, b are integers `backfill`(X, a, b, r, c) returns the lex least non-negative integral table Y obeying margins r, c such that $Y_{i,j} = X_{i,j}$ for all $i > a$ and if $1 \le a \le m$ then $Y_{a,j} = X_{a,j}$ for all $j > b$. If there are no such tables `backfill` returns $\perp$.

`backfill` works the following simple observation. Let $T(r;c : X_{m,j_1} = v_1, \cdots X_{m,j_k} = v_k)$ be the set of tables $X \in T(r;c)$ such that $X_{m,j_1} = v_1, \cdots X_{m,j_k} = v_k$

Lemma 8 *$T(r;c : X_{m,j_1} = v_1, \cdots X_{m,j_k} = v_k) = \emptyset$ iff $T(r;c : X_{m,j_1} = v_1, \cdots X_{m,j_k} = v_k, X_{m,j_{k+1}} = v_{k+1}) = \emptyset$ where all the $j_i \quad (i = 1 \cdots k+1)$ are distinct and $v_{k+1} = \max(0, \min(c_{j_{k+1}}, r_m - \sum_{i=1}^{k} v_i))$.*

Proof: Clearly there is only one direction to prove (as adding a constraint can not make an infeasible system feasible). So assume $T(r; c : X_{m,j_1} = v_1, \cdots X_{m,j_k} = v_k) \neq \emptyset$ and let $X \in T(r; c; X_{m,j_1} = v_1, \cdots X_{m,j_k} = v_k)$ be a table with maximal $X_{m,j_{k+1}}$. It is clear that $X_{m,j_{k+1}} \leq v_{k+1}$. If $X_{m,j_{k+1}} < v_{k+1}$ then there are a, b such that $X_{a,j_{k+1}} > 0$ and $X_{m,b} > 0$ and b distinct from all j and we could build a new table by increasing $X_{m,j_{k+1}}$ and $X_{a,b}$ by 1 and decreasing $X_{a,j_{k+1}}$ and $X_{m,b}$ by 1. Thus, since $X_{m,j_{k+1}}$ was maximal, have a feasible table with $X_{m,j_{k+1}} = v_{k+1}$. $\square$

So `backfill` can operate by filling in the highest index row first (the most significant entries fall in this row) and trying to put as much mass as possible in the least significant (left most) positions in the row (using lemma 8). If no sums are violated then the table created is the lex least matching the specified partial table, otherwise there are no such tables (and `backfill` should return $\perp$).

One can efficiently step through all the tables in $T(r; c)$ by calling `backfill`$(0, m + 1, 0, r, c)$ to get the lex-first table Y_1. To proceed to the next table one sets (a, b) to the least significant entry (1, 1), sets Y to be the current table with the (a, b)th entry incremented by 1 and calls `backfill`(Y, a, b, r, c) and determines if the returned table is feasible. If the table is infeasible (a, b) is advances to the next significant entry, which is increased by 1, and the call is repeated. The first legal table returned is the next table in our lexicographic order. This method imitates the propagation of carries found in counting, so even though a simple analysis indicates that it could take $nm\sharp(T(r; c))$ fill in attempts to run through all $\sharp(T(r; c))$ tables it is easy to see (for tables where r and c have all largish entries) this method actually lists all the tables using only $O(\sharp(T(r; c)))$ calls to `backfill`.

3.7 Divide and conquer.

Let X_i denote the ith row of a matrix X. The basic method we are using to count m by n tables ($m \geq n$) is as follows: let $k = \lfloor m/2 \rfloor$ and $q_1 = \sum_{i=1}^{k} r_i$ and $q_2 = \sum_{i=k+1}^{m} r_i$ we then notice that

$$\sharp(T(r_1, \cdots r_m; c_1, \cdots c_n)) = \sum_{X \in T(q_1, q_2; c_1, \cdots c_n)} \sharp(T(r_1, \cdots r_k; X_1)) \times \sharp(T(r_{k+1}, \cdots r_m; X_2)).$$

The idea is simple: we can split each table into two independent set of rows if we know how much each column sums to in each set of rows. Thus if we sum over all ways the columns can split their totals between these sets (the set of such splits is equivalent to a set of $2 \times n$ tables) then we can multiply the number of ways to fill in the top and bottom portions- allowing us to count much faster than explicit enumeration.

This formula can be applied recursively and we can stop the recursion at the base cases where m or n is equal to one and the answer is one. This method is much faster than explicitly enumerating all of the tables (in fact it was able to calculate the number of tables consistent with Snee's hair and eye color margins in under 20 minutes on an HP720). This method is not polynomial even for fixed m and n. However it is sufficient to solve counting problems to infer the piecewise polynomial that is the true counting function.

Remark: 3 *Actually exhaustive enumeration would be sufficient to imply a polynomial time counting method for fixed m, n as the tables that need to be solved to infer the piecewise polynomial depend only on m and n and not on the margins of any particular table one wishes to know the count of. This would, of course, be prohibitive in practice so it is desirable to develop the divide and conquer technique.*

3.8 Inferring the piecewise polynomial.

At this point we have an effective scheme for identifying the regions that the counting function is a well behaved on. All the remains is to infer the polynomial for each piece. There are some results in commutative algebra that relate the polynomials to "Hilbert Series" and "Todd Classes", but these structures encode a lot of information and are in themselves often hard to compute. The strategy taken here is to assume access to a counting oracle (in this case simulated by the divide and conquer counter) and then recover the desired polynomial by interpolating the values known in a region.

3.8.1 Lagrange interpolation

The first interpolation method is not too sophisticated. We find a point b in the cone we are working with such that $b + sE^i$ (s is the degree of the polynomial) is in the cone for all i. We know such a point must exist because the cone we are using is pointed and full dimensional (and therefore contains arbitrarily big balls as we move away from the origin). We then count (by divide and conquer) all the tables corresponding to any set of margins found in the set of points $x \in Z^d$ ($d = r+c-1, w = rc, s = (r-1)(c-1)$) such that $x \geq b$ and $\vec{1} \cdot (x-b) \leq s$. The counts at these points are used to determine the polynomial by using "Lagrange interpolation." For $b, x \in Z^d$ and $s \in Z$ such that $x \geq b$ and $\vec{1} \cdot (x-b) \leq s$ the Lagrange polynomial $l_{b,x,s}$ is the unique degree s polynomial in variables $v_1 \cdots v_d$ such that $l_{b,x,s}(x) = 1$ and $l_{b,x,s}(y) = 0$ for any integral $y \neq x$ such that $y \in Z^d,\ y \geq b,\ \vec{1} \cdot (y-b) \leq s$. $l_{b,x,s}$ has a very simple definition

$$l_{b,x,s}(v) = \begin{cases} \prod_{i=1}^{s} \left(1 + \frac{\vec{1}}{i} \cdot b - \frac{\vec{1}}{i} \cdot v\right) & x = b \\ \frac{(v_i - b_i)}{(x_i - b_i)} l_{b+E^i,x,s-1} & x_i > b_i,\ x_j = b_j\ \forall j < i \end{cases} \tag{3.13}$$

It is then easy to see that

$$p_C = \sum_{x \in Z^d,\ x \geq b,\ \vec{1} \cdot (x-b) \leq s} F(x) l_{b,x,s} \tag{3.14}$$

where $F(x)$ is the number of integer points in P_x.

This technique is important because we do not have to explicitly store a linear transformation to fit the polynomial coefficients to the known evaluations (in the 4×4 case there are $\binom{16}{7} = 11440$ coefficients to fit which would require a 11440×11440 matrix if the linear transformation were stored explicitly. If each entry took just 16 bytes to store this would still represent almost 2 gigabytes of storage). Also if extra storage is available this method can change k polynomial values into k coefficients in time $O(k^2)$. This can be accomplished by building a table of all k Lagrange polynomials in a sort of Gray-code order (where most polynomials in the table can determined by dividing a known one by a binomial and multiplying by another binomial in $O(k)$ time). Even though this requires as much storage as the direct linear algebra approach it is much faster. The primary weakness of this method is that b may be very large and require the solution of very difficult counting problems (though the set of problems requiring solution is very structured, allowing some savings).

A nice generalization, which would help with "large b" problem mentioned above, would be to find how to easily compute Lagrange polynomials for an arbitrary set S such that S determines p_C and S has minimal cardinality. The polynomials would then be indexed by x, S and d where $l_{x,S,s}$ would be the unique degree s polynomial such that $l_{x,S,s}(x) = 1$ and $l_{x,S,s}$ is zero on $S \setminus x$. By linear algebra we know that such a basis for the space of all degree s (or less) polynomials exists- but it is not clear that it can be competed with limited space and time (i.e. with less space than the number of terms squared).

One conjectured method to efficiently perform this calculation would be to hope that all minimal S have structure similar to the simplex we used in the last section. The recursive formula given for the Lagrange polynomials depended on finding a set of degree 1 polynomials ($(v_i - b_i),\ i = 1 \cdots d$ and $(\vec{1} \cdot v - s - \vec{1} \cdot b)$) that each zeroed out at least $\binom{n+d-1}{d}$ points in S but have no common zero in S. Even for the simple case of $d = 2$, $s = 2$ this is not always possible. Take

$$S = \{(0,0),(1,0),(0,1),(1,1),(3,2),(2,3)\}. \tag{3.15}$$

We expand elements of S to vectors in $\mathbf{R}^6$ by taking $(x,y) \rightarrow (1,x,y,xy,x^2,y^2)$ and write down the 6×6 matrix gotten by expanding all of S:

$$M = \left\{ \begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 2 & 6 & 9 & 4 \\ 1 & 2 & 3 & 6 & 4 & 9 \end{array} \right\}. \tag{3.16}$$

We see that $\det(M) = 32 \neq 0$ so S is a minimal basis of the kind we wanted. But no $\binom{d+s-1}{s} = \binom{3}{2} = 3$ points in S are collinear- so there is no way to build the Lagrange polynomials up from degree 1 polynomials as before.

3.8.2 An improvement

A simple improvement is to notice that if $M : \mathbf{R}^d \rightarrow \mathbf{R}^d$ is a full rank linear transformation and $p : \mathbf{R}^d \rightarrow \mathbf{R}$ is a rational polynomial of degree no more than s then $\bar{p}$ defined such that $\bar{p}(x) = p(M^{-1}x)$ is also a rational polynomial of degree no more than s. For each cone $C \in \mathbf{\Sigma}$ we let M_C be the $d \times d$ matrix whose columns are the first d integral vectors in C such that M is full rank (vectors ordered in graded lex order). Then we see if $S = \left\{ x \in Z^d \;\middle|\; x \geq 0 \,, \vec{1} \cdot x \leq s \right\}$ then $M_C(S)$ is a set of points that determine the interpolation polynomial for the cone C and $\bar{p}$ can be inferred by the above naive method. The advantage is that this set of problems may be much smaller (and therefore easier) that the set of problems that the original interpolation method would require. Finding a basis for

Z^d in the given cone is of course a hard problem, but it is a problem in Z^d not in the much larger space of coefficients and evaluations. In practice d is so much smaller than the number of coefficients needed that the time needed to search for the small basis by brute force has been negligible compared to any of the other steps in the algorithm. It would be interesting to use more subtle techniques such as computing a Hilbert basis for the cone and examining it for an acceptable short basis.

3.9 Sampling from counting.

For many statistical tests it is important to be able to generate a table uniformly at random that obeys given row and column sums. It is well know [38] that sampling and counting are intimately related. For this problem the relationship can be exposed by combining the fixed dimensional counter and the "lex" ordering introduced here for brute force enumeration. The method is as follows.

Suppose we want to generate a table, X, uniformly at random obeying rows sums $r_1, r_2, \cdots r_m$ and column sums $c_1, c_2, \cdots c_n$ where m, n are fixed. We will assume that we have pre-computed a representation of the piecewise polynomial counting function, so we can count tables at will. First we compute N, the total number of tables obeying r, c. We generate a integer k in the set $\{1, 2, \cdots N\}$ uniformly at random. We are going to return the lex kth table as our uniform random sample. First we look at the most significant entry of the table: $X_{m,n}$. If for an integer b such that more than k tables obeyed r, c and the structural constraint $X_{m,n} \leq b$ then we would know that more than the lex kth table must have $X_{m,n} \leq b$. Similarly we know that if fewer that k tables obeyed r, c and $X_{m,n} \leq b$ then the lex kth table must have $X_{m,n} > b$. Thus we can, using binary search on b, find in time $\log_2(\min(r_m, c_n))$ find the true value of $X_{m,n}$ for the lex k'th table. We then add $X_{m,n} = b$ as a structural constraint and use the same technique to find the correct value for the next most significant entry. This process is repeated until the entire table is constrained and then X is the lex kth table as desired. This method can generate a true uniform sample by solving no more than $(m-1)(n-1)\log_2(\min(r_m, c_n))$ structural counting problems. We can either assume that we had the $(m-1)(n-1)$ different types of counting polynomials pre-computed, or by using the formula 3.11 judiciously and converting each structural problem into no more that 2^{n-1} counting problems (notice that we apply structural constraints to only one row at a time until the row is completely constrained).

3.10 4×4 results.

We have solved the $m = n = 4$ case. This case splits into (after some symmetries are removed) 3694 cones such that $\sharp_{4,4}$ restricts to a degree 9 polynomial in 7 variables in each cone. This means each polynomial is determined by 11440 coefficients so can be interpolated from 11440 sufficiently general evaluations. Each of the 3694 tasks seems to represent about 3 hours of pmax cpu time, so the total job represents about 1.5 pmax cpu years. This task was completed in just over 6 weeks using Peter Stout's **WAX** system [50] which effectively simulates a coarse-grain parallel supercomputer (employing the numerous idle workstations at CMU).

This decomposition is able to count the number of tables compatible with the hair/eye color table 1.1 in 30 seconds.[3]

3.10.1 Snee's table

This table 1.1 has been our main test case. It can be counted in under 20 minutes using the divide and conquer technique. The divide and conquer counter has been augmented to generate a batch of uniform random samples while counting. This is accomplished by generating a large set uniform random integers from the range 1 to the total number of tables compatible with the margins. The divide and conquer procedure is then run with the batch being divided into the appropriate divisions until the tables are completely filled out. The indices used here are not in the lex order used in other sections but in an arbitrary, but constant, order determined by the coding of the divide and conquer algorithm. This method when dealing with batches of 10000 tables can generate about 20 tables a second. The pre-computed 4×4 solution can count Snee's tables in about 30 seconds. And this method can generate a uniform random sample in a couple of minutes. The generator can be converted into a batch process (so many tables share the first few hard counting problems) to generate uniform random tables in a batch. The savings would come from many tables sharing the harder early counting problems.

The results to date on this table, which clearly has been over studied, are that best estimate of the chi-square significance of the table is 15.5% with a 3 standard deviation confidence interval of $\pm.15$.

This agrees well with results from two Markov chains (described later) one

[3]All CPU times given in this section are HP720 CPU seconds unless otherwise noted.

based on a natural “King’s move” walk and one based on the newer “ball move” walk. These chains returned estimates of 15.6 ($\pm$3 std dev interval 15.2 $\cdots$ 15.9) and 15.6 ($\pm$3 std dev interval 14.1 $\cdots$ 17.1) respectively. The second confidence interval is much wider as the chain ran for a much shorter time.

The confidence intervals would not have been possible without the proof of convergence. One could design and run one of these chains without the convergence proof. However for a given length run one would not be able to estimate how many truly random samples the set of samples drawn for the chain was equivalent to. So there would be no way to draw confidence intervals around the estimate. Thus one could be in the unhappy circumstance of having two chains in wild disagreement and be able to draw any conclusions.

3.11 5×4.

The 5×4 problem is large enough to be considered impractical. However for any one table one can identify a cone its right hand side lies in and infer the polynomial for this cone alone. In fact we have found significant savings in not inferring the true polynomial by computing $\sum f(x)l_x$ (where $f(x)$ is the true value and l_x is the Lagrange polynomial such that $l_x(x) = 1$ and $l_x(y) = 0$ for all other points in the interpolations set) but computing the count for the desired table, z, by computing $f(z) = \sum f(x)l_x(z)$. $l_x(z)$ is an integer (not a polynomial) and can be computed much quicker than l_x can and the process requires almost no storage.

This method was used to determine the number of tables compatible with the margins $[\ 182,\ 778,\ 3635,\ 9558,\ 11110\ ]$ and $[\ 3046,\ 5173,\ 6116,\ 10928\ ]$ is 2319643659612889757482961153193875 3 in 8 days on an HP720 workstation. While this may not seem quick it should be pointed out that computing the sum mentioned in the previously paragraph is “embarrassingly parallel” (most of the time would be spent in computing the terms of the sum without any need for communication) so this calculation could be done quickly on a parallel computer.

3.12 $m \times 2$ and $m \times 3$.

Brad Mann has developed effective formulas for both the $m \times 2$ and $m \times 3$ case. Mann’s results are inclusion/exclusion formulas over all partitions of m. These formulas have a number of terms comparable to the polynomials developed here for exact computation but have the distinct advantage of involving no pre-computation and use very little space.

3.13 Asymptotic performance of estimates.

Asymptotic behavior of the contingency table counting function can be easily read off our chamber decomposition. This has made possible an interesting comparison of several estimation schemes in the literature.

The first estimate is from [44] and is intended for sparse tables. Let $T = \sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$ and assume the row sums are all less than $\log(\min(m,n))^{1/4-\epsilon}$ then

$$\sharp(r;c) \sim E1(r;c) = \frac{T!}{\prod r_i! \prod c_j!} e^{\frac{2}{T^2}\sum_{i,j}\binom{r_i}{2}\binom{c_j}{2}} \tag{3.17}$$

as $m, n \to \infty$. When the sparsity condition is not met this estimate tends to be off by an exponential factor (and is not used for the type of contingency tables arising in statistics).

The next estimate is from [30] and is intended for non-sparse tables (with a large number of large rows). The idea is to take the number of tables obeying the row- constraints and to multiply it by an estimate of the odds of satisfying the column constraints.

$$\begin{aligned}
\sigma^2 &= \frac{(n-1)\sum r_i(r_i+n)}{(n+1)n^2} \\
Q &= \frac{n-1}{\sigma^2 n} \sum c_j^2 - T^2/n \\
\sharp(r;c) &\sim E2(r;c) = \sqrt{n} e^{-Q/2} \left(\frac{n-1}{2\pi\sigma^2 n}\right)^{(n-1)/2} \prod_{i=1}^{m} \binom{r_i+n-1}{n-1}
\end{aligned} \tag{3.18}$$

And the final estimate is from Diaconis and Efron [16]. This estimate is based on a volume times density of lattice argument- with the novel innovation that the volume is purposefully overestimated to compensate for some corner effects missed in this type of analysis. Because of this the Diaconis/Efron technique estimate is much harder to analyze and not much is know about it.

$$w = \frac{1}{1 + mn/2T}$$

$$
\begin{aligned}
\bar{r_i} &= \frac{1-w}{m} + \frac{wr_i}{T} \\
\bar{c_j} &= \frac{1-2}{n} + \frac{wc_j}{T} \\
k &= \frac{n+1}{n\sum_{i=1}^{m} \bar{r_i}^2} - \frac{1}{n} \\
\sharp(r;c) &\sim E3(r;c) = \left(\frac{2T+mn}{2}\right)^{(m-1)(n-1)} \times \\
& \left(\prod_{i=1}^{m} \bar{r_i}\right)^{n-1} \left(\prod_{j=1}^{n} \bar{c_j}\right)^{k-1} \frac{\Gamma(nk)}{\Gamma(n)^m \Gamma(k)^n} \qquad (3.19)
\end{aligned}
$$

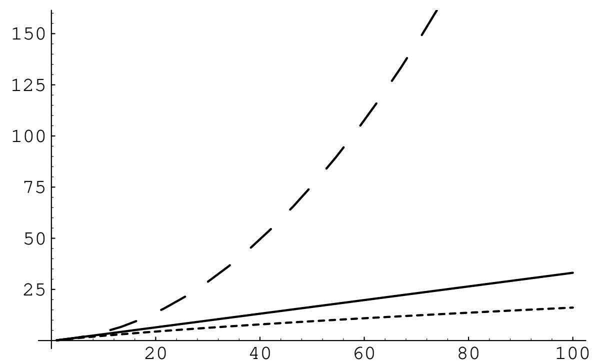
The first asymptotic trajectory we examine is counting the number of tables with row sums $[t, kt, kt]$ and column sums $[t, kt, kt]$ where k is a positive integer. By examining the complete 3 by 3 solution it is clear that the number of tables obeying these margins is

$$
\left(\frac{k}{3} - \frac{5}{24}\right) t^4 + \left(\frac{3k}{2} - \frac{3}{4}\right) t^3 + \left(\frac{13k}{6} - \frac{7}{24}\right) t^2 + \left(k + \frac{5}{4}\right) t + 1
$$

which is asymptoticly $\left(\frac{k}{3} - \frac{5}{24}\right) t^4$ as $t \to \infty$ and k is held constant. A little work shows the asymptotic behavior of the estimates on this trajectory is as follows.

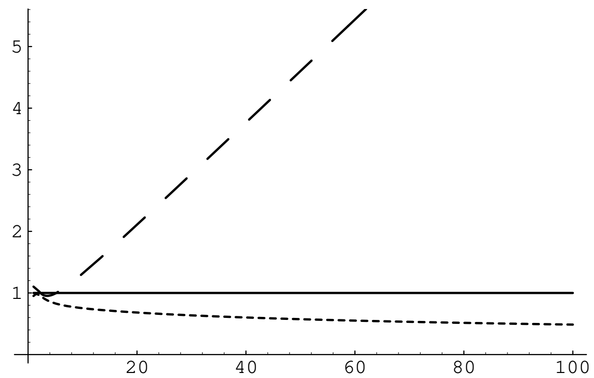
$$
\begin{aligned}
\log(E1) &\asymp \frac{(1+2k^2)^2}{2(1+2k)^2} t^2 && (3.20) \\
E2 &\asymp \frac{3\sqrt{3}k^2}{4\pi(2k^2+1)e^{4(k-1)^2/(2k^2+1)}} k^2 t^4 && (3.21) \\
E3 &\asymp \frac{2^{\frac{8\,k\,(2+k)}{1+2\,k^2}}\, k^{10/3} \, \frac{k^2}{(2+4\,k)^3}^{-\frac{4}{3} + \frac{4\,(1+2\,k)^2}{3\left(1+2\,k^2\right)}} \Gamma\left(\frac{3+16\,k+14\,k^2}{1+2\,k^2}\right)}{8\,(1+2\,k)^2\, \Gamma\left(\frac{3+16\,k+14\,k^2}{3+6\,k^2}\right)^3} k^{2/3} t^4 && (3.22)
\end{aligned}
$$

Estimate $E1$ is not really worth discussing (it was never intended for this case). Estimate $E2$ tends to overestimate the count by a factor of about k and $E3$ tends to underestimate by a factor of $k^{1/3}$. Figure 3.1 plots the behavior of these asymptotic expressions as a function of k. The graph is the value of equations 3.21 and 3.22 and the true asymptotic behavior (as $t \to \infty$), all divided by t^4, as functions of k. Figure 3.2 shows the same trend with the true count rescaled to be 1.



solid=true, dashed=E2, dotted=E3

Figure 3.1: Asymptotic behavior of counting estimates.



solid=true, dashed=E2, dotted=E3

Figure 3.2: Rescaled asymptotic behavior of counting estimates.

Another interesting trajectory is $[t, t^2, t^3]; [t, t^2, t^3]$. Along this trajectory the number of contingency tables is asymptoticly $t^5/3$, E2 is asymptoticly $3\sqrt{3}/(4\pi e^4)t^6$ and E3 is asymptoticly $t^6/4$.

3.14 Large tables with small margins.

The divide and conquer counter (when combined with Mann's formulas as base cases) is able to exactly tables from the literature.

For example, Mehta and Patel [42], in the course of performing significance tests, estimate counts for the following tables (using Gail/Mantel's method [30]).

Problem 1:

										Total
	1	1	1	0	0	0	1	3	3	10
	4	4	4	4	4	4	4	1	1	30
Total	5	5	5	4	4	4	5	4	4	40

Problem 2:

						Total
	2	0	1	2	6	11
	1	3	1	1	1	7
	1	0	3	1	0	5
	1	2	1	2	0	6
Total	5	5	6	6	7	29

Problem 3:

							Total
	2	0	1	2	6	5	16
	1	3	1	1	1	2	9
	1	0	3	1	0	0	5
	1	2	1	2	0	0	6
Total	5	5	6	6	7	7	36

Problem 4:

										Total
	1	1	1	0	0	0	1	2	4	10
	4	4	4	5	5	5	6	5	0	38
	1	1	1	0	0	0	1	2	4	10
Total	6	6	6	5	5	5	8	9	8	58

Problem 5:

							Total
	1	2	2	1	1	0	7
	2	0	0	2	3	0	7
	0	1	1	1	2	7	12
	1	1	2	0	0	0	4
	0	1	1	1	1	0	4
Total	4	5	6	5	7	7	34

Problem 6:

								Total
	1	2	2	1	1	0	1	8
	2	0	0	2	3	0	0	7
	0	1	1	1	2	7	3	15
	1	1	2	0	0	0	1	5
	0	1	1	1	1	0	0	4
Total	4	5	6	5	7	7	5	39

Our results working with these tables are as follows:

Problem	True count	reported est	Gail/Mantel ests	Diaconis/Efron ests
1	35353	40500	34534,73397	37992,40169
2	3187528	$1.1 * 10^6$	$3.01 * 10^6$, $3.84 * 10^6$	$3.30 * 10^6$,$3.32 * 10^6$
3	97080796	$68 * 10^6$	$125 * 10^6$,$68 * 10^6$	$110 * 10^6$, $112 * 10^6$
4	1326849651	$624 * 10^6$	$1963 * 10^6$, $519 * 10^6$	$1615 * 10^6$, $1757 * 10^6$
5	2159651513	$1.6 * 10^9$	$2.5 * 10^9$, $1.7 * 10^9$	$2.6 * 10^9$, $2.6 * 10^9$
6	108712356901	$64 * 10^9$	$132 * 10^9$, $64 * 10^9$	$144 * 10^9$, $149 * 10^9$

Table 3.1: Estimates from the literature.

Both the Gail/Mantel and Diaconis/Efron estimate are asymmetric (can give different estimated counts for a table and its transpose), so we have reported both estimates (with the better one first). Somebody actually using these estimates would not be able to, in all cases, identify the best of the two estimates. The differences between Mehta/Patel's reported values of the Gail/Mantel estimate and values given in lines 1,2 and 4 of this table are disturbing (the others are insignificant).

3.15 Magic squares.

As an exercise we were able to rederive all the known magic square Ehrhart polynomials. $p_n(x)$ is defined as the number of $n \times n$ non-negative integer matrices such that every row and every column adds up to x. $p_1 \cdots p_6$ are:

$$
\begin{aligned}
p_1(x) &= 1 && (3.23)\\
p_2(x) &= 1 + x && (3.24)\\
p_3(x) &= 1 + \frac{9\,x}{4} + \frac{15\,x^2}{8} + \frac{3\,x^3}{4} + \frac{x^4}{8} && (3.25)\\
p_4(x) &= 1 + \frac{65\,x}{18} + \frac{379\,x^2}{63} + \frac{35117\,x^3}{5670} + \frac{43\,x^4}{10} + \frac{1109\,x^5}{540} + \frac{2\,x^6}{3} + \frac{19\,x^7}{135} + \\
&\quad \frac{11\,x^8}{630} + \frac{11\,x^9}{11340} && (3.26)\\
p_5(x) &= 1 + \frac{725\,x}{144} + \frac{6229735\,x^2}{494208} + \frac{3028287247\,x^3}{145297152} + \frac{438177965089\,x^4}{17435658240} + \frac{664118435\,x^5}{28740096} + \\
&\quad \frac{3812839477\,x^6}{229920768} + \frac{196563587\,x^7}{20901888} + \frac{3541860299\,x^8}{836075520} + \frac{55426325\,x^9}{36578304} + \frac{125188639\,x^{10}}{292626432} + \\
&\quad \frac{984101\,x^{11}}{10450944} + \frac{72750523\,x^{12}}{4598415360} + \frac{112655\,x^{13}}{57480192} + \frac{1008757\,x^{14}}{5977939968} + \frac{188723\,x^{15}}{20922789888} + \\
&\quad \frac{188723\,x^{16}}{836911595520} && (3.27)\\
p_6(x) &= 1 + \frac{3899\,x}{600} + \frac{46584105377\,x^2}{2141691552} + \frac{12246206617138789\,x^3}{247365374256000} + \frac{382955230861099213\,x^4}{4517106834240000} + \\
&\quad \frac{15549846579377723056 7\,x^5}{1355132050272000000} + \frac{1422688636839855 1\,x^6}{112634352230400} + \frac{243245111626317349\,x^7}{2111894104320000} + \\
&\quad \frac{232132948167689\,x^8}{2634721689600} + \frac{253578194011961479\,x^9}{4446092851200000} + \frac{736591080322991\,x^{10}}{23433524674560} + \\
&\quad \frac{162650481872908 69\,x^{11}}{1098446469120000} + \frac{2000221303490489\,x^{12}}{334764638208000} + \frac{570713692223620411\,x^{13}}{2761808265216000000} + \\
&\quad \frac{8346012436199\,x^{14}}{13638559334400} + \frac{1424745952102609\,x^{15}}{9206027550720000} + \frac{77984295979769\,x^{16}}{2343352467456000} + \\
&\quad \frac{1062348478211833\,x^{17}}{175751435059200000} + \frac{18674864899\,x^{18}}{20324995891200} + \frac{2462417656967\,x^{19}}{21341245685760000} +
\end{aligned}
$$

$$\frac{141248912237\,x^{20}}{12014330904576000} + \frac{853529939221\,x^{21}}{901074817843200000} + \frac{4394656999\,x^{22}}{75690284698828800} +$$
$$\frac{158824242127\,x^{23}}{6244448487653376000} + \frac{9700106723\,x^{24}}{136783157348597760000} +$$
$$\frac{9700106723\,x^{25}}{1025873680114483200000} \tag{3.28}$$

3.16 Availability of software.

The url `http://mixing.sp.cs.cmu.edu/` provides an online demonstration of this work over the World Wide Web.

3.17 Problem hardness.

The problem of determining the number of contingency tables consistent with row and column totals is quite difficult. Martin Dyer and Ravi Kannan[24] established the following.

Theorem 9 *Determining the number of* $2 \times n$ *non-negative integer matrices* X *such that* $\sum_{j=1}^{n} X_{i,j} = r_i$ *(*$i = 1, 2$*) and* $X_{1,j} + X_{2,j} = c_j$ *(*$j = 1 \cdots n$*) is* $\sharp P$*-complete.*

Proof: In [22] it is shown that given the positive integers $a_1, a_2, \cdots, a_{n-1}, b$ it is $\sharp P$-hard to compute the $n-1$ dimensional volume of the polytope defined by

$$\begin{array}{c} \sum_{j=1}^{n-1} a_j y_j \leq b \\ 0 \leq y_j \leq 1 \;\; (j = 1, 2, \cdots n-1) \end{array} .$$

Thus if $a_n = b$ it is $\sharp P$-hard to compute the $n-1$ dimensional volume

$$\begin{array}{c} \sum_{j=1}^{n} a_j y_j = b \\ 0 \leq y_j \leq 1 \;\; (j = 1, 2, \cdots n) \end{array} .$$

Substituting $x_{1,j} = a_j y_j$, $x_{2,j} = a_j(1 - y_j)$ we encode this as a small set of contingency table problems. We identify $c_j = a_j \quad (j = 1, \cdots n)$, $,r_1 = b$ and $r_2 = \sum_{j=1}^{n-1} a_j$. We then use the fact [34] that for integer k the function $f(k) = \sharp(kr, kc)$ is a degree $n-1$ polynomial in k and the coefficient of k^{n-1} is the desired volume. Thus if we could solve the contingency table counting problem for $k = 0, 1, 2 \cdots n-1$ we could infer all of the coefficients of f, including the highest one. It is easy to see that the interpolation can be performed using rational numbers that require no more than $O(n(\log(r_1 + r_2) + n))$ bits to represent. $\Box$

3.18 Near uniform generation.

Consider $m > 1$ row by $n > 1$ column contingency tables obeying row and column sums ($r_i, c_j > 0$ and integer) such that $r_{i+1} \geq r_i \quad (i = 1 \cdots m-1)$ and $c_{j+1} \geq c_j \quad (j = 1 \cdots n-1)$.

Let $T = \sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$,

$$K = \left\{ X \in \mathbf{R}^{(m-1)(n-1)} \;\middle|\; \begin{array}{lrcll} (1) & \sum_{i=1}^{m-1} X_{i,j} & \leq & c_j & j = 1 \cdots n-1 \\ (2) & \sum_{j=1}^{n-1} X_{i,j} & \leq & r_i & i = 1 \cdots m-1 \\ (3) & \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} X_{i,j} & \geq & T - r_m - c_n & \\ (4) & X_{i,j} & \geq & 0 & j = 1 \cdots n-1, \\ & & & & i = 1 \cdots m-1 \end{array} \right\}.$$

There is an obvious $1-1$ correspondence between integral points in K and contingency tables obeying the given row and column sums (fill in last/row column using the linear dependencies). For $u \in Z^{(m-1)(n-1)}$ let

$$C(u) = \left\{ x \in \mathbf{R}^{(m-1)(n-1)} \;\middle|\; \begin{array}{ll} u_{i,j} - \frac{1}{2} \leq x_{i,j} < u_{i,j} + \frac{1}{2} & j = 1 \cdots n-1, \\ & i = 1 \cdots m-1 \end{array} \right\}$$

and

$$J = \bigcup_{u \in Z^{(m-1)(n-1)} \cap K} C(u)$$

Now we wish to find a convex body K' such that $J \subseteq K'$ and Vol(K') is not too much bigger than Vol(K). If we had such a body then to generate contingency table (from near uniform distribution) we would use standard techniques to generate a near uniform point from K' and rejection sample until we find a point in J and round to an integral vector.

Define

$$Y_{i,j} = \frac{r_i c_j}{T} + \frac{\min(r_i, c_j)}{2\max(n^2 - n, m^2 - m)}.$$

Lemma 9 *The coordinate aligned cube of (l_∞) diameter $\frac{\min(r_1,c_1)}{2\max(n^2-n,m^2-m)}$ centered at Y is contained in K.*

Proof: It suffices to check inequalities (1) and (2) with $X = Y^+$ and inequalities (3) and (4) with $X = Y^-$ where $Y^+_{i,j} = Y_{i,j} + \frac{\min(r_i,c_j)}{2\max(n^2-n,m^2-m)}$ and $Y^-_{i,j} = Y_{i,j} - \frac{\min(r_i,c_j)}{2\max(n^2-n,m^2-m)}(= \frac{r_i c_j}{T})$.

1.
$$\begin{aligned}\sum_{i=1}^{m-1}\left(\frac{r_i c_j}{T} + \frac{\min(r_i,c_j)}{\max(n^2-n,m^2-m)}\right) &\leq \frac{c_j(T-r_m)}{T} + \frac{c_j(m-1)}{m^2-m}\\ &\leq \frac{c_j(m-1)}{m} + \frac{c_j(m-1)}{m^2-m}\\ &= c_j.\end{aligned}$$

2.
$$\begin{aligned}\sum_{j=1}^{n-1}\left(\frac{r_i c_j}{T} + \frac{\min(r_i,c_j)}{\max(n^2-n,m^2-m)}\right) &\leq \frac{r_i(T-c_n)}{T} + \frac{r_i(n-1)}{n^2-n}\\ &\leq \frac{r_i(n-1)}{n} + \frac{r_i(n-1)}{n^2-n}\\ &= r_i.\end{aligned}$$

3.
$$\begin{aligned}\sum_{i=1}^{m-1}\sum_{j-1}^{n-1}\frac{r_i c_j}{T} &= \frac{(T-r_m)(T-c_n)}{T}\\ &= T - r_m - c_n + \frac{r_m c_n}{T}\\ &> T - r_m - c_n.\end{aligned}$$

4.
$$\frac{r_i c_j}{T} > 0$$

□

Corollary 2 *If* $\min(r_1,c_1) \geq 2\lambda(n-1)(m-1)\max(n^2-n,m^2-m)$ *then the dilation K' of K about Y by* $(1 + \frac{1}{\lambda(n-1)(m-1)})$ *contains J and* $\frac{\mathrm{Vol}(K')}{\mathrm{Vol}(K)} \leq e^{1/\lambda}$.

Proof: If $\min(r_1, c_1) \geq 2\lambda nm \max(n^2 - n, m^2 - m)$ then by the last lemma we know that a coordinate aligned cube of size $\lambda(n-1)(m-1)$ centered at Y is contained in K. Consider any $c \in \mathbf{R}^{(m-1)(n-1)}$ and b real such that $c \cdot (x - Y) \leq b \ \forall x \in K$. Take an arbitrary z such that $c \cdot (z - Y) \leq b$ and notice that $C(z)$ is contained in the half space $c \cdot (x - Y) \leq (1 + 1/(\lambda(n-1)(m-1)))$. From this it is easy to see that K' contains J. The volume of K' is precisely $(1 + 1/(\lambda(n-1)(m-1)))^{(n-1)(m-1)} \operatorname{Vol}(K)$ which is no more than $e^{1/\lambda} \operatorname{Vol}(K)$. $\square$

Let $\sharp(K)$ denote the number of points in K with integer coordinates.

Corollary 3 *If* $\min(r_1, c_1) \geq 2\lambda nm \max(n^2 - n, m^2 - m)$ *then*

$$e^{-1/\lambda} \operatorname{Vol}(K) \leq \sharp(K) \leq e^{1/\lambda} \operatorname{Vol}(K).$$

Proof: The right inequality follows immediately from the last corollary. Consider any $c \in \mathbf{R}^{(m-1)(n-1)}$ and b real such that $c \cdot (x - Y) \leq b \ \forall x \in K$. Take an arbitrary z such that $c \cdot (z - Y) \leq b$ notice that $C(z)$ is in the complement of the half space $c \cdot (z - Y) \leq (1 - 1/(\lambda(n-1)(m-1)))$. $\square$

3.19 Counting from generation.

The corollaries of the previous section extend into an approximate counting scheme, if we had a method of approximating the ratio of the number of tables with row/column totals (r, c) $(r_i \geq nm \max(n^2 - n, m^2 - m), \ r_j \geq nm \max(n^2 - n, m^2 - m) \ i = 1 \cdots m, \ j = 1 \cdots n)$ to the number of tables with row/column totals (r', c')

$$r'_i = \begin{cases} r_i + \Delta & i = a \\ r_i & \text{otherwise} \end{cases}$$

$$c'_j = \begin{cases} c_j + \Delta & j = b \\ c_j & \text{otherwise} \end{cases}$$

for arbitrary a, b and some non-negligible Δ.

Suppose we wish to count to within a relative error of $1 + \epsilon$ $(0 < \epsilon < 1)$, let $N = \left\lceil \frac{20nm \max(n^2 - n, m^2 - m)}{\epsilon} \right\rceil$ by the above arguments the number of integer tables is between $e^{-\epsilon/10} \operatorname{Vol}$ and $e^{\epsilon/10} \operatorname{Vol}$. This means we certainly have the volume of this body approximates the number of lattice points with a relative error of no more than $1 + \epsilon/3$. volume of a table with all row sums and column sums $\geq N$ approximates the number of lattice points to within a relative error of $1 + \frac{\epsilon}{3}$. All that remains would be to use standard technique (as in Lovasz/Simonovits or Kannan/Lovasz/Simonovits) to calculate the volume of this body to sufficient accuracy and approximate the ratios of lattice points as outline above.

If we could boost all r_i and c_j up to at least N in a polynomial number of stages and compute the ratios of all these stages so that the product of these ratios has a relative error of no more than $1 + \frac{\epsilon}{3}$ then we would have a good estimate of the number of tables obeying the original row and column sums. We will show that we can take $\Delta = \left\lfloor \min(\frac{r_i}{n}, \frac{c_j}{m} \right\rfloor$ when trying to boost a row and guarantee that the ratio of contingency tables obeying the original row/column totals to the boosted row/column totals is in the range $[\frac{1}{(m-1)(n-1)+1}, 1]$. This is enough to allow us to approximately count using no more than $O\left(nm \ln(\frac{\max(N, r_m)}{r_1})\right)$ stages (for m rows double each row in about n stages, raise row to $\max(N, r_m)$ in about $\ln(\frac{\max(N, r_m)}{r_1})$ doublings) each being computed to a relative error no finer than $\Omega\left(\frac{\epsilon}{((m-1)(n-1)+1)nm \ln(\frac{\max(N, r_m)}{r_1})}\right)$ yielding an approximation scheme that runs in time polynomial in $\frac{1}{\epsilon}, n, m$.

boosting a table To compute the ratio of tables obeying (r, c) to (r', c') (as defined before) we identify tables X obeying (r, c) with X' (equals X with $X_{a,b}$ increased by Δ) this defines 1-1 map from tables obeying (r, c) to a subset of the tables obeying (r', c'). So we generate a table X' from the uniform distribution that obeys the (r', c') and check if we are in the subset corresponding to the (r, c) tables (ie. is $X'_{a,b} \geq \Delta$?). All that remains is to show that this subset of all (r', c') tables is not too small.

Theorem 10 *If X is uniform random variable corresponding to a table obeying the row/column sums (r, c) then $X_{a,b} \geq \left\lfloor \min(\frac{r_a}{n}, \frac{c_b}{m}) \right\rfloor$ at least $\frac{1}{(m-1)(n-1)+1}$ of the time.*

Proof: Let $\Delta = \left\lfloor \min(\frac{r_a}{n}, \frac{c_b}{m}) \right\rfloor$. We will define a map f that maps all tables obeying (r, c) into the set of tables X obeying (r, c) such that $X_{a,b} \geq \Delta$. If $X_{a,b} \geq \Delta$ let $f(X) = X$ otherwise by the pigeonhole principle we see that there exists x, y such that $X_{a,y} \geq \left\lfloor \frac{r_a}{n} \right\rfloor$ and $X_{x,b} \geq \left\lfloor \frac{c_b}{m} \right\rfloor$. Let i, j be the smallest indices obeying these inequalities ($i \neq a, j \neq b$). And define

$$f(X)_{u,v} = \begin{cases} X_{u,v} + \Delta & \text{if } u = a \text{ and } v = b \\ X_{u,v} + \Delta & \text{if } u = x \text{ and } v = y \\ X_{u,v} - \Delta & \text{if } u = x \text{ and } v = b \\ X_{u,v} - \Delta & \text{if } u = a \text{ and } v = y \\ X_{u,v} & \text{otherwise} \end{cases}$$

It is easy to see that f maps at most $(m-1)(n-1)+1$ tables to any point in its range. $\square$

a trick to speed it up We can, in some cases, speed up the reduction by avoiding the volume computation. Assume $n = m$ and let K_l denote the polytope corresponding to $r_i = l, c_j = l$ for all i, j. Notice the polytope K_l is just lK_1. Also, K_1 has only integral vertices because the matrix formed by the left hand side of the constraints given in the definition of K is totally unimodular. Thus the theory of lattice point enumerators ([34] pp. 135-140) can be brought in and we see that a polynomial p_l of degree exactly $(m-1)(n-1)$ passes through the integers $\sharp(K_l) \quad l = 1 \cdots \infty$. So if we knew p_l (as a non-uniform piece of information in the circuit theory sense or from an oracle) we would not have to boost the table until all rows and columns were above N but only until they were equal (in fact we could even decrease them all be be min(r_1, c_1) allowing the more natural reduction direction: towards *smaller* problems). Now by a brute force technique (similar to the one developed by the authors of STATEXACT [42]) we have explicitly computed p_l for 2×2, 3×3, 4×4, 5×5 and 6×6 tables (which can be used to help enumerate any tables with max$(n, m) \leq 6$.

3.20 Difficulty in generating Fisher Yates.

Generating Fisher-Yates distributed tables in time polynomial in $log(T)$ (instead of polynomial in T) seems to be more difficult.

The main problem is that the density function varies quite rapidly. It is easy to show the density function is log-concave (replace $x!$ with $\Gamma(x+1)$ everywhere and notice that $\psi^{(1)}(x) = \frac{d^2[\ln(\Gamma(x))]}{dx^2}$ is non-negative for $x > 0$ [1] 6.4.10). But if a step is taken such that an entry is increased from 0 to d the density function can vary by as much as a multiplicative factor of d.

In fact the density function can go through an incredible range. Consider a n by n table with row/column sums all equal to T/n (assume n^2 divides T) and consider two fill-ins E where $E_{i,j} = T/n^2$ and U where $U_{i,i} = T/n$ and $U_{i,j} = 0$ for $i \neq j$.

$$\begin{aligned}
\frac{D(E)}{D(U)} &= \frac{\left(\frac{T}{n}!\right)^n}{\left(\frac{T}{n^2}!\right)^{n^2}} \\
&\geq \frac{\left(\sqrt{2\pi}\left(\frac{T}{n}\right)^{T/n+1/2} e^{-T/n}\right)^n}{\left(\sqrt{2\pi}\left(\frac{T}{n^2}\right)^{T/n^2+1/2} e^{-T/n^2+n/(12T)}\right)^{n^2}} \\
&= \frac{n^{T+n/2}}{\left(\sqrt{2\pi}\right)^{n^2-n}\left(\frac{T}{n^2}\right)^{n^2/2-n/2} e^{n^3/(12T)}}
\end{aligned}$$

which grows as $n^{\Omega(T)}$ for large enough fixed n and growing T.

3.21 Local geometry of Fisher Yates.

3.21.1 Bounds on variation

Consider m row by n column contingency tables, with row sum vector r, column sum vector c and non-negativity. Take the function F defined over all such tables X such that:

$$F(X) = \sum_{i,j} \ln \Gamma(X_{i,j} + 1) \tag{3.29}$$

We note that for $X > 0$ that F is strictly convex [32, 8.363 8].

Lemma 10 *If X is a legal table and $X > 0$ then for any Δ such that $X + \Delta$ is a legal table and $X + \Delta > 0$*

$$F(X + \Delta) - F(X) \leq \sum_{i,j} \frac{\Delta_{i,j}^2}{X_{i,j}} + \sum_{i,j} \Delta_{i,j} \psi(X_{i,j} + 1) \tag{3.30}$$

Furthermore, if for all i, j $\Delta_{i,j}/(X_{i,j}+1) \geq -3/4$ *then*

$$F(X+\Delta) - F(X) \geq \sum_{i,j} \frac{\Delta_{i,j}^2}{5X_{i,j}+3} + \sum_{i,j} \Delta_{i,j}\psi(X_{i,j}+1). \tag{3.31}$$

Proof: First define:

$$f(t) = F(X+t\Delta) = \sum_{i,j} \ln\Gamma(X_{i,j}+1+t\Delta_{i,j})$$

then by [32, 8.360]

$$f'(t) = \sum_{i,j} \Delta_{i,j}\psi(X_{i,j}+1+t\Delta_{i,j})$$

and [32, 8.363 8]

$$\begin{aligned} f''(t) &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}^2 \psi^{(1)}(X_{i,j}+1+t\Delta_{i,j}) \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}^2 \sum_{k=0}^{\infty} \frac{1}{(X_{i,j}+1+t\Delta_{i,j}+k)^2} \end{aligned} \tag{3.32}$$

Also

$$\begin{aligned} f'(t) &= f'(0) + \int_0^t f''(y)\,\mathrm{d}\,y \\ &= \int_0^t f''(y)\,\mathrm{d}\,y + \sum_{i,j} \Delta_{i,j}\psi(X_{i,j}+1) \end{aligned} \tag{3.33}$$

To prove inequality 3.30 assume that $X > 0$ and $X + \Delta > 0$. Continuing from equation 3.32

$$\begin{aligned} f''(t) &\leq \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}^2 \int_0^{\infty} \frac{\mathrm{d}\,y}{(X_{i,j}+t\Delta_{i,j}+y)^2} \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\Delta_{i,j}^2}{X_{i,j}+t\Delta_{i,j}}. \end{aligned}$$

Combining with equation 3.33

$$\begin{aligned}
f'(t) &\leq \int_0^t \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\Delta_{i,j}^2}{X_{i,j}+y\Delta_{i,j}}\,\mathrm{d}\,y + \sum_{i,j}\Delta_{i,j}\psi(X_{i,j}+1) \\
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}(\ln(X_{i,j}+t\Delta_{i,j}) - \ln(X_{i,j})) + \\
&\quad \sum_{i,j}\Delta_{i,j}\psi(X_{i,j}+1).
\end{aligned}$$

So it is enough to get an upper bound on:

$$\begin{aligned}
&\int_0^1 \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (\ln(X_{i,j}+t\Delta_{i,j}) - \ln(X_{i,j}))\Delta_{i,j}\,\mathrm{d}\,t \\
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} ((X_{i,j}+\Delta_{i,j})\ln(X_{i,j}+\Delta_{i,j}) - \\
&\quad (X_{i,j}+\Delta_{i,j}) - \Delta_{i,j}\ln(X_{i,j}) - X_{i,j}\ln(X_{i,j}) + X_{i,j}) \\
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+\Delta_{i,j})(\ln(X_{i,j}+\Delta_{i,j}) - \ln(X_{i,j})) \\
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+\Delta_{i,j})\ln\left(1+\frac{\Delta_{i,j}}{X_{i,j}}\right)
\end{aligned}$$

(the free $\Delta_{i,j}$ disappears from that second to last line because $\sum_{i,j\,:\,\Delta_{i,j}\neq 0}\Delta_{i,j} = 0$). Continuing with [1, 4.133], $\ln(1+x) \leq x$ for $x > -1$:

$$\begin{aligned}
&\leq \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+\Delta_{i,j})\frac{\Delta_{i,j}}{X_{i,j}} \\
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \left(\Delta_{i,j} + \frac{\Delta_{i,j}^2}{X_{i,j}}\right) \\
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\Delta_{i,j}^2}{X_{i,j}}
\end{aligned}$$

To prove inequality 3.31 further assume that $\Delta_{i,j}/(X_{i,j}+1) \geq -3/4$. Again from equation 3.32

$$f''(t) \geq \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}^2 \int_0^\infty \frac{\mathrm{d}\,y}{(X_{i,j}+1+t\Delta_{i,j}+y)^2}$$

$$= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\Delta_{i,j}^2}{X_{i,j}+1+t\Delta_{i,j}}.$$

Combining with equation 3.33

$$\begin{aligned} f'(t) &\geq \int_0^t \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\Delta_{i,j}^2}{X_{i,j}+1+y\Delta_{i,j}}\,\mathrm{d}\,y \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}(\ln(X_{i,j}+1+t\Delta_{i,j})-\ln(X_{i,j}+1)). \end{aligned}$$

Now it is enough to get a lower bound on:

$$\begin{aligned} &\int_0^1 \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (\ln(X_{i,j}+1+t\Delta_{i,j})-\ln(X_{i,j}+1))\Delta_{i,j}\,\mathrm{d}\,t \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} ((X_{i,j}+1+\Delta_{i,j})\ln(X_{i,j}+1+\Delta_{i,j}) - \\ &\quad (X_{i,j}+1+\Delta_{i,j}) - \Delta_{i,j}\ln(X_{i,j}+1) - \\ &\quad (X_{i,j}+1)\ln(X_{i,j}+1) + X_{i,j}+1) \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+1+\Delta_{i,j})(\ln(X_{i,j}+1+\Delta_{i,j})-\ln(X_{i,j}+1)) \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+1+\Delta_{i,j})\ln\left(1+\frac{\Delta_{i,j}}{X_{i,j}+1}\right) \end{aligned}$$

(the free $\Delta_{i,j}$ disappears from that second to last line because $\sum_{i,j\,:\,\Delta_{i,j}\neq 0}\Delta_{i,j}=0$). Using $\ln(1+x)\geq x/(1+\frac{2}{3}x)$ for $x\geq -3/4$.

$$\begin{aligned} &\geq \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+1+\Delta_{i,j})\frac{\frac{\Delta_{i,j}}{X_{i,j}+1}}{1+\frac{2}{3}\frac{\Delta_{i,j}}{X_{i,j}+1}} \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} (X_{i,j}+1+\Delta_{i,j})\frac{\frac{\Delta_{i,j}}{X_{i,j}+1}}{\frac{X_{i,j}+1+\Delta_{i,j}-\frac{\Delta_{i,j}}{3}}{X_{i,j}+1}} \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}\frac{X_{i,j}+1+\Delta_{i,j}}{X_{i,j}+1+\Delta_{i,j}-\frac{\Delta_{i,j}}{3}} \\ &= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \Delta_{i,j}\left(1+\frac{\frac{\Delta_{i,j}}{3}}{X_{i,j}+1+\Delta_{i,j}-\frac{\Delta_{i,j}}{3}}\right) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\frac{\Delta_{i,j}^2}{3}}{X_{i,j}+1+\Delta_{i,j}-\frac{\Delta_{i,j}}{3}} \\
&\geq \sum_{i,j\,:\,\Delta_{i,j}\neq 0} \frac{\Delta_{i,j}^2}{5X_{i,j}+3}
\end{aligned}
$$

□

We have the following theorem:

Theorem 11 *If X is a legal table minimizing equation 3.29 (over all legal tables) and $X > 0$ then for any Δ such that $X + \Delta$ is a legal table and $X + \Delta > 0$*

$$F(X+\Delta) - F(X) \leq \sum_{i,j} \frac{\Delta_{i,j}^2}{X_{i,j}} \tag{3.34}$$

Furthermore, if for all i, j $\Delta_{i,j}/(X_{i,j}+1) \geq -3/4$ then

$$F(X+\Delta) - F(X) \geq \sum_{i,j} \frac{\Delta_{i,j}^2}{5X_{i,j}+3} \tag{3.35}$$

Proof: It is clear that $\sum_{i,j} \Delta_{i,j}\psi(X_{i,j}+1) = 0$ in this case and we get the result from lemma 10. □

3.21.2 Location of the minimum

Let X be a m by n non-negative matrix obeying row sums r, column sums c and $T = \sum_{i=1}^m r_i = \sum_{j=1}^n c_j$. Define a m by n matrix, $S(X)$, as follows:

$$S(X)_{i,j} \stackrel{\text{def}}{=} \begin{cases} +1 & X_{i,j} \geq \frac{r_i c_j}{T} + 1 \\ -1 & X_{i,j} \leq \frac{r_i c_j}{T} - 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.36}$$

We call an m by n matrix Δ a "non-trivial balanced sub marking" of $S(X)$ if:

- $\exists i \in [1 \cdots m]\ \exists j \in [1 \cdots n]$ s.t. $\Delta_{i,j} \neq 0$ ("non-trivial")
- $\begin{array}{lll} \forall i \in [1 \cdots m] & \sum_{j=1}^n \Delta_{i,j} & = 0 \\ \forall j \in [1 \cdots n] & \sum_{i=1}^m \Delta_{i,j} & = 0 \end{array}$ ("balanced")

- $\forall i \in [1 \cdots m] \; \forall j \in [1 \cdots n] \;\; (\Delta_{i,j} = S(X)_{i,j}) \vee (\Delta_{i,j} = 0)$ ("sub-marking").

Lemma 11 *Take X is a m by n non-negative matrix obeying row/column sums r, c, $T = \sum_i r_i = \sum_j c_j$, Δ a non-trivial balanced sub marking of $S(X)$ and $\forall i, j \;\; \frac{r_i c_j}{T} > 1$. Then*

$$F(X) > F(X - \Delta).$$

Proof: Let $\epsilon_{i,j} = X_{i,j} - \left(\frac{r_i c_j}{T} + \Delta_{i,j}\right)$. Notice that $(\epsilon_{i,j} < 0) \rightarrow (\Delta_{i,j} < 0)$ and $(\epsilon_{i,j} > 0) \rightarrow (\Delta_{i,j} > 0)$. Then

$$\begin{aligned}
F(X) - F(X - \Delta) &= \sum_{i,j} \left(\ln \Gamma\left(\frac{r_i c_j}{T} + \Delta_{i,j} + \epsilon_{i,j} + 1\right) - \right. \\
&\quad \left. \ln \Gamma\left(\frac{r_i c_j}{T} + \epsilon_{i,j} + 1\right)\right) \\
&= \sum_{i,j \,:\, \Delta_{i,j} > 0} \ln\left(\frac{r_i c_j}{T} + \epsilon_{i,j} + 1\right) - \\
&\quad \sum_{i,j \,:\, \Delta_{i,j} < 0} \ln\left(\frac{r_i c_j}{T} + \epsilon_{i,j}\right) \\
&\geq \sum_{i,j \,:\, \Delta_{i,j} > 0} \ln\left(\frac{r_i c_j}{T}\right) - \sum_{i,j \,:\, \Delta_{i,j} < 0} \ln\left(\frac{r_i c_j}{T} - 1\right) \\
&> \sum_{i,j \,:\, \Delta_{i,j} > 0} \ln\left(\frac{r_i c_j}{T}\right) - \sum_{i,j \,:\, \Delta_{i,j} < 0} \ln\left(\frac{r_i c_j}{T}\right) \\
&= 0.
\end{aligned}$$

□

Lemma 12 *If $S(X)$ is such that every row and every column has at least one $+1$ and at least one -1 then $S(X)$ has a non-trivial balanced sub marking.*

Proof: We can assume (without loss of generality) that $S(X)_{1,1} = +1$, $S(X)_{1,2} = -1$ and $S(X)_{2,1} = -1$. This gets us into the general case used in this proof where we have Δ_k the m by n matrix such that

$$(\Delta_k)_{i,j} = \begin{cases} +1 & (i,j) = (1,1) \\ (-1)^{\min(i,j)} & (1 \leq i \leq k) \wedge (1 \leq j \leq k) \wedge (i = j \pm 1) \\ 0 & \text{otherwise} \end{cases}$$

and $((\Delta_k)_{i,j} \neq 0) \rightarrow ((\Delta_k)_{i,j} = S(X)_{i,j})$. We are going to use an inductive argument on k. One such pattern ($n = m = 5, k = 4$) is:

$$\begin{bmatrix} +1 & -1 & 0 & 0 & 0 \\ -1 & 0 & +1 & 0 & 0 \\ 0 & +1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now if we know, as above, the initial k by k segment or $S(X)$ we know (by the given) there must be a $(-1)^k$ somewhere in column k. If $S(X)_{k,k} = (-1)^k$ then we see that if $(-1)^k$ is added to the (k, k) position of Δ_k then we have the desired non-trivial balanced sub marking. If $S(X)_{i,k} = (-1)^k$ for some $i < k - 1$ then we see that if we add $(-1)^k$ to the (i, k) position of Δ_k, zero out the column $a < k$ such that $(\Delta_k)_{i,a} = (-1)^k$ and then iteratively zero out all columns that have the only non-zero entry for any row then what we have left is again the desired non-trivial balanced sub marking. An example of this process is ($n = m = 5, k = 5, S(X)_{1,5} = -1$):

$$\begin{bmatrix} +1 & 0 & 0 & 0 & -1 \\ -1 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & +1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

So either $k = m$ and we must have a non-trivial balanced sub marking or we can assume (by reordering rows) that $S(X)_{k+1,k} = (-1)^k$. By a similar column argument we see that either we have the desired marking or (by reordering columns) $S(X)_{k,k+1} = (-1)^k$. Thus we can say that either we are done or Δ_{k+1} is such that $((\Delta_{k+1})_{i,j} \neq 0) \rightarrow ((\Delta_{k+1})_{i,j} = S(X)_{i,j})$. □

Theorem 12 *If X is the m by n non-negative matrix obeying row/column sums r, c minimizing equation 3.29 over all such tables and $T = \sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$ then for all i, j:*

$$|X_{i,j} - \frac{r_i c_j}{T}| \leq (m + n - 3) \max(m - 1, n - 1).$$

Proof: We proceed by induction on m and n. If m or n is less than two then the theorem is trivial. For $m = n = 2$ the theorem follows immediately from lemma 11. Assume the theorem is true for all a, b s.t. $((a < m) \vee (b < n)) \wedge (1 \leq a \leq m) \wedge (1 \leq b \leq n)$. Suppose every row and every column of X has an entry $X_{i,j}$ differing from $\frac{r_i c_j}{T}$ by at least max$(m-1, n-1)$. Then $S(X)$ must satisfy the conditions of lemma 12 allowing us to again apply lemma 11. So we assume that there is some row or column where every entry is within max$(m-1, n-1)$ of $\frac{r_i c_j}{T}$, for discussion assume it is row m. Notice that the $m-1$ by n initial submatrix of X is the unique matrix minimizing equation 3.29 over all non-negative $m-1$ by n matrices satisfying row sums r and column sums $c_j - X_{m,j}$. By our inductive hypothesis we know that $|X_{i,j} - \frac{r_i(c_j - X_{m,j})}{T - r_m}| \leq$ max$(m-2, n-1)(m+n-4)$ for $i = 1 \cdots m-1$. We quickly see that for $i < m$ we have $|X_{i,j} - \frac{r_i c_j}{T}| \leq \left(\frac{r_i}{T - r_m} + m + n - 4\right)$ max$(m-1, n-1)$. And for $i < m$ we have $\frac{r_i}{T - r_m} \leq 1$. $\Box$

Chapter 4

Open Problems

We would like to identify some important open problems in the areas touched on in this thesis. For the Markov chains (used in both the optimization chapters and contingency table chapters of the thesis) a number of important questions remain open.

4.1 Effective diameter.

Most of the current upper bounds on the mixing time for Markov chains with average step size δ whose states are contained in a convex body of diameter d involve a factor of $(d/\delta)^2$. It would be a big breakthrough to be able to measure both d and δ in the infinity (or max) metric without introducing factors of n. It would also be a big improvement to be able to use the "effective diameter" or the expected distance between two points in K instead of d which is the maximum distance between any two points in K. Any of these improvements would lead to much better bounds for mixing time.

4.2 Non-stationary processes.

A method to analyze non-stationary processes would be very useful. This would not only allow us to directly prove mixing rates for simulate annealing but would allow the development of adaptive algorithms. An adaptive algorithm would start with a simple Markov chain that could have a poor mixing rate due to states with low local conductance. Such a chain could be repaired when detected by forcing smaller steps near the bad states encountered. If good time bounds could be developed for such a scheme then one would only have to run the Markov chain for a time proportional to the rate bad states are actually encountered instead of proportional to a weak upper bound on this rate.

4.3 Convergence acceleration.

The linear operator interpretation of Markov chains opens the question if any of the many methods of convergence acceleration methods used to accelerate iterative linear systems can be applied to Markov chains. I do not know if this is possible but have thought about applying Chebyshev acceleration[35] to a Markov chain.

The idea in Chebyshev acceleration is that if one has a process $X_{i+1} = (X_i P)/(\|X_i P\|_1)$ that starts with X_0 and P has a unique eigenvector, π, with (largest) eigenvalue 1 such that $\pi = \lim_{t \to \infty} X_t$. Then if one applies the operator t times to get the vectors $X_0, X_1, \cdots X_t$ then X_t is *not* the best estimate for π. There is a estimate $\hat{\pi} = \sum_{i=0}^{t} \lambda_i X_i$ where the λ_i are explicit constants independent of X_i and P that is a better estimate of π than X_t is. Not all the λ_i are non-negative.

The problem for Markov chains is that we realize X_i as a probability vector, so the subtractions needed to compute $\hat{\pi}$ have no natural interpretation in this context. In addition the Markov chain is only able to estimate X_i- so if the λ_i are large (and they are) then errors in estimating X_i soon make $\hat{\pi}$ unusable.[1]

4.4 Unary polynomial time in row/column sums.

For the contingency table problem it would be nice to develop chains that ran in unary polynomial time in T, that total of all rows and columns. The current inability to do this represents a major weakness in current uniform generation techniques for contingency tables. A unary polynomial time algorithm to generate contingency tables from the uniform distribution would be useful for two simple reasons. First, even though T can be vary large in principle it is typically a number that some statistician has counted up to. In practice many contingency tables are constructed as summary statistics of lists of people. So even though it

[1]Chebyshev acceleration seems to improve bias at the expense of amplifying variance. In explicit eigenvector problems the only source of variance is rounding error- so this is a good trade. In random processes the variance is large to begin with.

takes no more that $nm\log(T)$ space to write down such a table it often took the statistician time proportional to T to generate it. Thus an algorithm that runs in time polynomial in T would be very useful to a statistician, though the computer scientist would of course like to see one that runs in time polynomial in $\log(T)$. In this same vein there is an obvious and efficient method to generate tables from the Fisher-Yates or hypergeometric distributions in time polynomial in T. The method is just to build a large complete bipartite graph with T left nodes and T right nodes. One labels each right node with an index from $1\cdots m$ such that r_i right nodes have index i. One labels each left node with an index from $1\cdots n$ such that c_j right nodes have index j. Then one chooses at random, from the uniform distribution, a perfect matching of the left to right nodes. The matrix X such that $X_{i,j}$ is the number of right nodes marked with i that are matched up with left nodes marked with j is then a contingency table consistent with (r, c) and is Fisher-Yates distributed. It would be nice to have a comparable algorithm for the uniform case.

4.5 Higher way contingency tables.

Contingency tables representing relationships between more the two variables have been proposed. There is some freedom in defining what constraints are in three (and higher) way tables. If we define $k-$way contingency tables to be k dimensional block of non-negative integers $X_{i_1,i_2,\cdots,i_k}$ constrained such that

$$\sum_{i_l} X_{i_1,i_2,\cdots,i_k} = c_{i_1,\cdots,i_{l-1},*,i_{l+1},\cdots,i_k} \quad \forall l$$

where the $c_{i_1,\cdots,i_{l-1},*,i_{l+1},\cdots,i_k}$ are constants. Then there is no known scheme for approximately counting the number of constrained 3-way tables as this would be sufficient to approximately compute the general permanent.[2] We show this by encoding the set of perfect matchings of an arbitrary bipartite graph in a 3-way table. Let G be a bipartite graph with vertex sets V_1, V_2 and edges set $E \subseteq V_1 \times V_2$. We define X, a 3 way table with dimensions $|V_1| \times |V_2| \times 2$, and set

$$\begin{aligned} c_{*,j,1} &= 1 \\ c_{i,*,1} &= 1 \end{aligned}$$

[2]Sinclair and Jerrum settled the question of computing the dense permanent, the general case remains open.

$$
\begin{aligned}
c_{i,j,*} &= \begin{cases} 1 & (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \\
c_{*,j,2} &= \sum_i c_{i,j,*} - 1 \\
c_{i,*,2} &= \sum_j c_{i,j,*} - 1
\end{aligned}
$$

Then for any X obeying these constraints we see that the set of (i, j) such that $X_{i,j,1} \neq 0$ is a perfect matching of G and all perfect matchings of G can be so encoded.

Similarly, a 4-way table has enough structure to encode a three dimensional matching, so the decision problem itself is *NP* complete.

4.5.1 Simpson's paradox

One might wonder if there is any actual necessity for higher way tables. Perhaps one could collapse the information in a 3 way table into a two way table with a comparable number of cells and still perform a meaningful analysis. We present here an example, exploiting the well known Simpson's paradox, that should illustrate that this is not the case.

Consider the 3 way table in table 4.1, which could arise from a drug trial run in two cities.

City A	Good	Bad
Drug	67	33
Control	590	410

City B	Good	Bad
Drug	550	450
Control	50	50

Table 4.1: Drug trial in two cities.

This data can obviously be organized into a 3 way table indexed by treatment (drug/control), result (good/bad) and city (A/B). The obvious way to pair this data into a two dimensional table would be to add the city A data to the city B data to form a summary table like table 4.2.

The problem is that there is no way to tell if important information has been thrown away in this summarizing step. In table 4.1 the drug has a lower rate of bad effects than control in both city A and city B. However in table 4.2 we see that drug has a higher rate of bad effects than control over all. It is impossible to determine from only examining the data if summarizing table 4.1 into table 4.2 was legitimate. If one believes that the populations in city A and city B are identical and that the selection process that divide people into drug and control trials was independent of any properties of the city populations (like only city B got funding for a big drug trial) then one might want to use the summary data. If one does not believe that the cities have identical populations and an indifferent selection mechanism then one can not draw any conclusions because for all practical purposes city A ran only a control group and city B ran only a drug group.

Cities A + B

	Good	Bad
Drug	617	483
Control	640	460

Table 4.2: Drug trial summary.

Bibliography

[1] ABRAMOWITZ, MILTON; STEGUN, IRENE, "Handbook of Mathematical Functions", Dover Publications, New York, 1972.

[2] AGRESTI, ALAN, "Categorical Data Analysis", Wiley-Interscience, New York, 1990.

[3] ALDOUS, D.; DIACONIS, P., *Shuffling Cards and Stopping Times*, Amer. Math. Monthly, Num 93, pp. 333-348.

[4] APPLEGATE, D. *Sampling, Integration and Computing the Volume*, PhD Thesis, School of Computer Science, Carnegie Mellon University, December 1991, CMU-CS-91-207.

[5] APPLEGATE, DAVID; KANNAN, RAVI, *Sampling and integration of near log-concave functions*, 23rd ACM Symposium on the Theory of Computing, 1991.

[6] BAGLIVO, JENNY; OLIVIER, DONALD; PAGANO, MARCELLO, *Methods for Exact Goodness-of-Fit Tests*, Journal of the American Statistical Association, Vol. 87, No. 418, June 1992.

[7] BAGLIVO, JENNY; OLIVIER, DONALD; PAGANO, MARCELLO, *Methods for the Analysis of Contingency Tables with Large and Small Cell Counts*, Journal of the American Statistical Association, Vol. 83, No. 404, December 1988.

[8] BARVINOK, A.I., *A Polynomial Time Algorithm for Counting Integral Points in Polyhedra When the Dimension is Fixed*, 34th Symposium on Foundations of Computer Science, IEEE Computer Society Press, Nov 1993.

[9] BÉLISLE, C.; ROMEIJN, H.; SMITH, R. *Hit-And-Run Algorithms For Generating Multivariate Distributions*, Mathematics of Operations Research, Vol 18 No. 2, pp. 255-266, 1993.

[10] BERGER, M., "Geometry", (translated by M. Cole and S. Levy), Vols I and II, Springer-Verlag, New York, 1987.

[11] BILLERA, L.J.; STURMFELS, B., *Fiber polytopes*, Annals of Mathematics, 135, pp. 527-549, 1992.

[12] BILLINGSLEY, P., "Probability and Measure," Second Edition, Wiley, New York, 1986.

[13] BLAKLEY, S., *Combinatorial remarks on partitions of a multipartite number*, Duke Math. J. #31 (1964), pp 335-340.

[14] BOPPANA, R. B. AND M. M. HALLDÓRSSON. *Approximating maximum independent sets by excluding subgraphs*, Proc. of 2nd Scand. Workshop on Algorithm Theory, pp. 13-25, Springer-Verlag Lecture Notes in Computer Science #447, July, 1990.

[15] DAHMEN, W., MICCHELLI, C.A. *The number of solutions to linear Diophantine equations and multivariate splines*, Transactions Amer. Math. Soc., 308 (1988) 509–532.

[16] DIACONIS, PERSI; EFRON, BRADLEY, *Testing For Independence in a Two-Way Table: New Interpretations of the Chi-Square Statistic*, Annals of Statistics, Vol 13, No. 3, pp. 845-874, 1985.

[17] DIACONIS, PERSI; EFRON, BRADLEY, *Testing For Independence in a Two-Way Table: Rejoinder*.

[18] DIACONIS, P.; GANGOLLI, A., *Rectangular Arrays with Fixed Margins*, manuscript, April, 1994.

[19] DIACONIS, P.; STROOCK, D. *Geometric Bounds for Eigenvalues of Markov Chains*, Annals of Applied Probability, 1991, Vol. 1, No. 1, pp. 36-61.

[20] DIACONIS, PERSI; STURMFELS, BERND, *Algebraic Algorithms for Sampling from Conditional Distributions*, pre-print.

[21] DYER, MARTIN; FRIEZE, ALAN, *Computing the volume of convex bodies: a case where randomness provably helps*, Proceedings of Symposia in Applied Mathematics, Vol 44, 1991.

[22] DYER, MARTIN; FRIEZE, ALAN, *On the complexity of computing the volume of a polyhedron*, SIAM J. Comput., **17** (1988), pp. 27-37.

[23] DYER, MARTIN; FRIEZE, ALAN; KANNAN, RAVI, *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.

[24] DYER, MARTIN; KANNAN, RAVI; MOUNT, JOHN, *unpublished manuscript*, 1995.

[25] DYER, MARTIN; STOUGIE, LEEN, *personal communication*, 1994.

[26] FAIGLE, U.; GADEMANN, N.; KERN, W., *A Random Polynomial Time Algorithm for Well-Rounding Convex Bodies*, to appear: Discrete Applied Mathematics.

[27] FILL, JAMES, *Eigenvalue Bounds on Convergence to Stationarity for Nonreversible Markov Chains, with an Application to the Exclusion Process.*, The Annals of Applied Probability, Vol. 1. No. 1, 1991.

[28] FRIEZE, A.; KANNAN, R.; POLSON, N., *Sampling from Log-Concave Distributions*, The Annals of Applied Probability, Vol 4 No 3, pp. 812-837, 1994.

[29] FULTON, W., "Introduction to Toric Varieties", Princeton University Press, Princeton, 1993.

[30] GAIL, M.; MANTEL, N., *Counting the number of $r \times c$ contingency tables with fixed margins,* Jour. Amer. Statist. Assoc. **72**, pp. 859-862, 1977.

[31] GAREY, MICHAEL; JOHNSON, DAVID "Computers and Intractability, A Guide to the Theory of NP-Completeness," W.H. Freeman, New York, 1979.

[32] GRADSHTEYN, I. S., RYZHIK, I.M., *Table of Integrals, Series, and Products*, Corrected and Enlarged edition, Academic Press, 1980.

[33] GRÖTSCHEL, MARTIN; LOVÁSZ, LÁSZLÓ; SCHRIJVER, ALEXANDER, "Geometric Algorithms and Combinatorial Optimization", Springer-Verlag, New York, 1988.

[34] GRUBER, P.M; LEKKERKERKER, C.G., "Geometry of Numbers", Second Edition, North-Holland, New York, 1987.

[35] HAGEMAN, YOUNG, "Applied Interative Methods", Academic Press.

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

[37] JAYARAMAN, J.; SRINIVASAN, R.; ROUNDY, R.; TAYUR, S. *Procurement of Common Components in a Stochastic Environment*, IBM Technical Report, November 1992.

[38] JERRUM, M.R.; VALIANT, L. G.; VAZIRANI, V.V., *Random Generation of Combinatorial Structures from a Uniform Distribution*, Theoretical Computer Science 43 (1986), 169-188.

[39] KANNAN, RAVI; MOUNT, JOHN; TAYUR, SRIDHAR *A Randomized Algorithm to Optimize Over Certain Convex Sets*, Mathematics of Operations Research, Vol. 20 (May 1995).

[40] LOVÁSZ, LÁSZLÓ; SIMONOVITS, MIKLÓS, *Random Walks in a Convex Body and an Improved Volume Algorithm*, Revised version, March 1993.

[41] LUBY, MICHAEL; NISAN, NOAM, *A Parallel Approximation Algorithm for Positive Linear Programming* .

[42] MEHTA, CYRUS R.; PATEL, NITIN R., *A Network Algorithm for Performing Fisher's Exact Test in $r \times c$ Contingency Tables* Journal of the American Statistical Association, Vol. 78, No. 382, June 1983.

[43] NEMIROVSKII, ARKADII SEMENOVICH; YUDIN, DAVID BORISOVICH, "Problem complexity and method efficiency in optimization", (translated by E.R. Dawson), Wiley, 1983.

[44] O'NEIL, P., *Asymptotics and random matrices with row-sum and column-sum restrictions,* Bull. Amer. Math. Soc. **75**, pp. 1276-1282, 1969.

[45] PRESS, W.; FLANNERY B.; TEUKOLSKY, S.; VETTERLING, W., "Numerical Recipes in C", University of Cambridge Press, 1988.

[46] SCHRIJVER, ALEXANDER, "Theory of Linear and Integer Programming", Wiley, New York, 1986.

[47] SINCLAIR, ALISTAIR; JERRUM, MARK *Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains*, Information and Computation 82, pp. 93-133, 1989.

[48] SMITH, R. *Efficient Monte Carlo Procedures for Generating Points Uniformly Distributed over Bounded Regions*, Operations Research, Vol 32 No. 6, pp. 1296-1308, 1984.

[49] SNEE *Graphical display of two-way contingency tables*, Amer. Statis. 38, pp. 9-12, 1974.

[50] STOUT, P., PhD thesis, Carnegie-Mellon Univ 1994.

[51] STURMFELS, B., *How to Count Integer Matrices*, December 13, 1993.

[52] STURMFELS, B., *On Vector Partition Functions*, March 29, 1994.

[53] VAIDYA, P., *A new Algorithm for Minimizing Convex Functions Over Convex Sets*, 30th Symposium on Foundations of Computer Science, IEEE Computer Society Press, 1989.

[54] WETS, R. *Stochastic Programming* in Handbooks in Operations Research and Management Science, Vol. 1, North Holland, 1989.

[55] ZABINSKY, B.; SMITH, R., *Pure adaptive search in global optimization*, Mathematical Programming 53, pp. 323-338, 1992.

[56] ZABINSKY, B.; SMITH, R.; MCDONALD, J.; ROMEIJN, H.; KAUFMAN, D., *Improving Hit-and-Run for Global Optimization*, Journal of Global Optimization, 3, pp. 171-192, 1993.

Appendix A

Notation

Notation used in this thesis.

$B_r(x)$	The ball of radius r centered at x.
$C_r(x)$	The cube of side $2r$ centered at x.
$C(u)$	Shorthand for $C_{1/2}(x)$.
$\delta_{i,j}$	Kronecker delta, 1 if $i = j$, 0 otherwise.
E^i	The i'th standard unit vector, $E^i_j = \delta_{i,j}$.
$\mathrm{erf}(x)$	Error function.
$\Gamma(x)$	Euler gamma function.
$\mathrm{Lin}(x_1, \cdots, x_n)$	Linearity space of $x_1, \cdots, x_n$.
$T(r_1, \cdots r_m; c_1, \cdots c_n)$	The set of all m by n non-negative integer matrices with rows sums $r_1, \cdots r_m$ and column sums $c_1, \cdots, c_n$.
$\mathbf{Q}^n$	Rational n space.
$\mathbf{Q}^{n+}$	The set of $x \in \mathbf{Q}^n$ such that $x \geq 0$.
$\mathbf{R}^n$	Euclidean n space.
$\mathbf{R}^{n+}$	The set of $x \in \mathbf{R}^n$ such that $x \geq 0$.
Z^n	The standard integer lattice in n space.
$\sharp(X)$	The number of elements in the set X.
$\sharp(r; c)$	Shorthand for $\sharp(T(r; c))$.
$< x_1, \cdots, x_n >$	Group generated by $x_1, \cdots, x_n$.
$\|X\|$	The number of elements in the set X or length of X (depending if X is a set, scalar or vector).

List of Figures

List of Tables

Appendix B

Complete 3×3 solution

The complete enumeration oracle for 3 by 3 contingency tables is used as follows. One starts with row sums $[x_1, x_2, x_3]$ and column sums $[x_4, x_5, x_1+x_2+x_3-x_4-x_5]$ and arranges them (by row swaps, columns swaps and transpose) to obey all of the relations in table B.1 (the table margins are sorted so the row/column demands are non-decreasing and the first row demand in no bigger than the first column demand).

$$\begin{array}{lrrrrrll}
[& 1 & 0 & 0 & -1 & 0 &] \cdot x & \leq & 0 \\
[& -1 & -1 & -1 & 1 & 2 &] \cdot x & \leq & 0 \\
[& 0 & 0 & 0 & 1 & -1 &] \cdot x & \leq & 0 \\
[& 0 & 1 & -1 & 0 & 0 &] \cdot x & \leq & 0 \\
[& 1 & -1 & 0 & 0 & 0 &] \cdot x & \leq & 0
\end{array}$$

Table B.1: Global relations for 3 by 3 contingency tables.

Then one navigates down the decision tree, figure B.1, moving down a dashed arc when an inequality from table B.2 is violated and down a solid arc otherwise.

The chamber labeled by the leaf reached then has the correct polynomial and the values $x_1, \cdots, x_5$ are substituted into the polynomial to get the desired count. A complete list of the 3 by 3 polynomials follows.

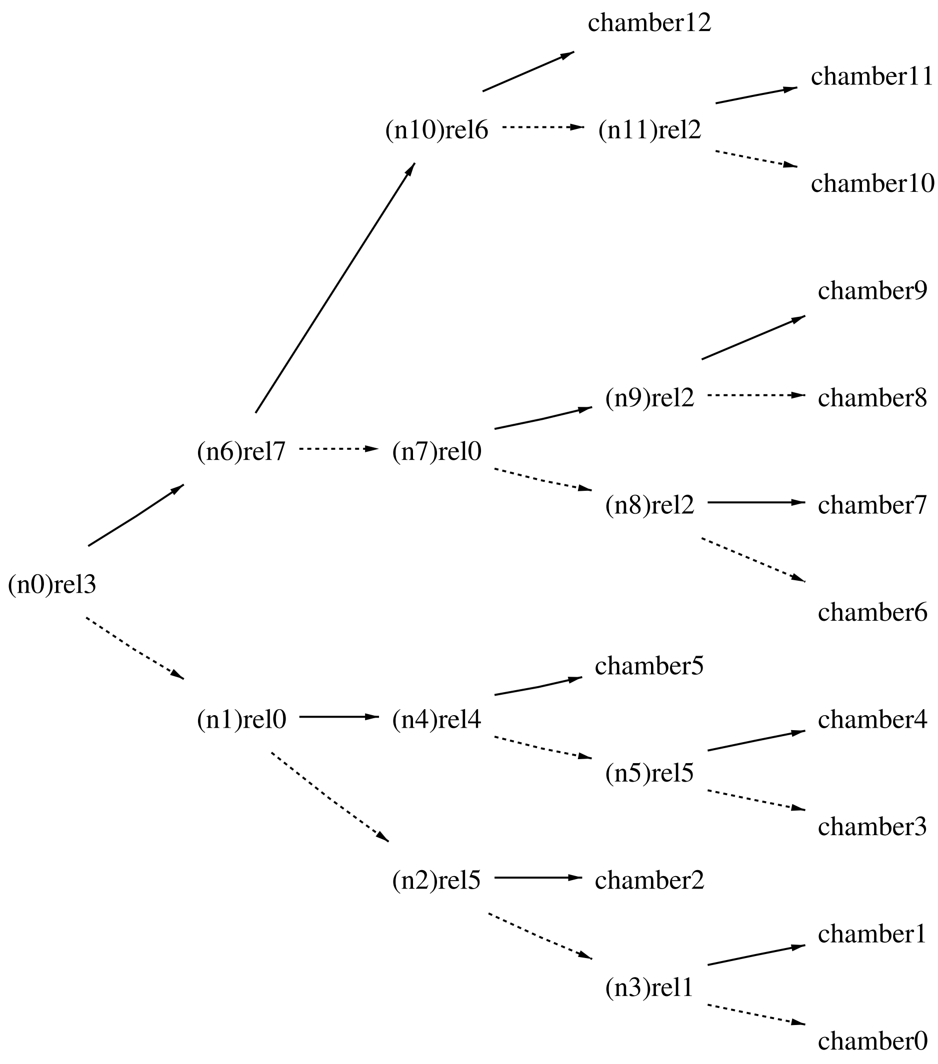


Figure B.1: Decision tree for 3 by 3 contingency tables.

$$
\begin{array}{lrrrrrrl}
\text{rel0}: & [\; 0 & 0 & 1 & -1 & -1 \;] \cdot x & \leq & 0 \\
\text{rel1}: & [\; 0 & 1 & 0 & -1 & -1 \;] \cdot x & \leq & 0 \\
\text{rel2}: & [\; 0 & 1 & 0 & -1 & 0 \;] \cdot x & \leq & 0 \\
\text{rel3}: & [\; 0 & 1 & 0 & 0 & -1 \;] \cdot x & \leq & 0 \\
\text{rel4}: & [\; 1 & 0 & 1 & -1 & -1 \;] \cdot x & \leq & 0 \\
\text{rel5}: & [\; 1 & 1 & 0 & -1 & -1 \;] \cdot x & \leq & 0 \\
\text{rel6}: & [\; 1 & 1 & 0 & -1 & 0 \;] \cdot x & \leq & 0 \\
\text{rel7}: & [\; 1 & 1 & 0 & 0 & -1 \;] \cdot x & \leq & 0
\end{array}
$$

Table B.2: Decision tree inequalities for 3 by 3 contingency tables.

chamber 0: $\frac{1}{24}x_1^4 -\frac{1}{6}x_1^3x_4 -\frac{1}{6}x_1^3x_5 +\frac{1}{2}x_1^2x_4x_5 -\frac{1}{4}x_1^3 +\frac{3}{2}x_1x_4x_5 -\frac{13}{24}x_1^2 +\frac{7}{6}x_1x_4 +\frac{7}{6}x_1x_5 +x_4x_5 +\frac{3}{4}x_1 +x_4 +x_5 +1$

chamber 1: $\frac{1}{24}x_1^4 -\frac{1}{6}x_1^3x_4 -\frac{1}{6}x_1^3x_5 +\frac{1}{2}x_1^2x_4x_5 -\frac{1}{24}x_2^4 +\frac{1}{6}x_2^3x_4 +\frac{1}{6}x_2^3x_5 -\frac{1}{4}x_2^2x_4^2 -\frac{1}{2}x_2^2x_4x_5 -\frac{1}{4}x_2^2x_5^2 +\frac{1}{6}x_2x_4^3 +\frac{1}{2}x_2x_4^2x_5 +\frac{1}{2}x_2x_4x_5^2 +\frac{1}{6}x_2x_5^3 -\frac{1}{24}x_4^4 -\frac{1}{6}x_4^3x_5 -\frac{1}{4}x_4^2x_5^2 -\frac{1}{6}x_4x_5^3 -\frac{1}{24}x_5^4 -\frac{1}{4}x_1^3 +\frac{3}{2}x_1x_4x_5 +\frac{1}{4}x_2^3 -\frac{3}{4}x_2^2x_4 -\frac{3}{4}x_2^2x_5 +\frac{3}{4}x_2x_4^2 +\frac{3}{2}x_2x_4x_5 +\frac{3}{4}x_2x_5^2 -\frac{1}{4}x_4^3 -\frac{3}{4}x_4^2x_5 -\frac{3}{4}x_4x_5^2 -\frac{1}{4}x_5^3 -\frac{13}{24}x_1^2 +\frac{7}{6}x_1x_4 +\frac{7}{6}x_1x_5 -\frac{11}{24}x_2^2 +\frac{11}{12}x_2x_4 +\frac{11}{12}x_2x_5 -\frac{11}{24}x_4^2 +\frac{1}{12}x_4x_5 -\frac{11}{24}x_5^2 +\frac{3}{4}x_1 +\frac{1}{4}x_2 +\frac{3}{4}x_4 +\frac{3}{4}x_5 +1$

chamber 2: $-\frac{1}{12}x_1^4 -\frac{1}{3}x_1^3x_2 +\frac{1}{6}x_1^3x_4 +\frac{1}{6}x_1^3x_5 -\frac{1}{4}x_1^2x_2^2 +\frac{1}{2}x_1^2x_2x_4 +\frac{1}{2}x_1^2x_2x_5 -\frac{1}{4}x_1^2x_4^2 -\frac{1}{4}x_1^2x_5^2 -\frac{1}{2}x_1^3 -\frac{3}{4}x_1^2x_2 +\frac{3}{4}x_1^2x_4 +\frac{3}{4}x_1^2x_5 -\frac{3}{4}x_1x_2^2 +\frac{3}{2}x_1x_2x_4 +\frac{3}{2}x_1x_2x_5 -\frac{3}{4}x_1x_4^2 -\frac{3}{4}x_1x_5^2 -\frac{5}{12}x_1^2 +\frac{1}{12}x_1x_2 +\frac{13}{12}x_1x_4 +\frac{13}{12}x_1x_5 -\frac{1}{2}x_2^2 +x_2x_4 +x_2x_5 -\frac{1}{2}x_4^2 -\frac{1}{2}x_5^2 +x_1 +\frac{1}{2}x_2 +\frac{1}{2}x_4 +\frac{1}{2}x_5 +1$

chamber 3: $\frac{1}{24}x_1^4 -\frac{1}{6}x_1^3x_4 -\frac{1}{6}x_1^3x_5 +\frac{1}{2}x_1^2x_4x_5 -\frac{1}{24}x_2^4 +\frac{1}{6}x_2^3x_4 +\frac{1}{6}x_2^3x_5 -\frac{1}{4}x_2^2x_4^2 -\frac{1}{2}x_2^2x_4x_5 -\frac{1}{4}x_2^2x_5^2 +\frac{1}{6}x_2x_4^3 +\frac{1}{2}x_2x_4^2x_5 +\frac{1}{2}x_2x_4x_5^2 +\frac{1}{6}x_2x_5^3 -\frac{1}{24}x_3^4 +\frac{1}{6}x_3^3x_4 +\frac{1}{6}x_3^3x_5 -\frac{1}{4}x_3^2x_4^2 -\frac{1}{2}x_3^2x_4x_5 -\frac{1}{4}x_3^2x_5^2 +\frac{1}{6}x_3x_4^3 +\frac{1}{2}x_3x_4^2x_5 +\frac{1}{2}x_3x_4x_5^2 +\frac{1}{6}x_3x_5^3 -\frac{1}{12}x_4^4 -\frac{1}{3}x_4^3x_5 -\frac{1}{2}x_4^2x_5^2 -\frac{1}{3}x_4x_5^3 -\frac{1}{12}x_5^4 -\frac{1}{4}x_1^3 +\frac{3}{2}x_1x_4x_5 +\frac{1}{4}x_2^3 -\frac{3}{4}x_2^2x_4 -\frac{3}{4}x_2^2x_5 +\frac{3}{4}x_2x_4^2 +\frac{3}{2}x_2x_4x_5 +\frac{3}{4}x_2x_5^2 +\frac{1}{4}x_3^3 -\frac{3}{4}x_3^2x_4 -\frac{3}{4}x_3^2x_5 +\frac{3}{4}x_3x_4^2 +\frac{3}{2}x_3x_4x_5 +\frac{3}{4}x_3x_5^2 -\frac{1}{2}x_4^3 -\frac{3}{2}x_4^2x_5 -\frac{3}{2}x_4x_5^2 -\frac{1}{2}x_5^3 -\frac{13}{24}x_1^2 +\frac{7}{6}x_1x_4 +\frac{7}{6}x_1x_5 -\frac{11}{24}x_2^2 +\frac{11}{12}x_2x_4 +\frac{11}{12}x_2x_5 -\frac{11}{24}x_3^2 +\frac{11}{12}x_3x_4 +\frac{11}{12}x_3x_5 -\frac{11}{12}x_4^2 -\frac{5}{6}x_4x_5 -\frac{11}{12}x_5^2 +\frac{3}{4}x_1 +\frac{1}{4}x_2 +\frac{1}{4}x_3 +\frac{1}{2}x_4 +\frac{1}{2}x_5 +1$

chamber 4: $-\frac{1}{12}x_1^4 -\frac{1}{3}x_1^3x_2 +\frac{1}{6}x_1^3x_4 +\frac{1}{6}x_1^3x_5 -\frac{1}{4}x_1^2x_2^2 +\frac{1}{2}x_1^2x_2x_4 +\frac{1}{2}x_1^2x_2x_5 -\frac{1}{4}x_1^2x_4^2 -\frac{1}{4}x_1^2x_5^2 -\frac{1}{24}x_3^4 +\frac{1}{6}x_3^3x_4 +\frac{1}{6}x_3^3x_5 -\frac{1}{4}x_3^2x_4^2 -\frac{1}{2}x_3^2x_4x_5 -\frac{1}{4}x_3^2x_5^2 +\frac{1}{6}x_3x_4^3 +\frac{1}{2}x_3x_4^2x_5 +\frac{1}{2}x_3x_4x_5^2 +\frac{1}{6}x_3x_5^3 -\frac{1}{24}x_4^4 -\frac{1}{6}x_4^3x_5 -\frac{1}{4}x_4^2x_5^2 -\frac{1}{6}x_4x_5^3 -\frac{1}{24}x_5^4 -\frac{1}{2}x_1^3 -\frac{3}{4}x_1^2x_2 +\frac{3}{4}x_1^2x_4 +\frac{3}{4}x_1^2x_5 -\frac{3}{4}x_1x_2^2 +\frac{3}{2}x_1x_2x_4 +\frac{3}{2}x_1x_2x_5 -\frac{3}{4}x_1x_4^2 -\frac{3}{4}x_1x_5^2 +\frac{1}{4}x_3^3 -\frac{3}{4}x_3^2x_4 -\frac{3}{4}x_3^2x_5 +\frac{3}{4}x_3x_4^2 +\frac{3}{2}x_3x_4x_5 +\frac{3}{4}x_3x_5^2 -\frac{1}{4}x_4^3 -\frac{3}{4}x_4^2x_5 -\frac{3}{4}x_4x_5^2 -\frac{1}{4}x_5^3 -\frac{5}{12}x_1^2 +\frac{1}{12}x_1x_2 +\frac{13}{12}x_1x_4 +\frac{13}{12}x_1x_5 -\frac{1}{2}x_2^2 +x_2x_4 +x_2x_5 -\frac{11}{24}x_3^2 +\frac{11}{12}x_3x_4 +\frac{11}{12}x_3x_5 -\frac{23}{24}x_4^2 -\frac{11}{12}x_4x_5 -\frac{23}{24}x_5^2 +x_1 +\frac{1}{2}x_2 +\frac{1}{4}x_3 +\frac{1}{4}x_4 +\frac{1}{4}x_5 +1$

chamber 5: $-\frac{5}{24}x_1^4 -\frac{1}{3}x_1^3x_2 -\frac{1}{3}x_1^3x_3 +\frac{1}{2}x_1^3x_4 +\frac{1}{2}x_1^3x_5 -\frac{1}{4}x_1^2x_2^2 +\frac{1}{2}x_1^2x_2x_4 +\frac{1}{2}x_1^2x_2x_5 -\frac{1}{4}x_1^2x_3^2 +\frac{1}{2}x_1^2x_3x_4 +\frac{1}{2}x_1^2x_3x_5 -\frac{1}{2}x_1^2x_4^2 -\frac{1}{2}x_1^2x_4x_5 -\frac{1}{2}x_1^2x_5^2 -\frac{3}{4}x_1^3 -\frac{3}{4}x_1^2x_2 -\frac{3}{4}x_1^2x_3 +\frac{3}{2}x_1^2x_4 +\frac{3}{2}x_1^2x_5 -\frac{3}{4}x_1x_2^2 +\frac{3}{2}x_1x_2x_4 +\frac{3}{2}x_1x_2x_5 -\frac{3}{4}x_1x_3^2 +\frac{3}{2}x_1x_3x_4 +\frac{3}{2}x_1x_3x_5 -\frac{3}{2}x_1x_4^2 -\frac{3}{2}x_1x_4x_5 -\frac{3}{2}x_1x_5^2 -\frac{7}{24}x_1^2 +\frac{1}{12}x_1x_2 +\frac{1}{12}x_1x_3 +x_1x_4 +x_1x_5 -\frac{1}{2}x_2^2 +x_2x_4 +x_2x_5 -\frac{1}{2}x_3^2 +x_3x_4 +x_3x_5 -x_4^2 -x_4x_5 -x_5^2 +\frac{5}{4}x_1 +\frac{1}{2}x_2 +\frac{1}{2}x_3 +1$

chamber 6: $-\frac{1}{12}x_1^4 -\frac{1}{3}x_1^3x_2 +\frac{1}{6}x_1^3x_4 +\frac{1}{6}x_1^3x_5 -\frac{1}{4}x_1^2x_2^2 +\frac{1}{2}x_1^2x_2x_4 +\frac{1}{2}x_1^2x_2x_5 -\frac{1}{4}x_1^2x_4^2 -\frac{1}{4}x_1^2x_5^2 -\frac{1}{6}x_1x_2^3 +\frac{1}{2}x_1x_2^2x_5 -\frac{1}{2}x_1x_2x_5^2 +\frac{1}{6}x_1x_5^3 -\frac{1}{24}x_2^4 +\frac{1}{6}x_2^3x_5 -\frac{1}{4}x_2^2x_5^2 +\frac{1}{6}x_2x_5^3 -\frac{1}{24}x_5^4 -\frac{1}{2}x_1^3 -\frac{3}{4}x_1^2x_2 +\frac{3}{4}x_1^2x_4 +\frac{3}{4}x_1^2x_5 -\frac{3}{4}x_1x_2^2 +\frac{3}{2}x_1x_2x_4 +\frac{3}{2}x_1x_2x_5 -\frac{3}{4}x_1x_4^2 -\frac{3}{4}x_1x_5^2 -\frac{1}{4}x_2^3 +\frac{3}{4}x_2^2x_5 -\frac{3}{4}x_2x_5^2 +\frac{1}{4}x_5^3 -\frac{5}{12}x_1^2 +\frac{1}{4}x_1x_2 +\frac{13}{12}x_1x_4 +\frac{11}{12}x_1x_5 -\frac{11}{24}x_2^2 +x_2x_4 +\frac{11}{12}x_2x_5 -\frac{1}{2}x_4^2 -\frac{11}{24}x_5^2 +x_1 +\frac{3}{4}x_2 +\frac{1}{2}x_4 +\frac{1}{4}x_5 +1$

chamber 7: $-\frac{1}{12}x_1^4 - \frac{1}{3}x_1^3x_2 + \frac{1}{6}x_1^3x_4 + \frac{1}{6}x_1^3x_5 - \frac{1}{4}x_1^2x_2^2 + \frac{1}{2}x_1^2x_2x_4 + \frac{1}{2}x_1^2x_2x_5 - \frac{1}{4}x_1^2x_4^2 - \frac{1}{4}x_1^2x_5^2 - \frac{1}{3}x_1x_2^3 + \frac{1}{2}x_1x_2^2x_4 + \frac{1}{2}x_1x_2^2x_5 - \frac{1}{2}x_1x_2x_4^2 - \frac{1}{2}x_1x_2x_5^2 + \frac{1}{6}x_1x_4^3 + \frac{1}{6}x_1x_5^3 - \frac{1}{12}x_2^4 + \frac{1}{6}x_2^3x_4 + \frac{1}{6}x_2^3x_5 - \frac{1}{4}x_2^2x_4^2 - \frac{1}{4}x_2^2x_5^2 + \frac{1}{6}x_2x_4^3 + \frac{1}{6}x_2x_5^3 - \frac{1}{24}x_4^4 - \frac{1}{24}x_5^4 - \frac{1}{2}x_1^3 - \frac{3}{4}x_1^2x_2 + \frac{3}{4}x_1^2x_4 + \frac{3}{4}x_1^2x_5 - \frac{3}{4}x_1x_2^2 + \frac{3}{2}x_1x_2x_4 + \frac{3}{2}x_1x_2x_5 - \frac{3}{4}x_1x_4^2 - \frac{3}{4}x_1x_5^2 - \frac{1}{2}x_2^3 + \frac{3}{4}x_2^2x_4 + \frac{3}{4}x_2^2x_5 - \frac{3}{4}x_2x_4^2 - \frac{3}{4}x_2x_5^2 + \frac{1}{4}x_4^3 + \frac{1}{4}x_5^3 - \frac{5}{12}x_1^2 + \frac{5}{12}x_1x_2 + \frac{11}{12}x_1x_4 + \frac{11}{12}x_1x_5 - \frac{5}{12}x_2^2 + \frac{11}{12}x_2x_4 + \frac{11}{12}x_2x_5 - \frac{11}{24}x_4^2 - \frac{11}{24}x_5^2 + x_1 + x_2 + \frac{1}{4}x_4 + \frac{1}{4}x_5 + 1$

chamber 8: $-\frac{1}{12}x_1^4 - \frac{1}{3}x_1^3x_2 + \frac{1}{6}x_1^3x_4 + \frac{1}{6}x_1^3x_5 - \frac{1}{4}x_1^2x_2^2 + \frac{1}{2}x_1^2x_2x_4 + \frac{1}{2}x_1^2x_2x_5 - \frac{1}{4}x_1^2x_4^2 - \frac{1}{4}x_1^2x_5^2 - \frac{1}{6}x_1x_2^3 + \frac{1}{2}x_1x_2^2x_5 - \frac{1}{2}x_1x_2x_5^2 + \frac{1}{6}x_1x_5^3 - \frac{1}{24}x_2^4 + \frac{1}{6}x_2^3x_5 - \frac{1}{4}x_2^2x_5^2 + \frac{1}{6}x_2x_5^3 - \frac{1}{24}x_3^4 + \frac{1}{6}x_3^3x_4 + \frac{1}{6}x_3^3x_5 - \frac{1}{4}x_3^2x_4^2 - \frac{1}{2}x_3^2x_4x_5 - \frac{1}{4}x_3^2x_5^2 + \frac{1}{6}x_3x_4^3 + \frac{1}{2}x_3x_4^2x_5 + \frac{1}{2}x_3x_4x_5^2 + \frac{1}{6}x_3x_5^3 - \frac{1}{24}x_4^4 - \frac{1}{6}x_4^3x_5 - \frac{1}{4}x_4^2x_5^2 - \frac{1}{6}x_4x_5^3 - \frac{1}{12}x_5^4 - \frac{1}{2}x_1^3 - \frac{3}{4}x_1^2x_2 + \frac{3}{4}x_1^2x_4 + \frac{3}{4}x_1^2x_5 - \frac{3}{4}x_1x_2^2 + \frac{3}{2}x_1x_2x_4 + \frac{3}{2}x_1x_2x_5 - \frac{3}{4}x_1x_4^2 - \frac{3}{4}x_1x_5^2 - \frac{1}{4}x_2^3 + \frac{3}{4}x_2^2x_5 - \frac{3}{4}x_2x_5^2 + \frac{1}{4}x_3^3 - \frac{3}{4}x_3^2x_4 - \frac{3}{4}x_3^2x_5 + \frac{3}{4}x_3x_4^2 + \frac{3}{2}x_3x_4x_5 + \frac{3}{4}x_3x_5^2 - \frac{1}{4}x_4^3 - \frac{3}{4}x_4^2x_5 - \frac{3}{4}x_4x_5^2 - \frac{5}{12}x_1^2 + \frac{1}{4}x_1x_2 + \frac{13}{12}x_1x_4 + \frac{11}{12}x_1x_5 - \frac{11}{24}x_2^2 + x_2x_4 + \frac{11}{12}x_2x_5 - \frac{11}{24}x_3^2 + \frac{11}{12}x_3x_4 + \frac{11}{12}x_3x_5 - \frac{23}{24}x_4^2 - \frac{11}{12}x_4x_5 - \frac{11}{12}x_5^2 + x_1 + \frac{3}{4}x_2 + \frac{1}{4}x_3 + \frac{1}{4}x_4 + 1$

chamber 9: $-\frac{1}{12}x_1^4 - \frac{1}{3}x_1^3x_2 + \frac{1}{6}x_1^3x_4 + \frac{1}{6}x_1^3x_5 - \frac{1}{4}x_1^2x_2^2 + \frac{1}{2}x_1^2x_2x_4 + \frac{1}{2}x_1^2x_2x_5 - \frac{1}{4}x_1^2x_4^2 - \frac{1}{4}x_1^2x_5^2 - \frac{1}{3}x_1x_2^3 + \frac{1}{2}x_1x_2^2x_4 + \frac{1}{2}x_1x_2^2x_5 - \frac{1}{2}x_1x_2x_4^2 - \frac{1}{2}x_1x_2x_5^2 + \frac{1}{6}x_1x_4^3 + \frac{1}{6}x_1x_5^3 - \frac{1}{12}x_2^4 + \frac{1}{6}x_2^3x_4 + \frac{1}{6}x_2^3x_5 - \frac{1}{4}x_2^2x_4^2 - \frac{1}{4}x_2^2x_5^2 + \frac{1}{6}x_2x_4^3 + \frac{1}{6}x_2x_5^3 - \frac{1}{24}x_3^4 + \frac{1}{6}x_3^3x_4 + \frac{1}{6}x_3^3x_5 - \frac{1}{4}x_3^2x_4^2 - \frac{1}{2}x_3^2x_4x_5 - \frac{1}{4}x_3^2x_5^2 + \frac{1}{6}x_3x_4^3 + \frac{1}{2}x_3x_4^2x_5 + \frac{1}{2}x_3x_4x_5^2 + \frac{1}{6}x_3x_5^3 - \frac{1}{12}x_4^4 - \frac{1}{6}x_4^3x_5 - \frac{1}{4}x_4^2x_5^2 - \frac{1}{6}x_4x_5^3 - \frac{1}{12}x_5^4 - \frac{1}{2}x_1^3 - \frac{3}{4}x_1^2x_2 + \frac{3}{4}x_1^2x_4 + \frac{3}{4}x_1^2x_5 - \frac{3}{4}x_1x_2^2 + \frac{3}{2}x_1x_2x_4 + \frac{3}{2}x_1x_2x_5 - \frac{3}{4}x_1x_4^2 - \frac{3}{4}x_1x_5^2 - \frac{1}{2}x_2^3 + \frac{3}{4}x_2^2x_4 + \frac{3}{4}x_2^2x_5 - \frac{3}{4}x_2x_4^2 - \frac{3}{4}x_2x_5^2 + \frac{1}{4}x_3^3 - \frac{3}{4}x_3^2x_4 - \frac{3}{4}x_3^2x_5 + \frac{3}{4}x_3x_4^2 + \frac{3}{2}x_3x_4x_5 + \frac{3}{4}x_3x_5^2 - \frac{3}{4}x_4^2x_5 - \frac{3}{4}x_4x_5^2 - \frac{5}{12}x_1^2 + \frac{5}{12}x_1x_2 + \frac{11}{12}x_1x_4 + \frac{11}{12}x_1x_5 - \frac{5}{12}x_2^2 + \frac{11}{12}x_2x_4 + \frac{11}{12}x_2x_5 - \frac{11}{24}x_3^2 + \frac{11}{12}x_3x_4 + \frac{11}{12}x_3x_5 - \frac{11}{12}x_4^2 - \frac{11}{12}x_4x_5 - \frac{11}{12}x_5^2 + x_1 + x_2 + \frac{1}{4}x_3 + 1$

chamber 10: $-\frac{1}{24}x_1^4 - \frac{1}{6}x_1^3x_2 + \frac{1}{6}x_1^3x_4 + \frac{1}{2}x_1^2x_2x_4 - \frac{1}{4}x_1^2x_4^2 - \frac{1}{4}x_1^3 + \frac{3}{4}x_1^2x_4 + \frac{3}{2}x_1x_2x_4 - \frac{3}{4}x_1x_4^2 + \frac{1}{24}x_1^2 + \frac{7}{6}x_1x_2 + \frac{13}{12}x_1x_4 + x_2x_4 - \frac{1}{2}x_4^2 + \frac{5}{4}x_1 + x_2 + \frac{1}{2}x_4 + 1$

chamber 11: $-\frac{1}{24}x_1^4 - \frac{1}{6}x_1^3x_2 + \frac{1}{6}x_1^3x_4 + \frac{1}{2}x_1^2x_2x_4 - \frac{1}{4}x_1^2x_4^2 - \frac{1}{6}x_1x_2^3 + \frac{1}{2}x_1x_2^2x_4 - \frac{1}{2}x_1x_2x_4^2 + \frac{1}{6}x_1x_4^3 - \frac{1}{24}x_2^4 + \frac{1}{6}x_2^3x_4 - \frac{1}{4}x_2^2x_4^2 + \frac{1}{6}x_2x_4^3 - \frac{1}{24}x_4^4 - \frac{1}{4}x_1^3 + \frac{3}{4}x_1^2x_4 + \frac{3}{2}x_1x_2x_4 - \frac{3}{4}x_1x_4^2 - \frac{1}{4}x_2^3 + \frac{3}{4}x_2^2x_4 - \frac{3}{4}x_2x_4^2 + \frac{1}{4}x_4^3 + \frac{1}{24}x_1^2 + \frac{4}{3}x_1x_2 + \frac{11}{12}x_1x_4 + \frac{1}{24}x_2^2 + \frac{11}{12}x_2x_4 - \frac{11}{24}x_4^2 + \frac{5}{4}x_1 + \frac{5}{4}x_2 + \frac{1}{4}x_4 + 1$

chamber 12: $\frac{1}{4}x_1^2x_2^2 + \frac{3}{4}x_1^2x_2 + \frac{3}{4}x_1x_2^2 + \frac{1}{2}x_1^2 + \frac{9}{4}x_1x_2 + \frac{1}{2}x_2^2 + \frac{3}{2}x_1 + \frac{3}{2}x_2 + 1$