Derandomizing Semidefinite Programming Based Approximation Algorithms

Sanjeev Mahajan H. Ramesh*
Max Planck Institut für Informatik
Saarbrücken, Germany, 66123

Abstract

Remarkable breakthroughs have been made recently in obtaining approximate solutions to some fundamental NP-Complete problems, namely Max-Cut, Max k-Cut, Max-Sat, Max-Dicut, Max-Bisection, k Vertex Coloring, Independent Set, etc. These breakthroughs all involve polynomial time randomized *algorithms based upon* Semidefinite Programming*, a technique pioneered by Goemans and Williamson[8].*

In this paper, we give techniques to derandomize the above class of randomized algorithms, thus obtaining polynomial time deterministic algorithms with the same approximation ratios for the above problems. At the heart of our technique is the use of spherical symmetry to convert a nested sequence of n integrations, which cannot be approximated sufficiently well in polynomial time, to a nested sequence of just a constant number of integrations, which can be approximated sufficiently well in polynomial time.

1 Introduction

The application of *Semidefinite Programming* to obtaining approximation algorithms for NP-Complete problems was pioneered by Goemans and Williamson[8]. This technique involves relaxing an integer program (solving which is an NP-Complete problem) to a semidefinite program (which can be solved with a sufficiently small error in polynomial time). In a remarkable breakthrough, Goemans and Williamson showed how this technique could be used to give a *randomized* approximation algorithm for the Max-Cut problem with an approximation ratio of .878. This must be contrasted with the previously best known approximation ratio of .5 obtained by the simple random cut algorithm. Subsequently, semidefinite programming based techniques have led to randomized algorithms with substantially better approximation ratios for a number of fundamental problems.

Goemans and Williamson[8] obtained a .878 approximation algorithm for Max-2Sat and an .758 approximation algorithm for Max-Sat, improving upon the previously best known bound of 3/4 [18] for both. They also obtain a .796 approximation algorithm for Max-Dicut, improving upon the previously best known ratio of .25 given by the random cut algorithm. Feige and Goemans [5] obtained improved approximation algorithms for Max-2Sat and Max-Dicut.

*Present Address: Indian Institute of Science, Bangalore, India

Karger, Motwani and Sudan obtained an algorithm for coloring any k-colorable graph with $O(n^{1-3/(k+1)}\log n)$ colors[11]; in particular, for 3-colorable graphs, this algorithm requires $O(n^{.25}\log n)$ colors. This improves upon the deterministic algorithm of Blum[2] which requires $O(n^{1-\frac{1}{k-4/3}}\log^{\frac{8}{5}} n)$ colors for k-colorable graphs.

Frieze and Jerrum[6] obtained a .65 approximation algorithm for Max-Bisection improving the previous best known bound of .5 given by the random bisection algorithm. They also obtained a $1-\frac{1}{k}+2\frac{\ln k}{k^2}$ approximation algorithm for the Max k-Cut problem, improving upon the previously best known ratio of $1-\frac{1}{k}$ given by a random k-Cut.

Alon and Kahale[1] obtained an approximation algorithm for the independent set problem. For any constant $k \geq 3$, if the given graph has an independent set of size $n/k+m$, where n is the number of vertices, they obtain an $\Omega(m^{\frac{3}{k+1}}\log m)$ sized independent set, improving the previously known bound of $\Omega(m^{\frac{1}{k-1}})$ due to Boppana and Halldorsson[3].

All the new developments mentioned above are *randomized* algorithms. All of them share the following common paradigm. First, a semidefinite program is solved to obtain a collection of n vectors in n dimensional space satisfying some properties dependent upon the particular problem in question. This step is deterministic (in the Feige, Goemans paper [5], there is another intermediate step of generating a new set of vectors from the vectors obtained above). Second, a set of independent random vectors is generated, each vector being *spherically symmetric*, i.e., equally likely to pass through any point on the n dimensional unit sphere centered at the origin. Finally, the solution is obtained using some computation on the n given vectors and the random vectors.

It is not obvious how to derandomize the above randomized algorithms, i.e., to obtain a "good" set of random vectors deterministically. A natural way to derandomize is to use the method of Conditional Probabilities[14, 16]. The problem that occurs then is to compute the conditional probabilities in polynomial time.

• The main contribution of this paper is a technique which enables derandomization of all the semidefinite programming based approximation algorithms listed above. This leads to deterministic approximation algorithms for Max-Cut, Max k-Cut, Max Bisection, Max-2Sat, Max-Sat, Max-Dicut, k Vertex Coloring, and Independent Set with the same approximation ratios as their randomized counterparts mentioned above.

Our derandomization uses the conditional probability technique. We compute conditional probabilities as follows. First, we show how to express each conditional probability computation as a sequence of $O(n)$ nested integrals. Performing this sequence of integrations with a small enough error seems hard to do in polynomial time. The key observation which facilitates conditional probability computation in polynomial ´time is that, using spherical symmetry properties, the above sequence of $O(n)$ nested integrals can be reduced to evaluating an expression with just a constant number of nested integrals for each of the semidefinite based approximation algorithms mentioned above. This new sequence of integrations can be performed with a small enough error in polynomial time. A host of precision issues also crops up in the derandomization. Conditional probabilities must be computed only at a polynomial number of points. Further each conditional probability computation must be performed within a small error. We show how to handle these precision issues in polynomial time.

As mentioned above, our derandomization techniques apply to all the semidefinite programming based approximation algorithms mentioned above. Loosely speaking, we believe our techniques are even more general, i.e., applicable to any scheme which follows the above paradigm and in which the critical performance analysis boils down to an "elementary event" involving just a constant number of the n vectors at a time. For example, in the graph coloring algorithm, only two vectors, corresponding to the endpoints of some edge, need to be considered at a time. An example of an elementary event involving 3 vectors is the Max-Dicut algorithm of Goemans and Williamson. Another example of the same is the algorithm of Kelsen, Mahajan, and Ramesh [12] for coloring 2-colorable 3-uniform hypergraphs approximately.

The paper is organized as follows. In Section 2, we outline the Goemans and Williamson Max-Cut algorithm and the Karger, Motwani, Sudan coloring algorithm. We then describe our derandomization scheme. Since the Karger, Motwani, Sudan coloring algorithm appears to be the hardest to derandomize amongst the algorithms mentioned above, our exposition concentrates on this algorithm. The derandomization of the other algorithms is similar. Section 3 describes the derandomization procedure. The following sections describe the derandomization procedure in detail.

2 The Semidefinite Programming Paradigm

It is known that any concave polynomial time computable function can be maximized (within some tolerance) over a convex set with a weak separation oracle in polynomial time [7]. One such convex set is the set of *semidefinite matrices*, i.e., those matrices whose eigenvalues are all non-negative. A set formed by the intersections of half-spaces and the set of semidefinite matrices is also a convex set. Further, this convex set admits a weak separation oracle. A semidefinite program involves maximizing a polynomial time computable concave function over one such convex set. Semidefinite programs are therefore solvable (up to an arbitrarily small additive error) in polynomial time. Goemans and Williamson first used this fact to obtain an approximation algorithm for Max-Cut.

The Goemans-Williamson Max-Cut Algorithm. Goemans and Williamson took a natural integer program for Max-Cut and showed how to relax it to a semidefinite program. The solution to this program is a set of n unit vectors, one corresponding to each vertex of the graph in question. These vectors emanate from the origin. We call these vectors *vertex vectors*. These are embedded in n dimensional space. This leads to the question as to how a large cut is obtained from these vectors.

Goemans and Williamson choose a random hyperplane through the origin whose normal is spherically symmetrically distributed; this hyperplane divides the vertex vectors into 2 groups, which define a cut in the obvious manner. The expected number $E(W)$ of edges[1] across the cut is $\sum_{(v,w)\in E} \arccos(v \cdot w)/\pi = \sum_{(v,w)\in E} Pr(sign(v \cdot R) \neq sign(w \cdot R))$, where E is the set of edges in the graph and v, w denote both vertices in the graph and the associated vertex vectors. Goemans and Williamson show that $E(W)$ is at least .878 times the maximum cut.

Note that the n random variables involved above are the n coordinates which define the normal R to the random hyperplane. Let $R_1, R_2, \ldots, R_n$ be these random variables. For R to be spherically symmetrically distributed, it suffices that the R_i's are independent and identi-

[1] For simplicity, we consider the unweighted Max-Cut problem.

cally distributed with a mean 0 and variance 1 normal distribution, i.e., the density function is $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ [4]. Derandomizing the above algorithm thus requires obtaining values for $R_1, \ldots, R_n$ deterministically so that the value of the cut given by the corresponding hyperplane is at least $E(W)$.

The Karger-Motwani-Sudan Coloring Algorithm. The Karger, Motwani, Sudan algorithm shows how to color a 3-colorable graph of n vertices with $O(n^{1/4}\log n)$ colors. The authors use a semidefinite program to obtain a set of vertex vectors such that $v \cdot w \leq -\frac{1}{2}$, for all edges (v, w). Note that if these vectors are somehow constrained to be in 2 dimensions, then there are at most 3 distinct vectors, which would specify a 3-coloring. However, the output to the semidefinite program are vectors in an n-dimensional space. It remains to be described how a coloring is obtained from these vectors. This is done as follows.

Karger, Motwani and Sudan choose r vectors, $t_1, \ldots, t_r$, independently and at random; each is spherically symmetric. These vectors are called *centers*. Let the jth coordinate of t_i be denoted by $t_i[j]$, $1 \leq j \leq n$. Spherical symmetry is obtained by the following procedure: each $t_i[j]$ is chosen independently at random from a normal distribution with mean 0 and variance 1. The color that vertex v gets is simply c, where $t_c \cdot v = \max_{1 \leq i \leq r} t_i \cdot v$. In other words, the color assigned to a vertex v corresponds to that amongst the r centers which has the largest projection on the vector v.

To determine how good the above procedure is, it is necessary to determine the probability that an edge is *bad*, i.e., both its endpoints get the same color. Consider two vertex vectors v, w, such that (v, w) is an edge e in $G = (V, E)$. The probability that v and w get the same color in the algorithm is given by $Pr(E^e) = \sum_{k=1}^{r} Pr(E^e_k)$, where E^e_k is the event that both get color t_k. E^e_k can be written as:

E^e_k: $t_k \cdot v = \max\{t_1 \cdot v, \ldots, t_r \cdot v\} \wedge t_k \cdot w = \max\{t_1 \cdot w, \ldots, t_r \cdot w\}$

Karger, Motwani and Sudan[11] show that $\sum_{e \in E} Pr(E^e) < n/4$, for $r = d^{1/3}\log^{4/3} d$, where d is the maximum degree of the graph. Thus, at the end of the above procedure, the expected number of bad edges is less than $n/4$. All vertices except those upon which these bad edges are incident are discarded (the colors assigned to them are final). The remaining vertices, which are at most $n/2$ in number, are recolored by repeating the above procedure $O(\log n)$ times, using a fresh set of colors each time. This gives an $O(d^{1/3}\log^{4/3} d \log n)$ coloring of a 3-colorable graph. This combined with Wigderson's trick [17] gives an $O(n^{1/4}\log n)$ coloring of a 3-colorable graph.

Derandomizing the above algorithm entails deterministically obtaining values for $t_i[j]$'s so that the number of bad edges is at most the expected number of bad edges above, i.e., $n/4$. Actually it suffices to obtain values for $t_i[j]$'s such that the number of bad edges is at most $n/4 + \tau$, for some constant τ. This is what we will do.

Note that the Goemans-Williamson algorithm uses a random hyperplane while the Karger-Motwani-Sudan algorithm uses a set of random centers. Although these 2 methods seem different, the hyperplane method can be interpreted as just the center method with 2 centers.

3 The Derandomization Scheme

For simplicity, we restrict our exposition here to the derandomization of the Karger, Motwani, Sudan algorithm for coloring 3-colorable graphs. Our procedure easily generalizes to all other known semidefinite programming based approximation algorithms listed in Section 1.

Notation. For a vector u, we denote by $u[l \dots m]$ the vector formed by the lth to mth coordinates of u.

The Derandomization Scheme. The scheme is essentially to use the method of conditional expectations to deterministically find values for the vectors $t_1, \dots, t_r$ so that the number of bad edges is just $n/4 + \tau$, for some constant τ.

We define a total order $<$ on the conditional variables as follows: $t_1[1] < \dots < t_1[n] < t_2[1] \dots < t_2[n] < \dots < t_r[1] < \dots < t_r[n]$. The values of these are fixed one by one, in order. So suppose that the values $t_1[1 \dots n], t_2[1 \dots n], \dots, t_i[1 \dots j-1]$ have been determined. We will show how a value for $t_i[j]$ is determined.

Notations. Let $\mathcal{E}$ be an event. Then $Pr(\mathcal{E}|i, j, \delta)$ denotes the probability that the event $\mathcal{E}$ occurs when the values for all conditional variables $< t_i[j]$ have been fixed as above and $t_i[j]$ itself is assigned value δ. So, for example, $Pr(E_k^e|i, j, \delta)$ denotes the probability that event E_k^e occurs (i.e., that both endpoints of edge e get the color associated with center t_k), when the values for all conditional variables $< t_i[j]$ have been fixed as above and $t_i[j]$ itself is assigned value δ. For notational brevity, sometimes we use $f(\delta)$ to denote $Pr(E_e^k|i, j, \delta)$.

Let $p(\delta)$ be the expected number of bad edges, when the values for all conditional variables before $t_i[j]$ are fixed as above and $t_i[j]$ is assigned value δ; $p(\delta) = \sum_{e \in E} \sum_{k=1}^{r} Pr(E_k^e|i, j, \delta)$.

Fixing $t_i[j]$. Let δ_{min} be the value of δ which minimizes $p(\delta)$, $-\infty \leq \delta \leq \infty$. We would like to fix $t_i[j]$ to δ_{min}. The question is how to compute δ_{min}. We do not actually compute δ_{min}. However, we will show the following.

Theorem 3.1 *A value κ for $t_i[j]$ satisfying the following property can be computed in polynomial time: $p(\delta_{min}) \leq p(\kappa) \leq p(\delta_{min}) + O(1/n^2)$.*

From the above theorem, we derive the following corollary.

Corollary 3.2 *After all $t_i[j]$'s have been fixed and colors assigned to vertices as in the randomized algorithm, the number of bad edges is at most $\frac{n}{4} + O(1)$.*

Proof. Note that the number of conditional variables $t_i[j]$ is $nr \leq n^2$ (actually for 3-colorable graphs r is much smaller, namely $d^{1/3} \log^{4/3} d$, where d is the maximum degree).

Recall that the expected number of bad edges before any of the random variables was fixed is at most $\frac{n}{4}$. By Theorem 3.1, the expected number of bad edges after the first conditional variable is fixed is at most $\frac{n}{4} + O(\frac{1}{n^2})$. An easy inductive argument shows that the expected number of bad edges after the lth conditional variable is fixed is at most $\frac{n}{4} + O(\frac{l}{n^2})$. After all the $nr \leq n^2$ conditional variables have been fixed, the expected number of bad edges (which is just the number of bad edges since all conditional variables are now fixed) is at most $\frac{n}{4} + O(1)$. □

To show Theorem 3.1, we will perform the following steps.

Step 1. First we discretize the vertex vectors. This ensures that derivatives of the functions $f(\delta)$ and $p(\delta)$ are bounded by a polynomial in n. This, in turn, ensures that the values of the above functions between any two nearby points will not be too different from their values at these two points. This facilitates discrete evaluation. This discretization is described in Section 4.

Step 2. In Section 5, we show how to express $f(\delta)$ in terms of a function I defined as follows. This enables the probability to be computed using integrals with just constant nesting depth.

Definition. Let b, b' be vectors of the same dimension, which is at least 2. Let a be another vector of the same dimension whose entries are independent and normally distributed with mean 0 and variance 1. Let $x \leq y$ and $x' \leq y'$ be in the range $-\infty \ldots \infty$. Then $I(b, b', x, y, x', y')$ denotes $Pr((x \leq a \cdot b \leq y) \wedge (x' \leq a \cdot b' \leq y'))$.

Step 3. Computing the integral corresponding to I is the key question. As mentioned in the introduction, naive computation would require a nested sequence of $\Omega(n)$ integrals. Using spherical symmetry properties we show how to perform this integration using just integrals with constant nesting. This is described in Section 6.

Step 4. In order to compute κ, we can afford to evaluate $p(\delta)$ only for a polynomial number of points. So we have to discretize the range $-\infty \ldots \infty$ for δ. We must do so in a way such that the least value of $p(\delta)$ when restricted to just these discrete points is $O(\frac{1}{n^2})$ away from the actual minimum. This is done in Section 7.

Step 5. Finally we need to show how $f(\delta)$ can be approximated within an additive $O(\frac{1}{n^5})$ error in polynomial time; this will ensure that $p(\delta)$ can be evaluated within an $O(r|E|\frac{1}{n^5}) = O(\frac{1}{n^2})$ error, as required by Theorem 3.1. To do so, we need to show that the integrations in Section 5 which describe $f(\delta)$ can be evaluated within an additive $O(\frac{1}{n^5})$ error. This is again done by discretizing the range between the limits of the integration and is described in Section 8.

4 Discretizing the Input Vectors

For simplicity, we assume that $v \cdot w = -1/2$ for edge $e = (v, w)$. The Karger-Motwani-Sudan algorithm sets $v \cdot w \leq -1/2$, but the theoretical performance of the algorithm does not change if $v \cdot w$ is set to exactly $-1/2$. Our algorithm can be easily generalized as long as $|v \cdot w|$ is bounded away from 1. Note that the above assumption is not valid for the Goemans-Williamson Max-Cut algorithm. We show how to tackle this problem in Section 9.

Let ϵ be a number which is $\Theta(\frac{1}{n^3})$. The objective of discretizing the input vectors is the following. All the components of v, w must be made $\Omega(\epsilon)$ in absolute value. In addition, for each h, $1 \leq h < n$, when the coordinate system is rotated so as to make $v[h \ldots n] = (b_1, 0, \ldots, 0)$ and $w[h \ldots n] = (b'_1, b'_2, 0, \ldots, 0)$, $|b_1|$ and $|b'_2|$ are at least inverse polynomial in n. The last property will be necessary in our proofs later, in Lemma 6.2, in particular. This discretization is achieved as follows.

Rounding Procedure. First, all the entries in v, w are rounded upwards (in absolute value) to the nearest non-zero multiples of ϵ. Next, up to 2ϵ is added (in absolute value) to v_{n-1} so that $|v_{n-1}w_n - w_{n-1}v_n| > 2\epsilon^2$. Note that in this process $1 \leq |v|^2, |w|^2 \leq 2$, for small enough ϵ, i.e., for large enough n. Each component of v, w changes by at most 3ϵ in this process and is now non-zero; further, the absolute values of the components cannot decrease. Finally, we divide each vector by its new norm so as to make it a unit vector. Then $|v_{n-1}w_n - w_{n-1}v_n| = \Omega(\epsilon^2)$.

Lemma 4.1 *For each h, $1 \le h < n$, $|w'|^2 - \frac{(v'\cdot w')^2}{|v'|^2} = \Omega(\epsilon^4)$, where $v' = v[h \dots n]$ and $w' = w[h \dots n]$.*

Proof. First consider $h = n-1$. $(v' \cdot w')^2 = (v_{n-1}w_{n-1} + v_n w_n)^2 = (v_{n-1}^2 + v_n^2)(w_{n-1}^2 + w_n^2) - (v_{n-1}w_n - w_{n-1}v_n)^2 \le |v'|^2|w'|^2 - \Omega(\epsilon^4)$. Therefore, $|w'|^2 - \frac{(v'\cdot w')^2}{|v'|^2} = \Omega(\epsilon^4)$.

Next consider $h < n-1$. Let $l = v[h, \dots, n-2]$ and $m = w[h, \dots, n-2]$. Let $l' = v[n-1 \dots n]$ and $m' = w[n-1 \dots n]$. $(v' \cdot w')^2 = (l \cdot m + l' \cdot m')^2 = (l \cdot m)^2 + (l' \cdot m')^2 + 2(l' \cdot m')(l \cdot m) \le |l|^2|m|^2 + (l' \cdot m')^2 + 2|l'||m'||l||m|$. By the previous paragraph, $(l' \cdot m')^2 \le |l'|^2|m'|^2 - \Omega(\epsilon^4)$. Therefore, $(v' \cdot w')^2 \le |l|^2|m|^2 + |l'|^2|m'|^2 + |l'|^2|m|^2 + |l|^2|m'|^2 - \Omega(\epsilon^4) \le (|l|^2 + |l'|^2)(|m|^2 + |m'|^2) - \Omega(\epsilon^4) = |v'|^2|w'|^2 - \Omega(\epsilon^4)$. Therefore, $||w'|^2 - \frac{(v'\cdot w')^2}{|v'|^2}| = \Omega(\epsilon^4)$. $\square$

The above rounding changes the vectors and hence the value of $p(\delta_{min})$. We redefine δ_{min} so that it now corresponds to the new rounded vectors. Lemma 4.2 shows that the above rounding changes $p(\delta_{min})$ by just $O(1)$.

Lemma 4.2 *$p(\delta_{min})$ changes by at most $O(n\epsilon)|E| = O(1)$ due to the above rounding.*

Proof. Originally $v \cdot w = -1/2$ and v, w were unit vectors. The above rounding changes $v \cdot w$ by some ϵ' which is $O(n\epsilon)$. Using Taylor series and the fact that $Pr(E^e)$ is a function only of $v \cdot w$ and not the individual vectors themselves, $Pr(E^e) = \sum_{i=0}^{\infty} a_i (v \cdot w)^i$, where $\sum_i a_i = 1$, $a_i \ge 0$ (see [6], Lemma 3, for proof). Then, after the rounding, $|\Delta Pr(E^e)| = |\sum_{i=0}^{\infty} a_i(-\frac{1}{2} + \epsilon')^i - (-\frac{1}{2})^i| \le \sum_{i=0}^{\infty} a_i |(\frac{1}{2} - \epsilon')^i - (\frac{1}{2})^i|$

First, suppose $\epsilon' \ge 0$. Then $|\Delta Pr(E^e)| \le \sum_{i=0}^{\infty} a_i((\frac{1}{2})^i - (\frac{1}{2} - \epsilon')^i) \le \sum_{i=0}^{\infty}((\frac{1}{2})^i - (\frac{1}{2} - \epsilon')^i) \le 4\epsilon'/(1 + 2\epsilon')$. Similarly for $\epsilon' < 0$, $|\Delta Pr(E^e)| \le -4\epsilon'/(1 + 2\epsilon')$. $\square$

Note that we are using the fact that $v \cdot w = -1/2$ only in Lemma 4.2. The proof of this lemma will go through as long as $|v \cdot w|$ is bounded away from 1.

5 Computing Conditional Probabilities

We are required to compute $p(\delta)$, i.e., $\sum_{e \in E} \sum_{k=1}^{r} Pr(E_k^e | i, j, \delta)$. For $e = (v, w)$ and some fixed k, we show how to express $f(\delta) = Pr(E_k^e | i, j, \delta)$ in terms of the integrals $I()$ defined earlier. Recall that $Pr(E_k^e | i, j, \delta)$ is the probability that both v, w are assigned the color corresponding to center t_k, when the values for all conditional variables before $t_i[j]$ have been determined and $t_i[j]$ is assigned δ.

Notation. For vectors a, b, let $a \cdot b[l, m]$ denote $a[l \dots m] \cdot b[l \dots m] = \sum_{h=l}^{m} a[h]b[h]$. Let $\alpha' = t_i \cdot v[1 \dots j-1]$ and Let $\beta' = t_i \cdot w[1 \dots j-1]$.

Fact 1. Note that $t_{i+1}, t_{i+2}, \dots, t_r$ are all completely undetermined, independent, mutually and of $t_1, \dots, t_i$, and identically distributed in a spherically symmetric manner in n dimensions. $t_i[j+1 \dots n]$ is also undetermined and is spherically symmetrically distributed in $n-j$ dimensions and is independent of $t_{i+1}, \dots t_r$ and of all the previously fixed components of t_i.

There are 3 cases, depending upon whether $k < i$, $k = i$, or $k > i$. Each case has 3 subcases, depending upon whether $j < n-1$, $j = n-1$ or $j = n$. We have to consider these 3 subcases separately for the following reason. When $j < n-1$, we will express the above probability in terms of the integral $I()$. For $j = n-1$ and $j = n$, we cannot express the above probability in terms of $I()$ (recall that $I()$ was only defined when its argument vectors are at least 2-dimensional). Therefore, in these 2 subcases, we have to express the probability directly. These 2 subcases themselves need to be separated because the derivative of $f(\delta)$ behaves differently in these 2 subcases, and the behaviour is crucial to the analysis, especially the discretization of Section 7.

Note from Section 4 that $v[n], w[n]$ are non-zero. We will need to divide by these quantities at points.

Case 1. $k < i$. In this case, the center t_k has been already determined. Let $t_k \cdot v = \alpha$ and $t_k \cdot w = \beta$. Centers $t_1, \ldots, t_{i-1}$ have also been determined. If one of $t_1 \cdot v, \ldots, t_{i-1} \cdot v$ is greater than α or if one of $t_1 \cdot w, \ldots, t_{i-1} \cdot w$ is greater than β then $Pr(E_k^e|i,j,\delta)$ is 0. Otherwise, it is:

$f(\delta) = Pr(\wedge_{l=i}^{r}(t_l \cdot v \leq \alpha \wedge t_l \cdot w \leq \beta)|i,j,\delta)$

Note that the events $t_l \cdot v \leq \alpha \wedge t_l \cdot w \leq \beta$, $i \leq l \leq r$, are all independent.

Case 1.1. $j < n-1$. By Fact 1:

$$\begin{aligned}
f(\delta) &= Pr(t_i \cdot v \leq \alpha \wedge t_i \cdot w \leq \beta|i,j,\delta) \times Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= Pr(\alpha' + \delta v[j] + t_i \cdot v[j+1\ldots n] \leq \alpha \wedge \beta' + \delta w[j] + t_i \cdot w[j+1\ldots n] \leq \beta)\times \\
&\quad Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= Pr(t_i \cdot v[j+1\ldots n] \leq \alpha - \alpha' - \delta v[j] \wedge t_i \cdot w[j+1\ldots n] \leq \beta - \beta' - \delta w[j])\times \\
&\quad Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= I(v[j+1\ldots n], w[j+1\ldots n], -\infty, \alpha - \alpha' - \delta v[j], -\infty, \beta - \beta' - \delta w[j])\times \\
&\quad I^{r-i}(v, w, -\infty, \alpha, -\infty, \beta)
\end{aligned}$$

Case 1.2. $j = n-1$. We assume that both $v[n]$ and $w[n]$ are positive. The other cases are similar. By Fact 1 and the fact that $t_i[n]$ is normally distributed:

$$\begin{aligned}
f(\delta) &= Pr(t_i \cdot v \leq \alpha \wedge t_i \cdot w \leq \beta|i,n-1,\delta) \times Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= Pr(t_i[n]v[n] \leq \alpha - \alpha' - \delta v[n-1] \wedge t_i[n]w[n] \leq \beta - \beta' - \delta w[n-1])\times \\
&\quad Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= Pr(t_i[n] \leq \min\{\tfrac{\alpha-\alpha'-\delta v[n-1]}{v[n]}, \tfrac{\beta-\beta'-\delta w[n-1]}{w[n]}\}) \times Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= (\tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\min\{\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}, \frac{\beta-\beta'-w[n-1]\delta}{w[n]}\}} \mathrm{e}^{-z^2/2}\mathrm{d}z) \times I^{r-i}(v, w, -\infty, \alpha, -\infty, \beta)
\end{aligned}$$

Note that the derivative of $f(\delta)$ with respect to δ is undefined at only one point, namely, the value of δ for which $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} = \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$.

Case 1.3. $j = n$. If $t_i \cdot v = \alpha' + v[n]\delta > \alpha$ or $t_i \cdot w = \beta' + w[n]\delta > \beta$, then t_i has a bigger dot product than t_k with at least one of v or w, and therefore, $Pr(E_k^e|i,j,\delta) = 0$. Otherwise:

$$\begin{aligned}
f(\delta) &= Pr(t_i \cdot v \leq \alpha \wedge t_i \cdot w \leq \beta|i,j,\delta)Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= Pr(t_r \cdot v \leq \alpha \wedge t_r \cdot w \leq \beta)^{r-i} \\
&= I^{r-i}(v, w, -\infty, \alpha, -\infty, \beta)
\end{aligned}$$

Note that the derivative of $f(\delta)$ with respect to δ is undefined only for two values, namely, when $\alpha = \alpha' + v[n]\delta$ and $\beta = \beta' + w[n]\delta$.

Case 2. $k > i$. Let $\max\{t_1 \cdot v, \ldots, t_{i-1} \cdot v\} = \alpha$ and $\max\{t_1 \cdot w, \ldots, t_{i-1} \cdot w\} = \beta$. $t_k \cdot v > \alpha$ and $t_k \cdot w > \beta$ for t_k to be the color assigned to both v and w. Then, let A be the event $t_k \cdot v \geq \alpha \wedge t_k \cdot w \geq \beta$ and B_l be the event $t_l \cdot v \leq t_k \cdot v \wedge t_l \cdot w \leq t_k \cdot w$, $l \geq i, l \neq k$.

Note that the events B_l in this case are not independent. However, they are independent for fixed values of $t_k \cdot v$ and $t_k \cdot w$. In what follows, we will, at appropriate points, fix $t_k \cdot v$ and $t_k \cdot w$ to be in some infinitesimal intervals, and then integrate over these intervals. Within such an integral, the values of $t_k \cdot v$ and $t_k \cdot w$ may be treated as fixed, and therefore, the events corresponding to the B_l's with the above values fixed become independent.

Case 2.1. $j < n - 1$.

$$
\begin{aligned}
f(\delta) &= Pr(A \wedge B_i \wedge \ldots B_{k-1} \wedge B_{k+1} \wedge \ldots B_r | i, j, \delta) \\
&= \textstyle\int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (Pr((x \leq t_k \cdot v \leq x + \mathrm{d}x) \wedge (y \leq t_k \cdot w \leq y + \mathrm{d}y)) \\
&\qquad\qquad Pr(t_i \cdot v \leq x \wedge t_i \cdot w \leq y | i, j, \delta)) \\
&\qquad\qquad \textstyle\prod_{l=i+1,\ldots k-1,k+1,\ldots r} Pr(t_l \cdot v \leq x \wedge t_l \cdot w \leq y | i, j, \delta)) \\
&= \textstyle\int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) \\
&\qquad\qquad Pr(\alpha' + \delta v[j] + t_i \cdot v[j+1, n] \leq x \wedge \beta' + \delta w[j] + t_i \cdot w[j+1, n] \leq y) \\
&\qquad\qquad I^{r-i-1}(v, w, -\infty, x, -\infty, y)) \\
&= \textstyle\int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) \\
&\qquad\qquad I(v[j+1 \ldots n], w[j+1 \ldots n], -\infty, x - \alpha' - v[j]\delta, -\infty, y - \beta' - w[j]\delta) \\
&\qquad\qquad I^{r-i-1}(v, w, -\infty, x, -\infty, y))
\end{aligned}
$$

Case 2.2. $j = n - 1$. Assume that $v[n]$ and $w[n]$ are positive. The remaining cases are similar.

$$
\begin{aligned}
f(\delta) &= Pr(A \wedge B_i \wedge \ldots B_{k-1} \wedge B_{k+1} \wedge \ldots B_r | i, n-1, \delta) \\
&= \textstyle\int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (Pr((x \leq t_k \cdot v \leq x + \mathrm{d}x) \wedge (y \leq t_k \cdot w \leq y + \mathrm{d}y)) \\
&\qquad\qquad \textstyle\prod_{l=i,\ldots k-1,k+1,\ldots r} Pr(t_l \cdot v \leq x \wedge t_l \cdot w \leq y | i, n-1, \delta)) \\
&= \textstyle\int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) \\
&\qquad\qquad Pr(\alpha' + \delta v[n-1] + t_i[n] v[n] \leq x \wedge \beta' + \delta w[n-1] + t_i[n] w[n] \leq y) \\
&\qquad\qquad \textstyle\prod_{l=i+1,\ldots k-1,k+1,\ldots r} Pr(t_l \cdot v \leq x \wedge t_l \cdot w \leq y)) \\
&= \textstyle\int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) (\frac{1}{\sqrt{2\pi}} \int_{z=-\infty}^{\min\{\frac{x-\alpha'-v[n-1]\delta}{v[n]}, \frac{y-\beta'-w[n-1]\delta}{w[n]}\}} \mathrm{e}^{-z^2/2} \mathrm{d}z) \\
&\qquad\qquad I^{r-i-1}(v, w, -\infty, x, -\infty, y)) \\
&= \textstyle\frac{1}{\sqrt{2\pi}} \int_{z=-\infty}^{\infty} \int_{x=\max\{\alpha, \alpha' + v[n]z + v[n-1]\delta\}}^{\infty} \int_{y=\max\{\beta, \beta' + w[n]z + w[n-1]\delta\}}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) \\
&\qquad\qquad I^{r-i-1}(v, w, -\infty, x, -\infty, y) \mathrm{e}^{-z^2/2}) \mathrm{d}z
\end{aligned}
$$

Note that the derivative of $f(\delta)$ with respect to δ is undefined only when $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} = \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$. We see this by the following argument. Consider the values of δ for which $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} < \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$. The above expression for $f(\delta)$ can then be split up into a sum of three terms described below. From the resulting expression, it is clear that it is differentiable for all values of δ such that $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} < \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$. A similar argument shows that $f(\delta)$ is is differentiable for all values of δ such that $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} > \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$.

$$
\begin{aligned}
f(\delta) = & \\
\frac{1}{\sqrt{2\pi}} & \int_{z=-\infty}^{\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}} \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y) \\
& \qquad I^{r-i-1}(v,w,-\infty,x,-\infty,y)\mathrm{e}^{-z^2/2}\mathrm{d}z) \\
+\frac{1}{\sqrt{2\pi}} & \int_{z=\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}}^{\frac{\beta-\beta'-w[n-1]\delta}{w[n]}} \int_{x=\alpha'+v[n]z+v[n-1]\delta}^{\infty} \int_{y=\beta}^{\infty} (I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y) \\
& \qquad I^{r-i-1}(v,w,-\infty,x,-\infty,y)\mathrm{e}^{-\frac{z^2}{2}}\mathrm{d}z) \\
+\frac{1}{\sqrt{2\pi}} & \int_{z=\frac{\beta-\beta'-w[n-1]\delta}{w[n]}}^{\infty} \int_{x=\alpha'+v[n]z+v[n-1]\delta}^{\infty} \int_{y=\beta'+w[n]z+w[n-1]\delta}^{\infty} (I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y) \\
& \qquad I^{r-i-1}(v,w,-\infty,x,-\infty,y)\mathrm{e}^{-\frac{z^2}{2}}\mathrm{d}z)
\end{aligned}
$$

Case 2.3. $j = n$. Since $t_i[n]$ is assigned to δ and all other components of t_i are fixed, $t_k \cdot v > \max\{\alpha, \alpha' + v[n]\delta\}$ and $t_k \cdot w > \max\{\beta, \beta' + w[n]\delta\}$ for t_k to be the color assigned to both v and w.

$$
\begin{aligned}
f(\delta) &= Pr(A \wedge B_i \wedge \ldots B_{k-1} \wedge B_{k+1} \wedge \ldots B_r | i, n, \delta) \\
&= \int_{y=\max\{\beta,\beta'+w[n]\delta\}}^{\infty} \int_{x=\max\{\alpha,\alpha'+v[n]\delta\}}^{\infty} (Pr((x \le t_k \cdot v \le x + \mathrm{d}x) \wedge (y \le t_k \cdot w \le y + \mathrm{d}y)) \\
&\qquad \textstyle\prod_{l=i+1,\ldots k-1,k+1,\ldots r} Pr(t_l \cdot v \le x \wedge t_l \cdot w \le y | i, n, \delta)) \\
&= \int_{y=\max\{\beta,\beta'+w[n]\delta\}}^{\infty} \int_{x=\max\{\alpha,\alpha'+v[n]\delta\}}^{\infty} (Pr((x \le t_k \cdot v \le x + \mathrm{d}x) \wedge (y \le t_k \cdot w \le y + \mathrm{d}y)) \\
&\qquad \textstyle\prod_{l=i+1,\ldots k-1,k+1,\ldots r} Pr(t_l \cdot v \le x \wedge t_l \cdot w \le y)) \\
&= \int_{\max\{\beta,\beta'+w[n]\delta\}}^{\infty} \int_{\max\{\alpha,\alpha'+v[n]\delta\}}^{\infty} (I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y) \\
&\qquad I^{r-i-1}(v,w,-\infty,x,-\infty,y)
\end{aligned}
$$

Note that the derivative of the above expression with respect to δ is undefined only for two values, namely, when $\alpha = \alpha' + v[n]\delta$ and $\beta = \beta' + w[n]\delta$.

Case 3. $k = i$. Let $\max\{t_1 \cdot v, \ldots, t_{i-1} \cdot v\} = \alpha$ and $\max\{t_1 \cdot w, \ldots, t_{i-1} \cdot w\} = \beta$. $t_i \cdot v > \alpha$ and $t_i \cdot w > \beta$, for t_i to be the color assigned to both v and w. Then, let A be the event $t_i \cdot v \ge \alpha \wedge t_i \cdot w \ge \beta$ and B_l be the event $t_l \cdot v \le t_i \cdot v \wedge t_l \cdot w \le t_i \cdot w$, $l > i$.

Again, note that the events B_l in this case are not independent. However, they are independent for fixed values of $t_i \cdot v$ and $t_i \cdot w$. Then, as in Case 2:

Case 3.1. $j < n - 1$.

$$
\begin{aligned}
f(\delta) &= Pr(A \wedge B_{i+1} \wedge \ldots \wedge B_r | i, j, \delta) \\
&= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v[j+1\ldots n], w[j+1\ldots n], x - \alpha' - v[j]\delta, x + \mathrm{d}x - \alpha' - v[j]\delta, \\
&\qquad y - \beta' - w[j]\delta, y + \mathrm{d}y - \beta' - w[j]\delta) \\
&\qquad I^{r-i}(v,w,-\infty,x,-\infty,y))
\end{aligned}
$$

Case 3.2. $j = n - 1$. Assume that $v[n]$ and $w[n]$ are positive. The other cases are similar.

$$
\begin{aligned}
f(\delta) &= Pr(A \wedge B_{i+1} \wedge \ldots \wedge B_r | i, n-1, \delta) \\
&= \frac{1}{\sqrt{2\pi}} \int_{z=\max\{\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}, \frac{\beta-\beta'-w[n-1]\delta}{w[n]}\}}^{\infty} \\
&\qquad (I^{r-i}(v,w,-\infty,\alpha'+v[n-1]\delta+v[n]z,-\infty,\beta'+w[n-1]\delta+w[n]z)\mathrm{e}^{-z^2/2}\mathrm{d}z)
\end{aligned}
.
$$

Note that the derivative of the above expression with respect to δ is undefined only when $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} = \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$.

Case 3.3. $j = n$. If $v[n]\delta + \alpha' < \alpha$ or $w[n]\delta + \beta' < \beta$ then this probability is 0. Otherwise,

$f(\delta) = Pr(A \wedge B_{i+1} \wedge \ldots \wedge B_r | i, j, \delta) = I^{r-i}(v, w, -\infty, \alpha' + v[n]\delta, -\infty, \beta' + w[n]\delta)$

Note that the derivative of the above expression with respect to δ is possibly undefined only for two at most 2 values, namely, when $\alpha = \alpha' + v[n]\delta$ and $\beta = \beta' + w[n]\delta$.

6 Computing $I(b, b', x, y, x', y')$

Recall that $I(b, b', x, y, x', y')$ denotes $Pr((x \leq a \cdot b \leq y) \wedge (x' \leq a \cdot b' \leq y'))$, where a is a vector whose entries are independent and normally distributed with mean 0 and variance 1. We show how to compute this probability.

Let b and b' be h dimensional. Note that $h \geq 2$. Consider the h dimensional coordinate system with respect to which b, b' are specified. Note that a naive way to compute I is to perform a sequence of h nested integrals. This seems hard to do in polynomial time with the required error. We use the following method instead.

Note that since each coordinate of a is normally distributed with mean 0 and variance 1, a has a spherically symmetric distribution. We rotate the coordinate system so that $b = (b_1, 0, \ldots, 0)$ and $b' = (b'_1, b'_2, 0 \ldots, 0)$, where $b_1, b'_2 \geq 0$. As we will show shortly, both b_1, b'_2 will be strictly positive for all our calls to I. Let $a' = (a'_1, ..., a'_n)$ be the coordinates of a under the rotated coordinate system. The following lemma is key.

Lemma 6.1 *The probability distribution of a' is identical to that of a. That is all the coordinates of a' are independently distributed according to the normal distribution with mean 0 and variance 1.*

Proof. Follows from the fact that a is spherically symmetric and a' is a rotation of a. $\square$

Note that $a' \cdot b = a_1 b_1$ and $a' \cdot b' = a_1 b'_1 + a_2 b'_2$. Now $I(b, b', x, y, x', y')$ denotes $Pr((x \leq a_1 b_1 \leq y) \wedge (x' \leq a_1 b'_1 + a_2 b'_2 \leq y'))$. This equals

$$Pr((\tfrac{x}{b_1} \leq a_1 \leq \tfrac{y}{b_1}) \wedge (\tfrac{x' - a_1 b'_1}{b'_2} \leq a_2 \leq \tfrac{y' - a_1 b'_1}{b'_2})) = \frac{1}{2\pi} \int_{x/b_1}^{y/b_1} \mathrm{e}^{-\frac{z^2}{2}} \left(\int_{(x' - z b'_1)/b'_2}^{(y' - z b'_1)/b'_2} \mathrm{e}^{-\frac{z'^2}{2}} \ \mathrm{d}z' \right) \ \mathrm{d}z$$

Lemma 6.2 $|b_1| = \Omega(\epsilon) = \Omega(\frac{1}{n^3})$ *and* $|b'_2| = \Omega(\epsilon^2) = \Omega(\frac{1}{n^6})$.

Proof. Recall that by the rounding in Section 4, $|b_1| = |b| = \Omega(\epsilon)$. $|b'_2| = \sqrt{||b'|^2 - \frac{(b \cdot b')^2}{|b|^2}|}$, since b'_2 is just the projection of b' on the line orthogonal to b in the plane containing b and b'. By Lemma 4.1 and the fact that b, b' are of the form $v[h \ldots n], w[h \ldots n]$ in all the calls we make to $I()$, $|b'_2| = \Omega(\epsilon^2)$. $\square$

7 Discretizing $t_i[j]$

For the purpose of this section, assume that integrations can be performed exactly. This will be dealt with in Section 8.

The values δ we choose for $t_i[j]$ will be multiples of $\Theta(1/n^{11})$ in the range $-n^{2.5} \ldots n^{2.5}$. We need to show that restricting and discretizing the range of δ causes $O(\frac{1}{n^2})$ error for each of the conditionality variables. This is shown in the rest of this section.

Lemma 7.1 shows that considering values between $-n^{2.5} \ldots n^{2.5}$ only causes an error of $O(\frac{1}{n^2})$ for each of the conditionality variables.

Lemma 7.1 *$\min_\delta p(\delta)$ can be estimated within an additive $|r|E|\frac{1}{n^5}| \leq |\frac{1}{n^2}|$ while only considering values between $-n^{2.5} \ldots n^{2.5}$ for δ.*

Proof. Note that $Pr(E_k^e|t_1[1]\ldots t_i[j-1]) == \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(\delta)\mathrm{e}^{-\frac{\delta^2}{2}}\mathrm{d}\delta$
$\leq \frac{1}{\sqrt{2\pi}}(\int_{-n^{2.5}}^{n^{2.5}} f(\delta)\mathrm{e}^{-\frac{\delta^2}{2}}\mathrm{d}\delta + 2\int_{n^{2.5}}^{\infty}\mathrm{e}^{-\frac{\delta^2}{2}}\mathrm{d}\delta)$, since $0 \leq f(\delta) \leq 1$. Now,
$\frac{2}{\sqrt{2\pi}}\int_{n^{2.5}}^{\infty}\mathrm{e}^{-\frac{\delta^2}{2}}\mathrm{d}\delta \leq \frac{1}{n^5}$, by Chebyschev's inequality. □

Next, we show that the discretization δ in the above range causes $O(\frac{1}{n^2})$ error for each of the conditionality variables. We consider the three cases, $j = n$, $j = n-1$ and $j < n-1$, separately. Lemma 7.2 proves the above when $j < n-1$. Note that for $j < n-1$, $f'(\delta)$ is always defined. For $j = n-1$ and $j = n$, a similar proof holds with the following difference. As shown in Section 5, $f'(\delta)$ is undefined at at most two values of δ when $j = n-1$ or $j = n$; so when $j = n-1, n$, $p'(\delta)$ is undefined only at $O(r|E|) = O(n^3)$ values of δ. We add these points to our discretization. These divide the range $-n^{2.5} \ldots n^{2.5}$ into $O(n^3)$ subranges, in each of which $f'(\delta)$ and $p'(\delta)$ are defined. In each of these ranges, a proof similar to that of Lemma 7.2 shows the needful.

Lemma 7.2 *Suppose $j < n-1$. Then $|\frac{\mathrm{d}Pr(E_k^e|i,j,\delta)}{\mathrm{d}\delta}| = |f'(\delta)| = O(n^6)$. Therefore $|f(\delta \pm O(\frac{1}{n^{11}})) - f(\delta)| \leq O(\frac{1}{n^{11}})O(n^6)$ and $|p(\delta \pm O(\frac{1}{n^{11}})) - p(\delta)| \leq O(\frac{1}{n^{11}})O(n^6)r|E| = O(\frac{1}{n^2})$.*

Proof. Note that the function f depends upon which of Cases 1.1, 2.1, 3.1 hold in Section 5. We show for one representative case, i.e., Case 2.1. The other cases can be shown similarly.

For Case 2.1: $f(\delta) = \int_\alpha^\infty \int_\beta^\infty g(x,y)h(x,y,\delta)\mathrm{d}y\mathrm{d}x$,
where $g(x,y)\mathrm{d}y\mathrm{d}x = I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y)I^{r-i-1}(v,w,-\infty,x,-\infty,y)$ and
$h(x,y,\delta) = I(v[j+1\ldots n], w[j+1\ldots n], -\infty, x - t_i\cdot v[1\ldots j-1] - v[j]\delta, -\infty, y - t_i\cdot w[1\ldots j-1] - w[j]\delta)$.

$|f'(\delta)| \leq \int_\alpha^\infty\int_\beta^\infty |g(x,y)||\frac{\partial h(x,y,\delta)}{\partial\delta}|\mathrm{d}y\mathrm{d}x \leq \max_{x,y}|\frac{\partial h(x,y,\delta)}{\partial\delta}|$, since $\int_\alpha^\infty\int_\beta^\infty g(x,y)\mathrm{d}y\mathrm{d}x$ is a probability and therefore ≤ 1.

We show that $|f'(\delta)| = O(n^6)$ by estimating $\max_{x,y}|\frac{\partial h(x,y,\delta)}{\partial\delta}|$.

Let $c(x,\delta) = x - t_i\cdot v[1\ldots j-1] - v[j]\delta = c' - v[j]\delta$ and
$d(y,\delta) = y - t_i\cdot w[1\ldots j-1] - w[j]\delta = d' - w[j]\delta$.
From Section 6,
$h(x,y,\delta) = \frac{1}{2\pi}\int_{-\infty}^{c(x,\delta)/b_1}\mathrm{e}^{-\frac{z^2}{2}}(\int_{-\infty}^{(d(y,\delta)-zb_1')/b_2'}\mathrm{e}^{-\frac{z'^2}{2}}\mathrm{d}z')\mathrm{d}z$, where b_1, b_1', b_2' are the obtained by rotating the coordinates, as in Section 6.

Let $G(y,\delta,l) = \frac{1}{\sqrt{2\pi}}\int \mathrm{e}^{-\frac{l^2}{2}}H(y,\delta,l)\mathrm{d}l$, where $H(y,\delta,l) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(d(y,\delta)-lb_1')/b_2'}\mathrm{e}^{-\frac{z'^2}{2}}\mathrm{d}z'$.

Then $|\frac{\partial h}{\partial\delta}| = |A+B| \leq |A| + |B|$, where $A = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c(x,\delta)/b_1}\mathrm{e}^{-\frac{l^2}{2}}\frac{\partial H(y,\delta,l)}{\partial\delta}\mathrm{d}l$
and $B = \frac{\partial G}{\partial l}_{l=c(x,\delta)/b_1}\frac{\partial l}{\partial\delta}_{l=c(x,\delta)/b_1}$.

Note that $|\frac{\partial G}{\partial l}| \leq 1$ for all l, since $H(y,\delta,l) \leq 1$.
Further, $|\frac{\partial l}{\partial\delta}_{l=c(x,\delta)/b_1}| = |v[j]/b_1| = O(n^3)$, by Lemma 6.2 and the discretization of the input

vectors in Section 4.
Therefore, $|B| = O(n^3)$.

$|A|$ is bounded as follows.
$|A| = |\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c(x,\delta)/b_1} \mathrm{e}^{-\frac{l^2}{2}} \frac{\partial H(y,\delta,l)}{\partial \delta} \mathrm{d}l| \leq \max_{y,\delta,l} |\frac{\partial H(y,\delta,l)}{\partial \delta}|$.

It remains to bound $\max_{y,\delta,l} |\frac{\partial H(y,\delta,l)}{\partial \delta}|$. This is done below using the same technique as above. Recall that:

$H(y,\delta,l) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(d(y,\delta)-lb'_1)/b'_2} \mathrm{e}^{-\frac{z'^2}{2}} \ \mathrm{d}z'$.

Let $J(m) = \frac{1}{\sqrt{2\pi}} \int \mathrm{e}^{-\frac{m^2}{2}} \ \mathrm{d}m$. Then

$|\frac{\partial H}{\partial \delta}| = |\frac{\mathrm{d}J}{\mathrm{d}m}_{m=(d(y,\delta)-lb'_1)/b'_2}||\frac{\partial m}{\partial \delta}_{m=(d(y,\delta)-lb'_1)/b'_2}| \leq |w[j]/b'_2| = O(n^6)$, by Lemma 6.2 and the discretization of the input vectors in Section 4.

Therefore, $|f'(\delta)| \leq |A| + |B| = O(n^6)$. The lemma follows. $\Box$

From the above 2 lemmas,

Corollary 7.3 $\min_\delta p(\delta)$ *can be estimated within an additive* $|O(\frac{1}{n^2})|$ *error while considering only multiples of* $O(\frac{1}{n^{11}})$ *between* $-n^3 \ldots n^3$ *and only* $O(n^3)$ *other points in this range for* δ*, provided integrations can be performed exactly.*

8 Performing Integrations

We show how to perform the integrations required to evaluate $f(\delta)$ (see Section 5) in polynomial time with just $O(\frac{1}{n^5})$ error. This implies that the error involved in computing $p(\delta)$ is $O(r|E|\frac{1}{n^5}) = O(\frac{1}{n^2})$ as required by Theorem 3.1.

The following lemmas will be used. The first of these lemmas describes the general form of each function we need to integrate.

Lemma 8.1 *Each integration we perform can be expressed in the following form:*

$$\int_l^m \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h)) \mathrm{d}h$$

for some function $\mathcal{G}(h)$*, where* $H()$ *is such that* $0 \leq H(e) \leq 1$*, for all* e*.*

Proof. This is easily verified by an inspection of the expressions to be integrated in Section 5 and the integral for $I()$ in Section 6. The functions $H()$ are always probabilities. The only fact to be noted is that $I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y)$ which appears in the integrals in Case 2 of Section 5 equals:

$$\frac{1}{2\pi} \int_{x/b_1}^{(x+\mathrm{d}x)/b_1} \mathrm{e}^{-\frac{z^2}{2}} (\int_{(y-zb'_1)/b'_2}^{(y+\mathrm{d}y-zb'_1)/b'_2} \mathrm{e}^{-\frac{z'^2}{2}} \ \mathrm{d}z') \ \mathrm{d}z$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{b_1} \mathrm{e}^{-\frac{x^2}{2b_1^2}} \frac{1}{\sqrt{2\pi}} \frac{1}{b'_2} \mathrm{e}^{-\frac{(y-x\frac{b'_1}{b_1})^2}{2(b'_2)^2}} \mathrm{d}y\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{h^2}{2}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(h')^2}{2}} \mathrm{d}h'\mathrm{d}h$$

where $v = (b_1, 0, \ldots, 0)$ and $w = (b_1', b_2', 0, \ldots, 0)$ in the rotated coordinate system as in Section 6 and the last equality is obtained by a change of variables $h = \frac{x}{b_1}$ and $h' = \frac{y - \frac{xb_1'}{b_1}}{b_2'}$. This change of variables affects the limits of the integration, but we are not claiming any special properties for the limits.

Similarly,

$$I(v[j+1\ldots n], w[j+1\ldots n], x-\alpha'-v[j]\delta, x+\mathrm{d}x-\alpha'-v[j]\delta, y-\beta'-w[j]\delta, y+\mathrm{d}y-\beta'-w[j]\delta)$$

which appears in the integrals in Case 3 of Section 5 equals

$$\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{h^2}{2}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(h')^2}{2}}\mathrm{d}h'\mathrm{d}h$$

where $v[j+1\ldots n] = (b_1, 0, \ldots, 0)$ and $w[j+1\ldots n] = (b_1', b_2', 0, \ldots, 0)$ in the rotated coordinate system as in Section 6, and the last equality is obtained by a change of variables $h = \frac{x-\alpha'-v[j]\delta}{b_1}$ and $h' = \frac{y-\beta'-w[j]\delta-\frac{(x-\alpha'-v[j]\delta)b_1'}{b_1}}{b_2'}$. $\square$

The next lemma shows that limits of each integration we perform can be clipped to some polynomial range.

Lemma 8.2

$$\int_{\max\{l,-n^a\}}^{\min\{m,n^a\}}\frac{1}{\sqrt{2\pi}}e^{-\frac{h^2}{2}}H(\mathcal{G}(h))\mathrm{d}h \le \int_l^m \frac{1}{\sqrt{2\pi}}e^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h$$
$$\le \int_{\max\{l,-n^a\}}^{\min\{m,n^a\}}\frac{1}{\sqrt{2\pi}}e^{-\frac{h^2}{2}}H(\mathcal{G}(h))\mathrm{d}h + O(\frac{1}{n^{2a}})$$

for all $a > 0$.

Proof. The first inequality is obvious. The second is derived using Chebyschev's inequality as follows.

$\int_l^m \frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h$

$\le \int_{\max\{l,-n^a\}}^{\min\{m,n^a\}}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h + \int_{-\infty}^{-n^a}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h + \int_{n^a}^{\infty}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h$

$\le \int_{\max\{l,-n^a\}}^{\min\{m,n^a\}}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h + \int_{-\infty}^{-n^a}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}})\mathrm{d}h + \int_{n^a}^{\infty}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}})\mathrm{d}h$

$\le \int_{\max\{l,-n^a\}}^{\min\{m,n^a\}}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h + 2\int_{n^a}^{\infty}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}})\mathrm{d}h$

$\le \int_{\max\{l,-n^a\}}^{\min\{m,n^a\}}\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h + 2\frac{1}{n^{2a}}$. Chebyschev's inequality is used in the last step above. The fact that $0 < H(e) < 1$ for all e is used in the second step. $\square$

The next lemma is classical and will be used to show that each integration can be converted to a summation by discretizing the range between the limits of integration.

Lemma 8.3 $|\int_l^{l+\rho}\frac{1}{\sqrt{2\pi}}(e^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h - \frac{1}{\sqrt{2\pi}}e^{-\frac{l^2}{2}}H(\mathcal{G}(l))\rho| \le M\rho^2$ *where* M *upper bounds the derivative of* $\frac{1}{\sqrt{2\pi}}(e^{-\frac{h^2}{2}}H(\mathcal{G}(h)))$ *with respect to* h.

Algorithm for Performing Integrations. The above three lemmas lead to the following algorithm for performing integrations. Consider a particular integral $\int_l^m \frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h$. We first replace the above integral by

$$\int_{\max\{l,-n^a\}}^{\min\{m,n^a\}} \frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))\mathrm{d}h$$

Here a will be fixed later. Next, we convert this integral to a sum by dividing the range between the limits of integration into steps of size $\frac{1}{n^b}$, for some b to be fixed later.

Suppose the derivative of $\frac{1}{\sqrt{2\pi}}(\mathrm{e}^{-\frac{h^2}{2}}H(\mathcal{G}(h)))$ is bounded by $O(n^e)$. This is yet to be shown; we will do so shortly. First, we compute the total error incurred above.

By Lemma 8.2, clipping the limits of integration incurs an error of $O(\frac{1}{n^{2a}})$. By Lemma 8.3, the error incurred in each step of the summation is $O(\frac{n^e}{n^{2b}})$, assuming there is no error in computing $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{l^2}{2}}H(\mathcal{G}(h))$. However, $H()$ itself may have been obtained as a result of performing a nested integration or as a product of $O(n)$ distinct integrations nested one level deeper (as in Case 2.2 of Section 5, for example). This implies that the value of $H()$ computed itself will have some error. So suppose we have computed each of these nested integrations within an error of $O(\frac{1}{n^f})$. Then the error in $H()$ is $O(\frac{1}{n^{f-1}})$. Therefore, the error incurred in each step of the summation is $O(\frac{n^e}{n^{2b}}+\frac{1}{n^{f-1}n^b})$; this sums to $O(\frac{n^{e+a}}{n^b}+\frac{n^a}{n^{f-1}})$ over all $2n^an^b$ steps. The total error is thus $O(\frac{1}{n^{2a}}+\frac{n^{e+a}}{n^b}+\frac{n^a}{n^{f-1}})$ and the time taken for this integration (ignoring the time taken for the nested integrals in $H()$) is $O(n^{a+b})$.

Finally, note that the depth of nesting in our integrals is at most 5 (in Case 2.2 of Section 5, it is 5). It can be easily seen that starting with the innermost integral and working outwards, values a, b can be chosen for these successive integrals based upon the respective e, f values so that the final error is $O(\frac{1}{n^5})$. This ensures that these integrations can be performed with $O(\frac{1}{n^2})$ error in polynomial time.

It now suffices to show that the derivatives of all integrands are bounded by $O(n^e)$, for some e.

Bounding Derivatives of Integrands in $I()$. Recall that

$$I(b,b',x,y,x',y') = \frac{1}{\sqrt{2\pi}}\int_{x/b_1}^{y/b_1} \mathrm{e}^{\frac{-z^2}{2}}\left(\int_{(x'-zb_1')/b_2'}^{(y'-zb_1')/b_2'} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-z'^2}{2}}\mathrm{d}z'\right)\mathrm{d}z.$$

Here b, b' have been rotated so that $b=(b_1,0,...,0)$ and $b'=(b_1',b_2',0,...,0)$

The derivative of $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-z'^2}{2}}$ with respect to z' is $-\frac{1}{\sqrt{2\pi}}z'\mathrm{e}^{\frac{-z'^2}{2}}$ which is bounded in absolute value by $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-1}{2}}$, a constant.

Next, we compute the derivative of the outer integrand. We first denote the inner integral by $h(z)$. Then the derivative of the function to be integrated, that is, $\frac{1}{\sqrt{2\pi}}h(z)\mathrm{e}^{\frac{-z^2}{2}}$ is

$$\frac{1}{\sqrt{2\pi}}(-z\mathrm{e}^{\frac{-z^2}{2}}h(z)+\frac{1}{2\pi}\mathrm{e}^{\frac{-z^2}{2}}(-b_1'/b_2')(\mathrm{e}^{\frac{-1}{2}(\frac{y'-zb_1'}{b_2'})^2}-\mathrm{e}^{-\frac{1}{2}(\frac{x'-zb_1'}{b_2'})^2}))$$

The first term in this sum is bounded in absolute value by a constant as $h(z)\le 1$ and the second term is bounded by $O(n^6)$ by Lemma 6.2. Hence the derivative is bounded by $O(n^6)$.

Bounding the Derivatives of other Integrands. We bound the derivatives for the integrands in Case 2.2 of Section 5. This is the most complicated case. For other cases, a similar procedure works.

Recall that in this case, the conditional probability $f(\delta)$ can be split into three terms. We show how the derivatives of the integrands involved in the first term can be bounded by polynomial functions of n. The remaining two terms are similar.

The first term is:

$$g(\delta)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}}\int_{x=\alpha}^{\infty}\int_{y=\beta}^{\infty}I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y)I^{r-i-1}(v,w,-\infty,x,-\infty,y)\mathrm{e}^{\frac{-z^2}{2}}\mathrm{d}z.$$

To simplify notation, we denote by c the value $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}$, As in Section 6, let the coordinate system be so rotated that the new coordinates of v are $(b_1,0,...,0)$ and the new coordinates of w are $(b_1',b_2',0,...,0)$ where $b_1,b_2'\geq 0$. Recall from Section 6 that:

$$I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y)=\frac{1}{\sqrt{2\pi}}\frac{1}{b_1}\mathrm{e}^{\frac{-x^2}{2b_1{}^2}}\frac{1}{\sqrt{2\pi}}\frac{1}{b_2'}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\mathrm{d}y\mathrm{d}x$$

Therefore

$$g(\delta)=\int_{z=-\infty}^{c}\int_{x=\alpha}^{\infty}\int_{y=\beta}^{\infty}I^{r-i-1}(v,w,-\infty,x,-\infty,y)\frac{1}{\sqrt{(2\pi)}b_1}\mathrm{e}^{\frac{-x^2}{2b_1{}^2}}\frac{1}{\sqrt{2\pi}b_2'}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}\mathrm{d}y\mathrm{d}x\mathrm{d}z$$

We first consider innermost integral, that is with respect to y. The term to be integrated is:

$$I^{r-i-1}(v,w,-\infty,x,-\infty,y)\frac{1}{\sqrt{2\pi}b_2'}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}$$

The other terms are independent of y. Its derivative with respect to y is:

$$\frac{1}{\sqrt{2\pi}b_2'}(r-i-1)I^{r-i-2}(v,w,-\infty,x,-\infty,y)\frac{\partial I(v,w,-\infty,x,-\infty,y)}{\partial y}\mathrm{e}^{-1/2(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}$$
$$-\frac{1}{\sqrt{2\pi}b_2'}I^{r-i-1}(v,w,-\infty,x,-\infty,y)(\frac{y-\frac{x}{b_1}b_1'}{b_2'{}^2})\mathrm{e}^{-1/2(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}$$

Now

$$\frac{\partial I(v,w,-\infty,x,-\infty,y)}{\partial y}=\frac{1}{2\pi b_2'}\int_{-\infty}^{\frac{x}{b_1}}\mathrm{e}^{-\frac{1}{2}z^2}\mathrm{e}^{-\frac{1}{2}(\frac{y-b_1'z}{b_2'})^2}\mathrm{d}z=O(\frac{1}{b_2'})$$

Observe that as the functions I, $x\mathrm{e}^{-x^2/2}$ are all bounded by constants, the value of the above derivative is bounded in absolute value by $O(\frac{(r-i-1)}{(b_2')^2}+\frac{1}{(b_2')^2})$. Since, $r-i-1\leq n$, $b_1=\Omega(\frac{1}{n^3})$, b_2' is $\Omega(\frac{1}{n^6})$ by Lemma 6.2, the above derivative is bounded by $O(n^{13})$.

The second innermost integral, i.e., the one with respect to x is considered next. The function inside the integral is $h(x)\frac{1}{\sqrt{(2\pi)}b_1}\mathrm{e}^{\frac{-x^2}{2b_1^2}}$, where

$$h(x)=\int_{y=\beta}^{\infty}I^{r-i-1}(v,w,-\infty,x,-\infty,y)\frac{1}{\sqrt{2\pi}b_2'}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\mathrm{d}y$$

Since $0 \leq I() \leq 1$, $h(x) = O(1)$. The derivative with respect to x is:

$$-\frac{x}{\sqrt{2\pi}b_1^3}\mathrm{e}^{-\frac{x^2}{2b_1^2}}h(x)$$
$$+\frac{1}{\sqrt{2\pi}b_1}\mathrm{e}^{-\frac{x^2}{2b_1^2}}\int_\beta^\infty(r-i-1)I^{r-i-2}(v,w,-\infty,x,-\infty,y)\frac{\partial I(v,w,-\infty,x,-\infty,y)}{\partial x}\frac{1}{\sqrt{2\pi}b_2'}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\mathrm{d}y$$
$$+\frac{1}{\sqrt{2\pi}b_1}\mathrm{e}^{-\frac{x^2}{2b_1^2}}\int_\beta^\infty I^{r-i-1}(v,w,-\infty,x,-\infty,y)\frac{1}{\sqrt{2\pi}b_2'}\frac{b_1'(y-\frac{xb_1'}{b_1})}{(b_2')^2b_1}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\mathrm{d}y$$

Here,

$$\frac{\partial I(v,w,-\infty,x,-\infty,y)}{\partial x} = \frac{1}{2\pi b_1}\mathrm{e}^{-\frac{x^2}{2b_1^2}}\int_{-\infty}^{\frac{y-xb_1'/b_1}{b_2'}}\mathrm{e}^{-\frac{1}{2}(z')^2}\mathrm{d}z' = O(\frac{1}{b_1})$$

Since $xe^{-\frac{x^2}{2}}, h(x), I()$ are all $O(1)$, $r-i-1 \leq n$, and

$$\int_\beta^\infty \frac{1}{b_2'}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\mathrm{d}y = O(1), \int_\beta^\infty \frac{b_1'(y-\frac{x}{b_1}b_1')}{b_1(b_2')^3}\mathrm{e}^{-\frac{1}{2}(\frac{y-\frac{x}{b_1}b_1'}{b_2'})^2}\mathrm{d}y = O(\frac{b_1'}{b_1b_2'}) = O(\frac{1}{b_1{b_2'}^2})$$

the above derivative is bounded by $O(\frac{n}{{b_1}^2}+\frac{1}{{b_2'}^2b_1^2}) = O(n^{18})$, by Lemma 6.2.

This leaves only the outermost integration, where the integrand is $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z^2}\int_{x=\alpha}^\infty h(x)\frac{1}{\sqrt{2\pi}b_1}\mathrm{e}^{-\frac{x^2}{2b_1^2}}\mathrm{d}x$, whose derivative with respect to z is $O(1)$.

From the above, it follows that Theorem 3.1 is true.

9 Comments on Derandomizing the Max-Cut Algorithm.

Recall that Lemma 4.2 of Section 4 works only when $|v \cdot w|$ is bounded away from 1. This may not be true for the case of Max Cut. However in this case, it can be easily seen that all the input vectors can be simultaneously discretized using a procedure analogous to that in Section 4, so that the following are achieved:

1. Each component of each vector is at least some inverse polynomial in absolute value.
2. The dot product of any pair of vectors changes only by an inverse polynomial in absolute value and is at most $1-\frac{1}{\mathrm{poly}(n)}$ in absolute value.
3. For each pair of vectors v, w and every h, $1 \leq h < n$, when the coordinate system is rotated so that $v[h\ldots n] = (b_1, 0, \ldots, 0)$ and $w[h\ldots n] = (b_1', b_2', 0, \ldots, 0)$, b_1 and b_2' are at least some inverse polynomial in absolute value.

To compensate for Lemma 4.2, we only need to observe that the value of the Goemans and Williamson objective function (that is, $\sum_{i,j} w_{ij}\frac{1-v_i.v_j}{2}$) for the discretized vector configuration is at least $(1-\frac{1}{\mathrm{poly}(n)})$ times that for the initial vector set (this is because, the sum of the edges weights is at most twice the value of the objective function for the initial vector set). The rest is just a matter of choosing the appropriate inverse polynomial terms.

10 Conclusions

We believe that the techniques used here can be used to derandomize a general class of randomized algorithms based on Semidefinite Programming. Loosely speaking, this class would comprise of those whose expected value calculations involve just a constant number of vectors in each "elementary" event. This class contains all randomized Semidefinite Programming based algorithms known so far. It would be nice to obtain a general theorem to this effect.

Acknowledgments

We thank Naveen Garg, Kurt Mehlhorn, David Williamson, Michel Goemans and Madhu Sudan for comments. We also thank Aravind Srinivasan for reading a draft of the manuscript and for his detailed comments.

References

[1] N. Alon, N. Kahale, *Approximating the Independence Number via the θ-Function*, Manuscript.

[2] A. Blum, *New Approximation Algorithms for Graph Coloring*, JACM, 41, 470-516, 1994.

[3] R. Boppana, M. Halldorsson, *Approximating Maximum Independent Sets by Excluding Subgraphs*, BIT, 32:180-196, 1992.

[4] W. Feller, *Introduction to Probability Theory and its Applications*, Vol. 1, Wiley Eastern Edition.

[5] U. Feige, M. Goemans, *Approximating the Value of Two Prover Proof Systems, with Applications to Max-2Sat and Max-Dicut*, 3rd Israeli Symposium on Theory and Computing Systems, pp. 182–189, 1995.

[6] A. Frieze, M. Jerrum, *Improved Approximation Algorithms for Max k-Cut and Max Bisection*, Integer Programming and Combinatorial Optimization, 1995.

[7] M. Grötschel, L. Lovász, A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, 1987.

[8] M. Goemans, D. Williamson, *0.878 Approximation Algorithms for Max CUT and Max 2SAT*, 26th Annual Symposium on the Theory of Computing, pp. 422-431, 1994. Also in Journal of the ACM, Vol. 42, No. 6, 1996, pp. 1115-1145.

[9] M. Halldorsson, *A Still Better Performance Guarantee for Approximate Graph Colouring*, IPL 45, 19-23, 1993.

[10] D.S. Johnson, *Worst Case Behaviour of Graph Coloring Algorithms*, Proceedings, 5th South-Eastern Conference on Combinatorics, Graph Theory and Computing, Congressus Numeratium X, 513-527, 1974.

[11] D. Karger, R. Motwani, M. Sudan, *Approximate Graph Coloring by Semidefinite Programming*, 35th IEEE Symposium on Foundations of Computer Science, pp. 1-10, October 1994.

[12] P. Kelsen, S. Mahajan, H. Ramesh, *Approximately Coloring 2-Colorable 3-Uniform Hypergraphs*, To appear, Scandinavian Workshop on Algorithm Theory, 1996.

[13] S. Khanna, N. Linial, S. Safra, *On the Hardness of Approximating The Chromatic Number*, 2nd Israeli Symposium on Theory and Computing Systems, pp. 250-260, 1992.

[14] P. Raghavan, *Probabilistic Construction of Deterministic Algorithms: Approximating Packing Integer Programs*, JCSS 37, 130-143, 1988.

[15] P. Raghavan, *Randomized Rounding*, Foundations of Theoretical Computer Science and Software Technology, 1994.

[16] J. Spencer, *Ten Lectures on the Probabilistic Method*, SIAM, Philadelphia, 1987.

[17] A. Wigderson, *Improving the Performance Guarantee of Approximate Graph Coloring*, JACM, 30, 729-735, 1983.

[18] M. Yannakakis, *On the Approximation of Maximum Satisfiability*, 3rd Annual ACM-SIAM Symposium on Discrete Algorithms, 1-9, 1992.