SIAM J. Comput.
Vol. 28, No. 5, pp. 1641–1663

DERANDOMIZING APPROXIMATION ALGORITHMS BASED ON SEMIDEFINITE PROGRAMMING*

SANJEEV MAHAJAN† AND H. RAMESH‡

Abstract. Remarkable breakthroughs have been made recently in obtaining approximate solutions to some fundamental NP-hard problems, namely Max-Cut, Max k-Cut, Max-Sat, Max-Dicut, Max-bisection, k-vertex coloring, maximum independent set, etc. All these breakthroughs involve polynomial time *randomized* algorithms based upon *semidefinite programming*, a technique pioneered by Goemans and Williamson.

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.

Key words. NP-hard, approximation algorithm, derandomization, semidefinite programming

AMS subject classification. 68Q25

PII. S0097539796309326

1. Introduction.

The application of *semidefinite programming* to obtaining approximation algorithms for NP-hard problems was pioneered by Goemans and Williamson [9]. This technique involves relaxing an integer program (solving which is an NP-hard problem) to a semidefinite program (which can be solved with a sufficiently small error in polynomial time).

Recall the Max-Cut problem which requires partitioning the vertex set of a given graph into two so that the number of edges going from one side of the partition to the other is maximized. In a remarkable breakthrough, Goemans and Williamson showed how semidefinite programming 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, techniques based upon semidefinite programming have led to randomized algorithms with substantially better approximation ratios for a number of fundamental problems.

Goemans and Williamson [9] obtained a .878 approximation algorithm for Max-2Sat and a .758 approximation algorithm for Max-Sat, improving upon the previously best known bound of 3/4 for both [18]. Max-2Sat requires finding an assignment to the variables of a given 2Sat formula which satisfies the maximum number of clauses. Max-Sat is the general version of this problem, where the clauses are no longer constrained to have two variables each. Goemans and Williamson [9] also obtained a .796 approximation algorithm for Max-Dicut, improving the previously best known ratio of .25 given by the random cut algorithm. This problem requires partitioning the vertex set of a given directed graph into two so that the number of edges going from the left side of the partition to the right is maximized. Feige and

*Received by the editors September 11, 1996; accepted for publication (in revised form) April 15, 1998; published electronically May 6, 1999. Part of this work was done while both authors were at Max Planck Institut für Informatik, Saarbrücken, Germany, 66123.

http://www.siam.org/journals/sicomp/28-5/30932.html

†LSI Logic, Milpitas, CA (msanjeev@lsil.com).

‡Indian Institute of Science, Bangalore, India (ramesh@csa.iisc.ernet.in).

Goemans [6] obtained further improved approximation algorithms for Max-2Sat and Max-Dicut.

Karger, Motwani, and Sudan obtained an algorithm for coloring any k-colorable graph with $O(n^{1-3/(k+1)}\log n)$ colors [12]; in particular, for 3-colorable graphs, this algorithm requires $O(n^{.25}\log n)$ colors. This improves upon the deterministic algorithm of Blum [3], which requires $O(n^{1-\frac{1}{k-4/3}}\log^{\frac{8}{5}} n)$ colors for k-colorable graphs.

Frieze and Jerrum [7] obtained a .65 approximation algorithm for Max-bisection improving the previous best known bound of .5 given by the random bisection algorithm. This problem requires partitioning the vertex set of a given graph into two parts of roughly equal size such that the number of edges going from one side of the partition to the other is maximized. 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 maximum independent set problem on graphs, which requires finding the largest subset of vertices, no two of which are connected by an edge. 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 [4].

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 and Goemans paper [6], 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.

Our contribution. The main contribution of this paper is a technique which enables derandomization of all approximation algorithms based upon semidefinite programming 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 maximum independent set with the same approximation ratios as their randomized counterparts mentioned above. However, we must mention that running times of our deterministic algorithms, though polynomial, are quite slow, for example, $O(n^{30})$ or so for 3-vertex coloring. In this paper, we do not make an effort to pinpoint the exact polynomial or to reduce the running time (within the realm of polynomials, that is).

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 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 three vectors is the Max-Dicut algorithm of Goemans and Williamson. Another example of the same is the algorithm of Alon et al. [2] 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 [8]. A weak separation oracle (see [8, p. 51],) is one which, given a point y, either asserts that y is in or close to the convex set in question or produces a hyperplane which "almost" separates all points well within the convex set from y.

One such convex set is the set of *semidefinite matrices*, i.e., those matrices whose eigenvalues are all nonnegative. 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 additive error exponentially small in the input length) 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 chose a random hyperplane through the origin whose normal is spherically symmetrically distributed; this hyperplane divides the vertex vectors into two groups, which define a cut in the obvious manner. The expected number $\mathbf{E}(W)$ of edges[1] across the cut is $\sum_{(v,w)\in E} \Pr(sign(v \cdot R) \neq sign(w \cdot R)) = \sum_{(v,w)\in E} \arccos(v \cdot w)/\pi$, 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 $\mathbf{E}(W)$ 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 identically distributed with a mean 0 and variance 1 normal distribution; i.e., the density function is $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ [5]. 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 $\mathbf{E}(W)$.

The Goemans–Williamson derandomization procedure. Goemans and Williamson actually gave the following derandomization procedure for their algorithm, which turned out to have a subtle bug described below.

From the initial set of vertex vectors in n dimensions, they obtain a new set of vertex vectors in $n-1$ dimensions satisfying the property that the expected size of the cut obtained by partitioning the new vertex vectors with a random hyperplane in $n-1$ dimensions is at least $\mathbf{E}(W)$. This procedure is repeated until the number of dimensions is down to 1, at which point partitioning the vectors becomes trivial. It remains to show how the new vertex vectors in $n-1$ dimensions are obtained from the older vertex vectors.

Consider an edge v, w, with vertex vector $v = (v_1, \ldots, v_{n-1}, v_n)$ and vertex vector $w = (w_1, \ldots, w_{n-1}, w_n)$. Recall that $R = (R_1, \ldots, R_n)$ is the normal to the random hyperplane. Goemans and Williamson obtain v', w', R' from v, w, R as follows: $R' = (R_1, \ldots, R_{n-2}, sign(R_{n-1})\sqrt{R_{n-1}^2 + R_n^2})$, $v' = (v_1, \ldots, v_{n-2}, x\cos(\alpha - \gamma))$, and $w' = (w_1, \ldots, w_{n-2}, y\cos(\beta - \gamma))$, where $\gamma = \arctan(R_n/R_{n-1})$, $\alpha = \arctan(v_n/v_{n-1})$, $\beta = \arctan(w_n/w_{n-1})$, $x = sign(v_{n-1})\sqrt{v_{n-1}^2 + v_n^2}$, and $y = sign(w_{n-1})\sqrt{w_{n-1}^2 + w_n^2}$.

These new definitions have the property that $v' \cdot R' = v \cdot R$ and $w' \cdot R' = w \cdot R$. To obtain the new vectors, one needs to decide on the value of γ. By the above property, there exists a value of γ such that $\sum_{(v,w)\in E} \Pr(sign(v' \cdot R') \neq sign(w' \cdot R')|\gamma) \geq \mathbf{E}(W)$. This value of γ is found by computing $\sum_{(v,w)\in E} \Pr(sign(v' \cdot R') \neq sign(w' \cdot R')|\gamma)$ for a polynomial number of points in the suitably discretized interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. At this point, Goemans and William claim that the vector R' is spherically symmetric for any fixed value of γ, and therefore $\Pr(sign(v' \cdot R') \neq sign(w' \cdot R')|\gamma) = (\frac{v' \cdot w'}{|v'||w'|})/\pi$ and is thus easy to compute.

The flaw lies in the fact that $r = sign(R_{n-1})\sqrt{R_{n-1}^2 + R_n^2}$ is not distributed normally with mean 0 and variance 1, even for a fixed value of γ. In fact, given γ, it can be shown to be distributed according to the density function $\frac{|r|}{2}e^{-r^2/2}$. It is not clear how $\Pr(sign(v' \cdot R') \neq sign(w' \cdot R')|\gamma)$ can be computed for this distribution of R'.

The Karger–Motwani–Sudan coloring algorithm. Our description of this algorithm is based on the conference proceedings version of their paper [12].

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

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

if these vectors are somehow constrained to be in two dimensions, then there are at most three 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*. The number of centers r will be spelled out in a few paragraphs. 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. Ties in choosing the vector with the largest projection occur with probability 0 and can therefore be ignored.

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 $\Pr(E^e) = \sum_{k=1}^{r} \Pr(E_k^e)$, where E_k^e is the event that both get color t_k. E_k^e can be written as

$$\mathrm{E}_k^e : t_k \cdot v = \max\{t_1 \cdot v, \ldots, t_r \cdot v\} \wedge t_k \cdot w = \max\{t_1 \cdot w, \ldots, t_r \cdot w\}.$$

Karger, Motwani, and Sudan [12] show the following theorem (see Theorem 7.7, Corollary 7.8, and Lemma 7.9 of [12]).

THEOREM 2.1. *For $r = O(d^{1/3} \log^{4/3} d)$, if edge $e = (v, w)$ satisfies $v \cdot w \leq -1/2 + O(1/\log r)$, then $\Pr(E^e) = O(\frac{1}{d})$, where d is the maximum degree of the graph. Thus $\sum_{e \in E} \Pr(E^e)$, i.e., the expected number of bad edges, can be made less than $n/4$ by an appropriate choice of constants.*

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 expected number of remaining vertices is at most $n/2$. Markov's inequality helps to bound their number with a reasonable probability. These vertices 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 a technique due to Wigderson [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 $n/4 + O(1)$. 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 two methods seem different, the hyperplane method can be interpreted as just the center method with two centers.

3. The derandomization scheme.

We give an overview of our derandomization scheme in this section. 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.

Our derandomization procedure has two steps. The first step described in section 3.1 is a discretization step. It is necessary for reasons which will become clear later. This step obtains a new set of vertex vectors which are "close" to the initial vertex vectors. The new vertex vectors satisfy the property that the Karger–Motwani–Sudan randomized algorithm continues to give the claimed theoretical performance on these vectors as well. This justifies the use of the new vectors in the actual derandomization process, which is the second step and is described in section 3.2.

3.1. Preprocessing vertex vectors: Discretization. Before we describe our derandomization scheme, we discretize the vertex vectors obtained from the semidefinite program so as to satisfy the following properties. Let ϵ be a parameter which is $\Theta(\frac{1}{n^2})$.

1. Each component of each vector is at least inverse polynomial (more precisely, $\Omega(\epsilon)$) in absolute value.
2. The dot product of any pair of vectors changes only by an inverse polynomial (more precisely, $O(n\epsilon)$) 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 $v[h \dots n] = (b_1, 0, \dots, 0)$ and $w[h \dots n] = (b'_1, b'_2, 0, \dots, 0)$, b_1 and b'_2 are at least some inverse polynomial (more precisely, $\Omega(\epsilon)$ and $\Omega(\epsilon^2)$, respectively) in absolute value.

The method for performing the discretization is given in Appendix 1. The purpose of the above discretization will become clear in the next subsection. Of course, we have to show that the above discretization does not cause much error. This is true because $v \cdot w$, which was at most $-\frac{1}{2}$ before the discretization, is at most $-\frac{1}{2} + O(\frac{1}{n})$ now for each edge $e = (v, w)$. Then, by Theorem 2.1, the theoretical bounds of the Karger–Motwani–Sudan randomized algorithm continue to hold for discretized vectors as well.

From now on, all our references to vectors will be to the discretized vectors.

3.2. Outline of the derandomization procedure. 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 $n/4 + O(1)$.

We order 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 $t_1[1 \cdots n], t_2[1 \cdots n], \dots, t_i[1 \cdots j-1]$ have been determined. We will show how a value for $t_i[j]$ is determined.

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

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; $p(\delta) = \sum_{e \in E} \sum_{k=1}^{r} f_{e,k}(\delta)$.

Note that both $f_{e,k}(\delta)$ and $p(\delta)$ implicitly refer to some fixed values of i and j. This will be the case throughout this paper.

Problem to be solved now. The problem now is to find a value of δ for which $p(\delta) \leq \sum_{e \in E} \sum_{k=1}^{r} \Pr(E_e^k | t_1[1 \dots n], t_2[1 \dots n], \dots, t_i[1 \dots j-1])$. In other words, we want a value of δ with the following property: the expected number of bad edges with $t_i[j]$ assigned δ and $t_1[1 \dots n], t_2[1 \dots n], \dots, t_i[1 \dots j-1]$ fixed as above, is at most the expected number of bad edges with just $t_1[1 \dots n], t_2[1 \dots n], \dots, t_i[1 \dots j-1]$ fixed as above.

Fixing $t_i[j]$. Let $\tau = \sum_{e \in E} \sum_{k=1}^{r} \Pr(E_e^k | t_1[1 \dots n], t_2[1 \dots n], \dots, t_i[1 \dots j-1])$. We want to compute a value δ such that $p(\delta) \leq \tau$. We will not be able to compute such a δ. However, we will show the following.

In Theorem 3.1, we claim that working with the discretized vertex vectors, we can compute a value κ, such that $p(\kappa)$ is within $O(1/n^2)$ of τ. Corollary 3.2 then shows that this suffices to obtain the required bound on the number of colors.

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

From the above theorem, we get 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 $n/4 + O(1)$.*

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 $n/4$. By Theorem 3.1, the expected number of bad edges after the first conditional variable is fixed is at most $n/4 + O(1/n^2)$. An easy inductive argument shows that the expected number of bad edges after the lth conditional variable is fixed is at most $n/4 + O(1/n^2)$. 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 $n/4 + O(1)$. ☐

Note that while the Karger–Motwani–Sudan algorithm ensures that the number of bad edges is less than $n/4$, our deterministic algorithm shows a slightly weaker bound, i.e., at most $n/4 + O(1)$. However, it can be seen easily that this weaker bound on the number of bad edges also suffices to obtain the bound of $O(n^{1/4} \log n)$ colors for coloring a 3-colorable graph deterministically.

The rest of this paper will be aimed at proving Theorem 3.1. This is accomplished by performing the following steps, each of which is elaborated in detail in the following sections.

Remark. The purpose of discretizing the vertex vectors earlier can be explained now. The intention is to ensure that derivatives of the functions $f_{e,k}(\delta)$ and $p(\delta)$ (with respect to δ) 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, thus facilitating discrete evaluation, which we will need to do.

3.3. Steps required to prove Theorem 3.1. The following steps are performed in our algorithm to obtain the value κ described above. Recall again that we are working with fixed values of i, j and assuming that $t_1[1 \dots n], t_2[1 \dots n], \dots, t_i[1 \dots j-1]$ have already been fixed.

Step 1. In order to compute κ, we would like to evaluate $p(\delta)$ at a number of points. However, we can afford to evaluate $p(\delta)$ only for a polynomial number of points. In section 4, we show how to obtain a set S of polynomial size such that

$\min_{\delta \in S} p(\delta) - \tau = O(\frac{1}{n^2})$. Therefore, in order to compute κ, it suffices to evaluate $p(\delta)$ at points in S, as long as the evaluation at each point in S is correct to within an additive $O(\frac{1}{n^2})$ error.

Step 2. We now need to show how $p(\delta)$ can be evaluated within an additive $O(\frac{1}{n^2})$ error for any particular point δ in polynomial time. Of course, we need to do this computation for points in S only, but the description of this step is for any general point δ.

To compute $p(\delta)$, we need to compute $f_{e,k}(\delta)$ to within an $O(\frac{1}{n^5})$ additive error, for each edge e and each center k (using the rather liberal upper bounds of $O(n^2)$ for the number of edges and $O(n)$ for the number of centers). We will describe this computation for a particular edge e and a particular center k. This will involve two substeps. The first substep will develop an expression involving a nested sequence of integrations for $f_{e,k}(\delta)$, and the second substep will actually evaluate this expression within the required error.

Substep 2a. We show how to write $f_{e,k}(\delta)$ as an expression involving a nested sequence of integrations of constant depth. This is done is two stages. In the first stage, in section 5, $f_{e,k}(\delta)$ is expressed as a nested sequence of integrations of constant depth, with the integrand comprising only basic functions and a function I, defined below. Subsequently, in section 6, we express the function I itself as an expression involving a constant number of nested integrations. This requires the use of spherical symmetry properties. The fact that the depth of integration is a constant will be significant in Substep 2b. If this were not the case, then it is not clear how $f_{e,k}(\delta)$ could be computed within the required error bounds at any point δ in polynomial time.

DEFINITION 3.3. *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* $\Pr((x \leq a \cdot b \leq y) \wedge (x' \leq a \cdot b' \leq y'))$.

Substep 2b. The expression for $f_{e,k}(\delta)$ obtained in Substep 2a is evaluated to within an $O(\frac{1}{n^5})$ additive error in this step in polynomial time. This is described in section 7.

Remarks on $\frac{\mathrm{d}f_{e,k}(\delta)}{\mathrm{d}\delta}$. We will also be concerned with the differentiability of $p(\delta)$ and therefore $f_{e,k}(\delta)$ for each edge e and each center k. From the expressions derived for $f_{e,k}(\delta)$ in section 5, the following properties about the differentiability of $f_{e,k}(\delta)$ will become clear. These properties will be used in section 4, i.e., to obtain the set S as described in Step 1. We state these properties in a lemma for future reference. The proof of this lemma in section 5 will not rely on any usage of the lemma in section 4.

LEMMA 3.4. *When $j < n-1$, $f_{e,k}(\delta)$ is differentiable (with respect to δ) for all δ. When $j = n-1$ or $j = n$, $f_{e,k}(\delta)$ is differentiable for all but at most* 2 *values of δ.*

4. Step 1: Determining set S.

We show how to obtain a set S of polynomial size such that $\min_{\delta \in S} p(\delta) - \tau = O(\frac{1}{n^2})$. Recall from section 3.2 that $p(\delta) = \sum_{e \in E} \sum_{k=1}^{r} f_{e,k}(\delta) = \sum_{e \in E} \sum_{k=1}^{r} \Pr(E_e^k | i, j, \delta)$.

We will use the following theorem from [5, Chapter 7, Lemma 2].

THEOREM 4.1. *For every $a > 0$, $\int_a^\infty \mathrm{e}^{\frac{-\delta^2}{2}} \mathrm{d}\delta \leq \frac{1}{a} \mathrm{e}^{-\frac{a^2}{2}}$.*

First, we show that we can restrict δ to the range $-3\sqrt{\ln n} \ldots 3\sqrt{\ln n}$.

LEMMA 4.2. $\min_{-3\sqrt{\ln n} < \delta < 3\sqrt{\ln n}} p(\delta) - \tau = O(\frac{1}{n^2})$.

Proof. Note that $\tau = \sum_{e \in E} \sum_{k=1}^{r} \Pr(E_k^e | t_1[1] \ldots t_i[j-1]) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} p(\delta) \mathrm{e}^{-\frac{\delta^2}{2}} \mathrm{d}\delta$.

Let δ_{min} be the point in the above range at which $p(\delta)$ is minimized in this range. Then

$$p(\delta_{min}) \leq \frac{\int_{-3\sqrt{\ln n}}^{3\sqrt{\ln n}} p(\delta) \mathrm{e}^{-\frac{\delta^2}{2}} \mathrm{d}\delta}{\int_{-3\sqrt{\ln n}}^{3\sqrt{\ln n}} \mathrm{e}^{-\frac{\delta^2}{2}} \mathrm{d}\delta} \leq \frac{\tau}{(1 - 2\int_{3\sqrt{\ln n}}^{\infty} \mathrm{e}^{-\frac{\delta^2}{2}} \mathrm{d}\delta)} \leq \frac{\tau}{1 - O(\frac{1}{n^{4.5}})}$$

by Theorem 4.1. Therefore, $p(\delta_{min}) \leq \tau(1 + O(\frac{1}{n^{4.5}})) \leq \tau + O(\frac{1}{n^{2.5}})$ as required. ∎

The set S we choose will comprise points which are multiples of $\Theta(1/n^9)$ in the range $-3\sqrt{\ln n} \dots 3\sqrt{\ln n}$. In addition, it will contain all those points in the above range at which $p(\delta)$ is not differentiable. By Lemma 3.4, there will be at most $2r|E| = O(n^3)$ such points, at most 2 for each edge-center pair. These points will be obtained in the process of writing down the expressions for $f_{e,k}(\delta)$ in section 5. The size of S is thus $O(n^9\sqrt{\ln n})$.

We need to show that evaluating $p(\delta)$ at just points in S suffices to approximate $\min_{-3\sqrt{\ln n}<\delta<3\sqrt{\ln n}} p(\delta)$ to within $O(\frac{1}{n^2})$. For this, we will need to bound the derivative of $p(\delta)$ with respect to δ, wherever it exists, in the above range. This is done in Lemma 4.3.

LEMMA 4.3. *At any point δ, $|\frac{\mathrm{d}p(\delta)}{\mathrm{d}\delta}| = O(n^7)$ whenever it exists. Consider any $\Delta = O(\frac{1}{n^9})$. Then it follows that $|p(\delta + \Delta) - p(\delta)| \leq O(\frac{1}{n^2})$, for any δ, such that $p(\delta)$ is differentiable everywhere in the range $]\delta \dots \delta + \Delta[$.*

Proof. The second part follows from the fact that $\frac{|p(\delta+\Delta)-p(\delta)|}{\Delta} = O(n^7)$. This is because the slope of $p()$ must equal $\frac{p(\delta+\Delta)-p(\delta)}{\Delta}$ at some point in the range $]\delta \dots \delta+\Delta[$.

To show the first part, we show in Appendix 2 that the derivative of each $f_{e,k}(\delta)$ is bounded by $O(n^4)$ wherever it exists. Thus the derivative of $p(\delta)$ will be bounded by $O(r|E|n^4) = O(n^7)$ as claimed. ∎

COROLLARY 4.4. *If δ_{min} is the point in the range $-3\sqrt{\ln n} < \delta < 3\sqrt{\ln n}$ at which $p(\delta)$ is minimized in this range, then $p(\delta_{min})$ can be approximated to within an $O(\frac{1}{n^2})$ additive error by evaluating $p(\delta)$ at the nearest point in S which is bigger than δ_{min}.*

From Lemma 4.2 and Corollary 4.4, we conclude that $\min_{\delta\in S} p(\delta) - \tau = O(\frac{1}{n^2})$ as required.

5. Substep 2a: Deriving expressions for $f_{e,k}(\delta)$. Recall that $f_{e,k}(\delta) = \Pr(E_k^e|i,j,\delta)$ is the probability that both endpoints v, w of e 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. For a fixed edge $e = (v, w)$ and some fixed k, we show how to express $f_{e,k}(\delta)$ in terms of some basic functions and the function $I()$ defined earlier. The exact expression will depend upon which of a number of cases occurs. However, the thing to note is that the expression in each case will have only a constant number of nested integrals. For each case, we will also determine points δ, if any, at which $f_{e,k}(\delta)$ is not differentiable. Recall that these points are necessary in defining S in section 4. Then, in section 6, we will show how $I()$ itself can be expressed in terms of basic functions and nested integrals of depth just 2.

Notation. For convenience, we use the notation $\Pr((x \leq t_k \cdot v \leq x + \mathrm{d}x) \wedge (y \leq t_k \cdot w \leq y + \mathrm{d}y))$ to denote the density function of the joint probability distribution of $t_k \cdot v$ and $t_k \cdot w$, multiplied by $\mathrm{d}x\mathrm{d}y$. Informally speaking, the above term denotes the probability that $t_k \cdot v$ is in the infinitesimal range $x \dots x + \mathrm{d}x$ and $t_k \cdot w$ is in the infinitesimal range $y \dots y+\mathrm{d}y$. Similarly, we use the notation $I(v, w, x, x+\mathrm{d}x, y, y+\mathrm{d}y)$

to denote $\Pr((x \le a \cdot b \le x + \mathrm{d}x) \wedge (y \le a \cdot b' \le y + \mathrm{d}y))$, where a is as in the definition of I (see section 3.3). So what $I(v, w, x, x+\mathrm{d}x, y, y+\mathrm{d}y)$ effectively denotes is the density function of the joint probability distribution of $a \cdot b, a \cdot b'$ multiplied by $\mathrm{d}x\mathrm{d}y$. The expression we derive for $f_{e,k}(\delta)$ will have terms of the form $I(v, w, x, x+\mathrm{d}x, y, y+\mathrm{d}y)$. In section 6, we will expand this term out in terms of the actual density function.

Before giving the expressions for $f_{e,k}(\delta)$, we need to reiterate a basic fact.

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

The cases to be considered. There are three cases, depending upon whether $k < i$, $k = i$, or $k > i$. Each case has three subcases, depending upon whether $j < n-1$, $j = n-1$, or $j = n$. We have to consider these three subcases separately for the following reason: When $j < n-1$, we will express the above probability in terms of the function $I()$. 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 two-dimensional.) Therefore, in these two subcases, we have to express the probability directly. These two subcases themselves need to be separated because the derivative of $f_{e,k}(\delta)$ behaves differently in these two subcases, and the behavior is crucial to the analysis (setting up S in section 4). Recall Lemma 3.4 in this context.

Note from property 1 of section 3.1 that $v[n], w[n]$ are nonzero. We will need to divide by these quantities at points.

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

Case 1 ($k < 1$). In this case, the center t_k already has been 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

$$f_{e,k}(\delta) = \Pr(\wedge_{l=i}^{r}(t_l \cdot v \le \alpha \wedge t_l \cdot w \le \beta) | i, j, \delta).$$

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

Case 1.1 ($j < n-1$). By Fact 1,

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

Based on the following lemma, we claim that the derivative of $f_{e,k}(\delta)$ with respect to δ is always defined for the above case. The same will be true for Cases 2.1 and 3.1.

LEMMA 5.1. *$I(b, b', \mathcal{A}(\delta), \mathcal{B}(\delta), \mathcal{C}(\delta), \mathcal{D}(\delta))$ is differentiable with respect to δ for all δ, where $\mathcal{A}(), \mathcal{B}(), \mathcal{C}(), \mathcal{D}()$ are linear functions of δ.*

Proof. The proof of Lemma 5.1 will be described in section 6, after the expression for $I()$ is derived. ∎

Case 1.2 ($j = n-1$). We derive the expression for $f_{e,k}(\delta)$ for the case when $v[n]$ and $w[n]$ are both positive. The other cases are similar. By Fact 1 and the fact that $t_i[n]$ is normally distributed,

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

Note that the derivative of $f_{e,k}(\delta)$ 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, $f_{e,k}(\delta) = 0$. Otherwise

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

Note that the derivative of $f_{e,k}(\delta)$ 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 $\max\{t_1 \cdot w, \ldots, t_{i-1} \cdot w\} = \beta$. $t_k \cdot v$ must be greater than α and $t_k \cdot w$ must be greater than β for t_k to be the color assigned to both v and w. Then let A be the event $t_k \cdot v \ge \alpha \wedge t_k \cdot w \ge \beta$ and B_l be the event $t_l \cdot v \le t_k \cdot v \wedge t_l \cdot w \le t_k \cdot w$, $l \ge i, l \ne k$.

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. Note that we do not do any discretization or approximation here. Rather, what we derive here is an exact integral using the slightly nonstandard but intuitively illustrative notation $\Pr((x \le t_k \cdot v \le x + \mathrm{d}x) \wedge (y \le t_k \cdot w \le y + \mathrm{d}y))$ defined earlier in this section.

Case 2.1 ($j < n-1$).

$$\begin{aligned}
f_{e,k}(\delta) &= \Pr(A \wedge B_i \wedge \cdots B_{k-1} \wedge B_{k+1} \wedge \cdots B_r | i, j, \delta) \\
&= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (\Pr((x \le t_k \cdot v \le x + \mathrm{d}x) \wedge (y \le t_k \cdot w \le y + \mathrm{d}y)) \\
&\qquad \times \quad \Pr(t_i \cdot v \le x \wedge t_i \cdot w \le y | i, j, \delta)
\end{aligned}$$

$$\times \prod_{l=i+1,\dots,k-1,k+1,\dots,r} \Pr(t_l \cdot v \le x \wedge t_l \cdot w \le y | i, j, \delta)$$

$$= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y)$$
$$\times \Pr(\alpha' + \delta v[j] + t_i \cdot v[j+1, n] \le x \wedge \beta' + \delta w[j] + t_i \cdot w[j+1, n] \le y)$$
$$\times I^{r-i-1}(v, w, -\infty, x, -\infty, y))$$
$$= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y)$$
$$\times I(v[j+1\dots n], w[j+1\dots n], -\infty, x - \alpha' - v[j]\delta, -\infty, y - \beta' - w[j]\delta)$$
$$\times I^{r-i-1}(v, w, -\infty, x, -\infty, y)).$$

By Lemma 5.1, $f(\delta)$ is always differentiable with respect to δ in this case.

Case 2.2 ($j = n - 1$). Assume that $v[n]$ and $w[n]$ are positive. The remaining cases are similar. Then $f_{e,k}(\delta) = \Pr(A \wedge B_i \wedge \cdots \wedge B_{k-1} \wedge B_{k+1} \wedge \cdots \wedge B_r | i, n-1, \delta)$

$$= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (\Pr((x \le t_k \cdot v \le x + \mathrm{d}x) \wedge (y \le t_k \cdot w \le y + \mathrm{d}y))$$
$$\times \prod_{l=i,\dots,k-1,k+1,\dots,r} \Pr(t_l \cdot v \le x \wedge t_l \cdot w \le y | i, n-1, \delta))$$
$$= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y)$$
$$\times \Pr(\alpha' + \delta v[n-1] + t_i[n]v[n] \le x \wedge \beta' + \delta w[n-1] + t_i[n]w[n] \le y)$$
$$\times \prod_{l=i+1,\dots,k-1,k+1,\dots,r} \Pr(t_l \cdot v \le x \wedge t_l \cdot w \le y))$$
$$= \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} (I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) \left(\frac{1}{\sqrt{2\pi}} \int_{z=-\infty}^{\min\{\frac{x-\alpha'-v[n-1]\delta}{v[n]}, \frac{y-\beta'-w[n-1]\delta}{w[n]}\}} \mathrm{e}^{-z^2/2} \mathrm{d}z \right)$$
$$\times I^{r-i-1}(v, w, -\infty, x, -\infty, y))$$
$$= \frac{1}{\sqrt{2\pi}} \int_{z=-\infty}^{\infty} \int_{x=\max\{\alpha, \alpha'+v[n]z+v[n-1]\delta\}}^{\infty} \int_{y=\max\{\beta, \beta'+w[n]z+w[n-1]\delta\}}$$
$$(I(v, w, x, x + \mathrm{d}x, y, y + \mathrm{d}y) \times I^{r-i-1}(v, w, -\infty, x, -\infty, y)\mathrm{e}^{-z^2/2})\mathrm{d}z.$$

Note that the derivative of $f_{e,k}(\delta)$ 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 $f_{e,k}(\delta)$ 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 $f_{e,k}(\delta)$ is differentiable for all values of δ such that $\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]} > \frac{\beta-\beta'-w[n-1]\delta}{w[n]}$.

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

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

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. Then $f_{e,k}(\delta) = \Pr(A \wedge B_i \wedge \cdots, B_{k-1} \wedge B_{k+1} \wedge \cdots, B_r | i, n, \delta)$

$$= \int_{y=\max\{\beta,\beta'+w[n]\delta\}}^{\infty}\int_{x=\max\{\alpha,\alpha'+v[n]\delta\}}^{\infty}(\Pr((x \le t_k\cdot v \le x+\mathrm{d}x)\wedge(y \le t_k \cdot w \le y+\mathrm{d}y))$$
$$\times\prod_{l=i+1,\ldots,k-1,k+1,\ldots,r}\Pr(t_l\cdot v \le x \wedge t_l \cdot w \le y | i,n,\delta))$$
$$= \int_{y=\max\{\beta,\beta'+w[n]\delta\}}^{\infty}\int_{x=\max\{\alpha,\alpha'+v[n]\delta\}}^{\infty}(\Pr((x \le t_k\cdot v \le x+\mathrm{d}x)\wedge(y \le t_k \cdot w \le y+\mathrm{d}y))$$
$$\times\prod_{l=i+1,\ldots,k-1,k+1,\ldots,r}\Pr(t_l\cdot v \le x \wedge t_l \cdot w \le y))$$
$$= \int_{\max\{\beta,\beta'+w[n]\delta\}}^{\infty}\int_{\max\{\alpha,\alpha'+v[n]\delta\}}^{\infty}(I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y)$$
$$\times\, I^{r-i-1}(v,w,-\infty,x,-\infty,y)).$$

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 $t_i \cdot w$.

Case 3.1 ($j < n - 1$).

$$f_{e,k}(\delta) = \Pr(A \wedge B_{i+1} \wedge \cdots \wedge B_r | i, j, \delta)$$
$$= \int_{x=\alpha}^{\infty}\int_{y=\beta}^{\infty}(I(v[j+1\ldots n], w[j+1\ldots n], x-\alpha'-v[j]\delta, x+\mathrm{d}x-\alpha'-v[j]\delta,$$
$$y-\beta'-w[j]\delta, y+\mathrm{d}y-\beta'-w[j]\delta)$$
$$\times\, I^{r-i}(v,w,-\infty,x,-\infty,y)).$$

By Lemma 5.1, $f(\delta)$ is always differentiable with respect to δ in this case.

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

$$f_{e,k}(\delta) = \Pr(A \wedge B_{i+1} \wedge \cdots \wedge B_r | i, n-1, \delta)$$
$$= \frac{1}{\sqrt{2\pi}} \int_{z=\max\{\frac{\alpha-\alpha'-v[n-1]\delta}{v[n]}, \frac{\beta-\beta'-w[n-1]\delta}{w[n]}\}}^{\infty}$$
$$\times \, (I^{r-i}(v, w, -\infty, \alpha'+v[n-1]\delta+v[n]z, -\infty, \beta' + w[n-1]\delta + w[n]z)\mathrm{e}^{-z^2/2}\mathrm{d}z).$$

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,

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

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

6. Substep 2a: Expressing $\boldsymbol{I(b, b', x, y, x', y')}$. Recall that $I(b, b', x, y, x', y')$ denotes $\Pr((x \le a \cdot b \le y) \wedge (x' \le a \cdot b' \le y'))$, where a is a vector whose entries are independent and normally distributed with mean 0 and variance 1. We show how to derive an expression for this probability in terms of nested integrals of depth 2, with the integrand comprising only basic functions. Let b and b' be h-dimensional. Note that $h \ge 2$. Consider the h-dimensional coordinate system with respect to which b, b' are specified.

The naive way. Note that a naive way to compute I is to perform a sequence of h nested integrals, one over each of $a[1] \ldots a[h]$, the coordinates of the vector a. The following is the naive expression for the case when $b[h]$ and $b'[h]$ are both positive. A similar expression holds for other cases.

$$\left(\frac{1}{\sqrt{2\pi}}\right)^h \int_{a[1]=-\infty}^{\infty} \int_{a[2]=-\infty}^{\infty} \cdots \int_{a[h-1]=-\infty}^{\infty} \int_{a[h]=\alpha}^{\beta} \mathrm{e}^{-\frac{\sum_{l=1}^{h} a[l]^2}{2}} \mathrm{d}a[h]\mathrm{d}a[h-1] \cdots \mathrm{d}a[1],$$

where

$$\alpha = \max\left\{\left(x - \sum_{1}^{h-1} a[i]b[i]\right)/b[h], \left(x' - \sum_{1}^{h-1} a[i]b'[i]\right)/b'[h]\right\},$$

$$\beta = \min\left\{\left(y - \sum_{1}^{h-1} a[i]b[i]\right)/b[h], \left(y' - \sum_{1}^{h-1} a[i]b'[i]\right)/b'[h]\right\}.$$

Computing this integral to within the required error bounds in polynomial time seems hard. We use the following method instead.

Our method. 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 $b_1, b'_2 \ge 0$. As we will show shortly, both b_1, b'_2 will be strictly positive for all our calls to I. Let $a'[1]a'[2] \ldots a'[h]$ 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. Let $x_1, \ldots, x_h$ denote the initial coordinate system, and let $x'_1, x'_2, \ldots, x'_h$ denote the coordinate system after rotation. Then $(x_1, x_2, \ldots, x_h)A = (x'_1, x'_2, \ldots, x'_h)$, where A is the orthonormal rotation matrix; i.e., $AA^T = I$.

Next, recall that the probability density function of $a[l]$ is $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x_l^2}{2}}$ for $l = 1 \ldots h$. We show that the probability density function of $a'[l]$ is $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x'_l)^2}{2}}$ for $l = 1 \ldots h$. We also show that the joint probability density function of the $a'[l]$s is just the product of their individual density functions. The lemma then follows.

The density function of the joint distribution of $a[1], \ldots, a[h]$ is $\frac{1}{(\sqrt{2\pi})^h}\mathrm{e}^{-\frac{\Sigma_1^h x_l^2}{2}}$. We now derive the joint distribution of $a'[1], \ldots, a'[h]$. Since $(x_1, x_2, \ldots, x_h)A = (x'_1, x'_2, \ldots, x'_h)$ and A is orthonormal, $\sum_1^h x_l^2 = \sum_1^h (x'_l)^2$. Using the standard method for performing coordinate transformation, the density function of the joint distribution of $a'[1], \ldots, a'[h]$ is $\frac{1}{(\sqrt{2\pi})^h}\mathrm{e}^{-\frac{\Sigma_1^h (x'_l)^2}{2}}\det(B)$, where B is the matrix whose p, qth entry is $\frac{\partial x_p}{\partial x'_q}$. Since $(x_1, x_2, \ldots, x_h) = (x'_1, x'_2, \ldots, x'_h)A^T$, the matrix B is easily seen to be identical to A. Since A is orthonormal, $\det(A) = 1$, and therefore, the density function of the joint distribution of $a'[1], \ldots, a'[h]$ is just $\frac{1}{(\sqrt{2\pi})^h}\mathrm{e}^{-\frac{\Sigma_1^h (x'_l)^2}{2}}$.

Finally, the density function of $a'[l]$ can be seen to be $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x'_l)^2}{2}}$ by integrating away the other terms, i.e.,

$$\int_{x_1=-\infty}^{\infty}\int_{x_2=-\infty}^{\infty}\cdots\int_{x_{l-1}=-\infty}^{\infty}\int_{x_{l+1}=-\infty}^{\infty}\cdots\int_{x_h=-\infty}^{\infty}\left(\frac{1}{(\sqrt{2\pi})^h}\mathrm{e}^{-\frac{\sum_1^h (x'_l)^2}{2}}\right)$$
$$\times \quad \mathrm{d}x'_h \mathrm{d}x'_{h-1}\cdots\mathrm{d}x'_{l+1}\mathrm{d}x'_{l-1}\cdots\mathrm{d}x'_1 = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x'_l)^2}{2}}. \qquad \square$$

Having rotated the coordinate axes, 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 $\Pr((x \le a'[1]b_1 \le y) \wedge (x' \le a'[1]b'_1 + a'[2]b'_2 \le y'))$. We give the expression for $I(b, b', x, y, x', y')$ in the following lemma for future reference.

LEMMA 6.2.

$$I(b, b', x, y, x', y') = \Pr((x \le a'[1]b_1 \le y) \wedge (x' \le a'[1]b'_1 + a'[2]b'_2 \le y'))$$

$$= \Pr\left(\left(\frac{x}{b_1} \le a'[1] \le \frac{y}{b_1}\right) \wedge \left(\frac{x' - a'[1]b'_1}{b'_2} \le a'[2] \le \frac{y' - a'[1]b'_1}{b'_2}\right)\right)$$

$$= \frac{1}{2\pi}\int_{\frac{x}{b_1}}^{\frac{y}{b_1}} e^{-\frac{z^2}{2}}\left(\int_{\frac{(x'-zb'_1)}{b'_2}}^{\frac{(y'-zb'_1)}{b'_2}} e^{-\frac{z'^2}{2}} \quad \mathrm{d}z'\right) \quad \mathrm{d}z.$$

LEMMA 6.3. $|b_1| = \Omega(\epsilon) = \Omega(\frac{1}{n^2})$ *and* $|b'_2| = \Omega(\epsilon^2) = \Omega(\frac{1}{n^4})$.

Proof. Note that b, b' are of the form $v[h \ldots n], w[h \ldots n]$ for vertex vectors v and w and $h \le n-1$ in all the calls we make to $I()$. The lemma now follows from property 3 of the discretization of vertex vectors in section 3.1. $\square$

Two remarks. We remark that Lemma 5.1 can be easily seen to hold by inspecting the expression derived in Lemma 6.2, with x, x', y, y' replaced by linear functions of δ. Second, recall from section 5 that some of the expressions involve occurrences of $I()$ with $y = x + \mathrm{d}x$ and $y' = x' + \mathrm{d}x'$. It can be easily seen that

$$\Pr((x \leq a \cdot b \leq x + \mathrm{d}x) \wedge (x' \leq a \cdot b' \leq x' + \mathrm{d}x')) = \frac{1}{2\pi} \frac{1}{b_1 b_2'} \mathrm{e}^{-\frac{(x/b_1)^2}{2}} \mathrm{e}^{-\frac{(x' - x b_1'/b_1)^2}{2(b_2')^2}} \ \mathrm{d}x' \ \mathrm{d}x.$$

7. Substep 2b: Evaluating $f_{e,k}(\delta)$. Sections 5 and 6 derived expressions for $f_{e,k}(\delta)$, which were nested integrals with constant depth (at most five). We now show how to evaluate these expressions at any given value of δ in polynomial time with just $O(\frac{1}{n^5})$ error.

First, in Lemma 7.1, we show that all the expressions we derived in sections 5 and 6 involve integrations where the integrand has a particular form. This enables us to focus on only integrations of this form. We show how to perform each such integration within an inverse polynomial error, with the polynomial being chosen so that the final error in computing $f_{e,k}(\delta)$ is $O(\frac{1}{n^5})$ as required.

To perform each such integration, we have to do two things: first, restrict the range of the limits of integration and, second, convert the integration to a summation and compute the summation. Each of these steps will incur an error, which we will show can be made inverse polynomial, with the actual polynomial chosen to keep the overall error within the stated bounds. The error in restricting the range of the limits of integration is bounded in Lemma 7.2. To bound the error in converting the integrations to summations, we give Lemma 7.3, which states that it suffices to bound the absolute value of the derivative of the integrands. In Lemma 7.4, we show that the derivative of each integrand is bounded by some polynomial in n. Together Lemmas 7.1, 7.2, 7.3, and 7.4 imply that each integration can be computed to within an inverse polynomial error. Finally, since expressions for $f_{e,k}(\delta)$ involve up to five nested integrations, the inverse polynomial error terms in each integration have to be chosen so that the final combined error of all five integrations is $O(\frac{1}{n^5})$. This is described under the heading Algorithm for Performing Integrations. Lemma 7.1 obtains the general form of each integration.

LEMMA 7.1. *Each integration we perform can be expressed in the following form:*

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

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

$$\frac{1}{\sqrt{2\pi}} \frac{1}{b_1} \mathrm{e}^{-\frac{x^2}{2b_1^2}} \frac{1}{\sqrt{2\pi}} \frac{1}{b_2'} \mathrm{e}^{-\frac{\left(y - x\frac{b_1'}{b_1}\right)^2}{2(b_2')^2}} \mathrm{d}y \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{h^2}{2}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(h')^2}{2}} \mathrm{d}h' \mathrm{d}h,$$

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 $h' = \frac{y - x b_1'/b_1}{b_2'}$. This change of variables affects the limits of the integration, but we are not claiming any special properties for the limits.

Similarly,

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

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 $v[j+1\dots n] = (b_1, 0, \dots, 0)$ and $w[j+1\dots n] = (b_1', b_2', 0, \dots, 0)$ 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 $h' = \frac{y-\beta'-w[j]\delta-\frac{(x-\alpha'-v[j]\delta)b_1'}{b_1}}{b_2'}$. ∎

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

LEMMA 7.2.

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

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

for all $a > 0$.

Proof. The first inequality is obvious. The second is derived using Theorem 4.1 as follows:

$$\begin{aligned}
&\int_l^m \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h\\
&\le \int_{\max\{l,-a\sqrt{\ln n}\}}^{\min\{m,a\sqrt{\ln n}\}} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h + \int_{-\infty}^{-a\sqrt{\ln n}} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h\\
&\quad + \int_{a\sqrt{\ln n}}^{\infty} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h\\
&\le \int_{\max\{l,-a\sqrt{\ln n}\}}^{\min\{m,a\sqrt{\ln n}\}} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h + \int_{-\infty}^{-a\sqrt{\ln n}} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}}\right)\mathrm{d}h\\
&\quad + \int_{a\sqrt{\ln n}}^{\infty} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}}\right)\mathrm{d}h\\
&\le \int_{\max\{l,-a\sqrt{\ln n}\}}^{\min\{m,a\sqrt{\ln n}\}} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h + 2\int_{a\sqrt{\ln n}}^{\infty} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}}\right)\mathrm{d}h\\
&\le \int_{\max\{l,-a\sqrt{\ln n}\}}^{\min\{m,a\sqrt{\ln n}\}} \frac{1}{\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{h^2}{2}} H(\mathcal{G}(h))\right)\mathrm{d}h + O\left(\frac{1}{n^{a^2/2}}\right).
\end{aligned}$$

Theorem 4.1 is used in the last step above. The fact that $0 \le H() \le 1$ is used in the second step. ∎

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 7.3. *$|\int_l^{l+\rho} \frac{1}{\sqrt{2\pi}}(e^{-\frac{h^2}{2}} H(\mathcal{G}(h)))\mathrm{d}h - \frac{1}{\sqrt{2\pi}} e^{-\frac{l^2}{2}} H(\mathcal{G}(l))\rho| \le \frac{M\rho^2}{2}$, 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.*

LEMMA 7.4. *The derivative of $\frac{1}{\sqrt{2\pi}}(e^{-\frac{h^2}{2}} H(\mathcal{G}(h)))$ with respect to h is at most a polynomial in n in absolute value in all our integrations.*

Proof. The proof of Lemma 3 is found in Appendix 3. □

ALGORITHM FOR PERFORMING INTEGRATIONS. The four lemmas above 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 with

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

Here a will be a constant to be fixed later. Next, we convert this integral to a sum by dividing the range between the limits of integration into steps of size $\Theta(\frac{1}{n^b})$ 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 $O(n^c)$. We compute the total error incurred above.

By Lemma 7.2, clipping the limits of integration incurs an error of $O(\frac{1}{n^{a^2/2}})$. By Lemma 7.3, the error incurred in each step of the summation is $O(\frac{n^c}{n^{2b}})$, 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 $O(\frac{n^c}{n^{2b}} + \frac{1}{n^{f-1}n^b})$; this sums to $O(\sqrt{\ln n}\frac{n^c}{n^b} + \frac{\sqrt{\ln n}}{n^{f-1}})$ over all $O(\sqrt{\ln n}n^b)$ steps. The total error is thus $O(\frac{1}{n^{a^2/2}} + \sqrt{\ln n}\frac{n^c}{n^b} + \frac{\sqrt{\ln n}}{n^{f-1}})$ and the time taken for this integration (ignoring the time taken for the nested integrals in $H()$) is $O(\sqrt{\ln n}n^b)$.

Finally, note that the depth of nesting in our integrals is at most five (in Case 2.2 of section 5, it is five). It can be easily seen that starting with the innermost integral and working outward, values a, b can be chosen for these successive integrals based upon the respective c, f values so that the final error is $O(\frac{1}{n^5})$.

8. Comments on derandomizing the Max-Cut algorithm. We need to show that discretizing the vectors in this case so as to satisfy properties 1–3 (with ϵ chosen appropriately) of section 3.1 is justified. In the case of the Karger–Motwani–Sudan algorithm, it was justified using Theorem 2.1.

To compensate for this theorem we need only 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 $(1 - \frac{1}{poly(n)})$ times that for the initial vector set (this is because the sum of the edge 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.

9. 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 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.

Also, it would be nice to obtain a more efficient derandomization scheme, since the running time of our algorithm is a large polynomial, around $O(n^{30})$ for the 3-coloring problem.

Appendix 1. Discretizing vertex vectors. Let ϵ be a parameter which is $\Theta(\frac{1}{n^2})$. In this section, we show how to discretize the vertex vectors so as to satisfy the three properties specified in section 3.1.

The vertex vectors are considered one by one in the order $v_1, v_2, \ldots, v_n$. We describe the processing of vector v_i.

First, each entry in v_i is rounded upward (in absolute value) to the nearest nonzero multiple of ϵ. Next, up to $2n\epsilon$ is added to $v_i[n-1]$ so that $|v_i[n-1]v_j[n] - v_i[n]v_j[n-1]| > \epsilon^2$ for every $j < i$. This is done to satisfy property 3. Property 1 is clearly satisfied for v_i now. Property 2 is also satisfied because each component of v_i other than $v_i[n-1]$ changes by $O(\epsilon)$ and $v_i[n-1]$ changes by $O(n\epsilon)$. However, note that in this process the vertex vectors no longer remain unit vectors. In fact, $1 \le |v_i|^2 \le 2$ now, for small enough ϵ, i.e., for large enough n. So we divide each vector v_i by its new norm and make it a unit vector. Since we divide by a number between one and two, property 1 and property 2 continue to hold.

It remains to show that property 3 holds. We need the following lemma.

LEMMA 9.1. *For each pair of vertex vectors v, w and every h, $1 \le h < n$, 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 (more precisely, $\Omega(\epsilon)$ and $\Omega(\epsilon^2)$, respectively) in absolute value.*

Proof. Let $v' = v[h \ldots n]$ and $w' = w[h \ldots n]$. Note that b_1 is just the norm of v' which is $\Omega(\epsilon)$ by property 1. Also note that $|b'_2| = \sqrt{||w'|^2 - \frac{(v' \cdot w')^2}{|v'|^2}|}$, since b'_2 is just the projection of w' on the line orthogonal to v' in the plane containing v' and w'. So we need to show that $||w'|^2 - \frac{(v' \cdot w')^2}{|v'|^2}| = \Omega(\epsilon^4)$ for every h $1 \le h < n$.

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 $l = v[h \ldots n-2]$ and $m = w[h \ldots n-2]$. Let $l' = v[n-1, n]$ and $m' = w[n-1, n]$; $(v' \cdot w')^2 = (l \cdot m + l' \cdot m')^2 = (l \cdot m)^2 + (l' \cdot m')^2 + 2(l' \cdot m')(l \cdot m) \le |l|^2|m|^2 + (l' \cdot m')^2 + 2|l'||m'||l||m|$. 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, $||w'|^2 - \frac{(v' \cdot w')^2}{|v'|^2}| = \Omega(\epsilon^4)$. ∎

Appendix 2. Proof of Lemma 4.3. We show that for each edge e and each center k, the derivative of $f_{e,k}(\delta)$ (with respect to δ) is $O(n^4)$.

Recall from section 5 that the expression for $f_{e,k}(\delta)$ depends upon which one of Cases 1, 2, and 3 and which one of the conditions $j < n-1, j = n-1, j = n$ hold.

We show the above claim only for one representative case, i.e., Case 2.1, where $j < n-1$. The other cases can be shown similarly. For Case 2.1

$$f_{e,k}(\delta) = \int_{\alpha}^{\infty} \int_{\beta}^{\infty} g(x,y)h(x,y,\delta)\mathrm{d}y\mathrm{d}x,$$

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

$$\begin{aligned}h(x,y,\delta) =& I(v[j+1\dots n],w[j+1\dots n],-\infty,x-t_i\cdot v[1\dots j-1]-v[j]\delta,-\infty,\\ & y-t_i\cdot w[1\dots j-1]-w[j]\delta).\end{aligned}$$

Now

$$|f'_{e,k}(\delta)| \le \int_\alpha^\infty \int_\beta^\infty |g(x,y)| \left|\frac{\partial h(x,y,\delta)}{\partial\delta}\right| \mathrm{d}y\mathrm{d}x \le \max_{x,y}\left|\frac{\partial h(x,y,\delta)}{\partial\delta}\right|$$

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 $|f'_{e,k}(\delta)| = O(n^4)$ by estimating $\max_{x,y}|\frac{\partial h(x,y,\delta)}{\partial\delta}|$. Let $c(x,\delta) = x - t_i\cdot v[1\dots j-1]-v[j]\delta = c' - v[j]\delta$ and $d(y,\delta) = y - t_i\cdot w[1\dots j-1]-w[j]\delta = d'-w[j]\delta$.

By Lemma 6.2, $h(x,y,\delta) = \frac{1}{2\pi}\int_{-\infty}^{c(x,\delta)/b_1} \mathrm{e}^{-\frac{z^2}{2}}(\int_{-\infty}^{(d(y,\delta)-zb'_1)/b'_2}\mathrm{e}^{-\frac{z'^2}{2}}\ \mathrm{d}z')\mathrm{d}z$, where b_1, b'_1, b'_2 are 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 $\frac{\partial h}{\partial\delta}$ can be expressed as $A+B$, 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

$$B = \frac{\partial G}{\partial l}_{l=c(x,\delta)/b_1}\frac{\partial l}{\partial\delta}_{l=c(x,\delta)/b_1}.$$

Therefore, $|\frac{\partial h}{\partial\delta}| = |A+B| \le |A|+|B|$. Next, we bound $|A|$ and $|B|$ separately.

$|B|$ is bounded as follows: Note that $|\frac{\partial G}{\partial l}| = |\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{l^2}{2}}H(y,\delta,l)| \le 1$ for all l, since $0 \le H(y,\delta,l)\le 1$. Further, $|\frac{\partial l}{\partial\delta}_{l=c(x,\delta)/b_1}| = |v[j]/b_1| = O(n^2)$, by Lemma 6.3. Therefore, $|B| = O(n^2)$.

$|A|$ is bounded as follows: $|A| = |\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c(x,\delta)/b_1}\mathrm{e}^{-\frac{l^2}{2}}\frac{\partial H(y,\delta,l)}{\partial\delta}\mathrm{d}l| \le \max_{y,\delta,l}|\frac{\partial H(y,\delta,l)}{\partial\delta}|$. 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 $|\frac{\partial H}{\partial\delta}| = |\frac{\mathrm{d}J}{\mathrm{d}m}_{m=(d(y,\delta)-lb'_1)/b'_2}||\frac{\partial m}{\partial\delta}_{m=(d(y,\delta)-lb'_1)/b'_2}| \le |w[j]/b'_2| = O(n^4)$, by Lemma 6.3. Therefore, $|f'_{e,k}(\delta)| \le |A|+|B| = O(n^4)$.

Appendix 3. Proof of Lemma 7.4.

Bounding derivatives of integrands in $I()$. Recall that

$$I(b,b',x,y,x',y') = \frac{1}{\sqrt{2\pi}} \int_{\frac{x}{b_1}}^{\frac{y}{b_1}} \mathrm{e}^{\frac{-z^2}{2}} \left(\int_{\frac{(x'-zb'_1)}{b'_2}}^{\frac{(y'-zb'_1)}{b'_2}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\frac{-z'^2}{2}} \mathrm{d}z' \right) \mathrm{d}z.$$

Here b, b' have been rotated so that $b = (b_1, 0, \ldots, 0)$ and $b' = (b'_1, b'_2, 0, \ldots, 0)$.

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 we compute the derivative of the function to be integrated; that is, $\frac{1}{\sqrt{2\pi}} h(z) \mathrm{e}^{\frac{-z^2}{2}}$ is

$$\frac{1}{\sqrt{2\pi}} \left(-z\mathrm{e}^{\frac{-z^2}{2}} h(z) + \frac{1}{2\pi} \mathrm{e}^{\frac{-z^2}{2}} (-b'_1/b'_2) \left(\mathrm{e}^{\frac{-1}{2}\left(\frac{y'-zb'_1}{b'_2}\right)^2} - \mathrm{e}^{-\frac{1}{2}\left(\frac{x'-zb'_1}{b'_2}\right)^2} \right) \right).$$

The first term in this sum is bounded in absolute value by a constant as $h(z) \leq 1$, and the second term is bounded by $O(n^4)$ by Lemma 6.3. Hence the derivative is bounded by $O(n^4)$.

Bounding 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 $f_{e,k}(\delta)$ 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

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

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 $(b_1, 0, \ldots, 0)$ and the new coordinates of w are $(b'_1, b'_2, 0, \ldots, 0)$, where $b_1, b'_2 \geq 0$. Recall from section 6 that

$$I(v,w,x,x+\mathrm{d}x,y,y+\mathrm{d}y) = \frac{1}{\sqrt{2\pi}} \frac{1}{b_1} \mathrm{e}^{\frac{-x^2}{2{b_1}^2}} \frac{1}{\sqrt{2\pi}} \frac{1}{b'_2} \mathrm{e}^{-\frac{1}{2}\left(\frac{y-\frac{x}{b_1}b'_1}{b'_2}\right)^2} \mathrm{d}y\mathrm{d}x.$$

Therefore,

$$\begin{aligned} g(\delta) = {} & \int_{z=-\infty}^{c} \int_{x=\alpha}^{\infty} \int_{y=\beta}^{\infty} I^{r-i-1}(v,w,-\infty,x,-\infty,y) \tfrac{1}{(2\pi)b_1 b'_2} \mathrm{e}^{\frac{-x^2}{2{b_1}^2}} \mathrm{e}^{-\frac{1}{2}\left(\frac{y-\frac{x}{b_1}b'_1}{b'_2}\right)^2} \\ & \times \tfrac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2} \mathrm{d}y\mathrm{d}x\mathrm{d}z. \end{aligned}$$

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

$$I^{r-i-1}(v,w,-\infty,x,-\infty,y) \frac{1}{\sqrt{2\pi} b'_2} \mathrm{e}^{-\frac{1}{2}\left(\frac{y-\frac{x}{b_1}b'_1}{b'_2}\right)^2}.$$

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

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

Now

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

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 $O\left(\frac{(r-i-1)}{(b_2')^2}+\frac{1}{(b_2')^2}\right)$. Since $r-i-1\le n$, $b_1=\Omega(\frac{1}{n^2})$, b_2' is $\Omega(\frac{1}{n^4})$ by Lemma 6.3, the above derivative is bounded by $O(n^9)$.

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

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

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

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

Here,

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

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

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

the above derivative is bounded by $O(\frac{n}{{b_1}^2} + \frac{1}{{b_2'}^2 b_1^2}) = O(n^{12})$, by Lemma 6.3.

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 $O(1)$.

Acknowledgments. We thank Naveen Garg, Kurt Mehlhorn, David Williamson, Michel Goemans, and Madhu Sudan for comments. We thank Aravind Srinivasan for reading a draft of the manuscript and for his detailed comments. We thank the anonymous referees for many comments which have improved the readability of this paper and also for pointing us to the tighter tail bounds for the normal distribution, which improves the running time of some parts of the algorithm.

REFERENCES

[1] N. Alon and N. Kahale, *Approximating the independence number via the θ-function*, Math. Programming, 80 (1998), pp. 253–264.

[2] N. Alon, P. Kelsen, S. Mahajan, and H. Ramesh, *Approximately coloring hypergraphs*, Nordic J. Comput. 3 (1996), pp. 425–439.

[3] A. Blum, *New approximation algorithms for graph coloring*, J. Assoc. Comput. Mach., 41 (1994), pp. 470–516.

[4] R. Boppana and M. Halldorsson, *Approximating maximum independent sets by excluding subgraphs*, BIT, 32 (1992), pp. 180–196.

[5] W. Feller, *Introduction to Probability Theory and its Applications*, Vol. 1, John Wiley, New York, 1966.

[6] U. Feige and M. Goemans, *Approximating the value of two prover proof systems, with applications to Max-*2*Sat and Max- Dicut*, Proc. 3rd Israeli Symposium on Theory of Computing and Systems, Tel Aviv, 1995, pp. 182–189.

[7] A. Frieze and M. Jerrum, *Improved approximation algorithms for Max k-Cut and Max Bisection*, Integer Programming and Combinatorial Optimization, Lecture Notes in Comput. Sci., Springer-Verlag, Berlin, 1995, pp. 1115–1165.

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

[9] M. Goemans and D. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach., 42 (1996), pp. 1115–1145.

[10] M. Halldorsson, *A still better performance guarantee for approximate graph colouring*, Inform. Process. Lett. 45 (1993), pp. 19–23.

[11] D.S. Johnson, *Worst case behaviour of graph coloring algorithms*, in Proc. 5th South-Eastern Conference on Combinatorics, Graph Theory and Computing, Congressus Numeratium X, (1974), pp. 513–527.

[12] D. Karger, R. Motwani, and M. Sudan, *Approximate graph coloring by semidefinite programming*, Proc. 35th IEEE Symposium on Foundations of Computer Science, Santa Fe, NM, 1994, pp. 1–10.

[13] S. Khanna, N. Linial, and S. Safra, *On the hardness of approximating the chromatic number*, Proc. 2nd Israeli Symposium on Theory and Computing Systems, 1993, pp. 250–260.

[14] P. Raghavan, *Probabilistic construction of deterministic algorithms: Approximating packing integer programs*, J. Comput. System Sci., 37 (1988), pp. 130–143.

[15] P. Raghavan, *Randomized Approximation algorithms in combinatorial optimization*, Foundations of Software Technology and Theoretical Computer Science, Lecture Notes in Comput. Sci., Springer-Verlag, New York, 1994, pp. 300–317.

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

[17] A. Wigderson, *Improving the performance guarantee of approximate graph coloring*, J. Assoc. Comput. Mach., 30 (1983), pp. 729–735.

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