On the Mean Radius of Permutation Polytopes

ALEXANDER BARVINOK AND RAY ROBB

Abstract. Let $X \subset S_n$ be a subset of the symmetric group S_n and let $Q_X \subset \mathbb{R}^{n^2}$ be the convex hull of the set of permutation matrices representing the elements of X in $\mathbb{R}^{n^2}$. We present some formulas relating the cardinality of X and the maximal value of a typical linear function on Q_X. Applications to the average case analysis of Combinatorial Optimization problems and to efficient counting are discussed.

1. Introduction

In this paper we address the question of how combinatorially interesting polytopes, such as the Birkhoff Polytope and the Traveling Salesman Polytope look like from the Banach Geometry point of view. Connections between optimization problems and the combinatorial structure of underlying polytopes have long been studied (see, for example, [3]). In [1] the first author considered the problem of optimizing an "average" linear function on a given polytope. It was shown that the famous "measure concentration phenomenon" (see [5]) implies that as the dimension grows the optimal value of almost any linear function sharply concentrates around certain "most popular" value. Some bounds for this value for the Birkhoff, the Traveling Salesman and other related polytopes were provided. In this paper we give an asymptotic formula for this most popular optimal value using the technique developed in [1]. Furthermore, we show that the optimal value of a typical optimization problem on a set of permutations roughly depends on the cardinality of the set alone. This observation suggests a new approach to efficient counting: a rough estimate of the cardinality of a set can be derived from the optimal value in a typical optimization problem on the set.

(1.1) Definitions and notation. Let S_n be the symmetric group, that is the set of permutations of $\{1, \ldots, n\}$. For a $\sigma \in S_n$ we denote by $\pi(\sigma)$ the permutation matrix

$$\pi_{ij}(\sigma) = \begin{cases} 1 & \text{if } \sigma(i) = j, \\ 0 & \text{if } \sigma(i) \neq j. \end{cases}$$

We consider $\pi(\sigma)$ as a point in $\mathbb{R}^{n^2}$. Our main object is the convex hull

$$Q_X = \operatorname{conv}\{\pi(\sigma) : \sigma \in X\},$$

where $X \subset S_n$ is a subset. For example, if $X = S_n$ then Q_X is the Birkhoff Polytope and if X is the subset consisting of $(n-1)!$ one-cycle permutations then Q_X is the Traveling Salesman Polytope (see, for example [2]).

We denote by $\langle \cdot, \cdot \rangle$ the standard scalar product in Euclidean space $\mathbb{R}^d$. Let $\| \cdot \|$ be the corresponding Euclidean norm in $\mathbb{R}^d$ and let

$$S^{d-1} = \Big\{ u \in \mathbb{R}^d : \|u\| = 1 \Big\}$$

be the unit sphere. We consider the (unique) rotation invariant Borel probability measure $\mu = du$ on S^{d-1}. With a polytope $P \subset \mathbb{R}^d$ we associate its support function

$$h_P : S^{d-1} \longrightarrow \mathbb{R}, \qquad h_P(u) = \max\big\{ \langle u, x \rangle : \ x \in P \big\}.$$

Let $\rho(P)$ denote the median of h_P, that is the unique number such that

$$\mu\Big\{ u \in S^{d-1} : \ h_P(u) \le \rho(P) \Big\} \ge \frac{1}{2} \quad \text{and} \quad \mu\Big\{ u \in S^{d-1} : \ h_P(u) \ge \rho(P) \Big\} \ge \frac{1}{2}.$$

We identify $\mathbb{R}^{n^2}$ with the space of $n \times n$ matrices. In particular,

$$\langle a, b \rangle = \sum_{i,j=1}^{n} a_{ij} b_{ij} \qquad \text{for matrices} \quad a = (a_{ij}), \quad b = (b_{ij}) \in \mathbb{R}^{n^2}.$$

Thus for $X \subset \mathbb{R}^n$ we have

$$h_{Q_X}(u) = \max\Big\{ \sum_{i=1}^{n} u_{i\sigma(i)} : \ \sigma \in X \Big\}, \qquad \text{where} \quad u = (u_{ij}) \in \mathbb{R}^{n^2}.$$

We denote by $|X|$ the cardinality of a finite set X.

In this paper we prove the following main result.

(1.2) Theorem. *Let $X \subset S_n$ be a subset of S_n and let $Q_X = \mathrm{conv}\{\pi(\sigma) : \sigma \in X\}$ be the convex hull of permutation matrices for the permutations from X.*
(1.2.1) *Let us define*

$$\delta(X) = 1 - \frac{\ln |X|}{n \ln n}.$$

Then

$$1 - \sqrt{\delta(X)} + o(1) \le \frac{\rho(Q_X)}{\sqrt{2 \ln n}} \le \sqrt{1 - \delta(X)} + o(1) \qquad \text{as} \quad n \longrightarrow +\infty.$$

(1.2.2) *We have*

$$\mu\Big\{ u \in S^{n^2-1} : \big| h_{Q_X}(u) - \rho(Q_X) \big| > \epsilon \Big\} \le 2 \exp\Big\{ -\frac{\epsilon^2 (n^2 - 2)}{2n} \Big\}$$

for any $\epsilon > 0$ and any n.

Thus (1.2.2) asserts that the maximal value of the linear function $\langle u, \cdot \rangle$ on Q_X for a "typical" unit vector u is very close to the median $\rho(Q_X)$ whereas (1.2.1) provides some bounds for this median in terms of the cardinality of X.

(1.3) Examples. Let us choose a sequence $X_n = S_n$, so $\lim\limits_{n\longrightarrow+\infty} \delta(X_n) = 0$ in (1.2.1). The problem

$$\text{Given } c = (c_{ij}) \in \mathbb{R}^{n^2}, \text{ find } \sigma \in S_n \text{ which maximizes } \sum_{i=1}^{n} c_{i\sigma(i)}$$

is known as the *Assignment Problem* (see [6]). Theorem 1.2 implies that if the entries c_{ij} are chosen independently from the standard normal distribution (that is, $c/\|c\|$ is chosen from the uniform distribution on the unit sphere in $\mathbb{R}^{n^2}$) then the optimal value is $\sqrt{2\ln n}\|c\|\big(1+o(1)\big)$ with the probability which tends to 1 as n grows.

Let us choose X_n to be the set of all one-cycle permutations $1 \to i_1 \to \ldots \to i_{n-1} \to 1$. Thus $|X_n| = (n-1)!$, so $\lim\limits_{n\longrightarrow+\infty} \delta(X_n) =$ in (1.2.1). The problem

$$\text{Given } c = (c_{ij}) \in \mathbb{R}^{n^2}, \text{ find } \sigma \in X_n \text{ which maximizes } \sum_{i=1}^{n} c_{i\sigma(i)}$$

is known as the *Traveling Salesman Problem* (see [6]). Theorem 1.2 implies that if the entries c_{ij} are chosen independently from the standard normal distribution then the optimal value is $\sqrt{2\ln n}\|c\|\big(1+o(1)\big)$ with the probability which tends to 1 as n grows.

It is interesting to note that the average case behavior of the polynomially solvable Assignment Problem and NP-hard Traveling Salesman Problem is the same. With respect to a typical linear function the Assignment (Birkhoff) Polytope and the Traveling Salesman Polytope both look like a ball of a radius about $\sqrt{2\ln n}$. Some ramifications of these examples will be discussed in Section 6.

The paper is organized as follows. In Section 2 we review some basic facts about measure concentration on the sphere S^{d-1}. In Section 3 we review the technique from [1] which allows us to compute $\rho(Q)$ from $\rho(P)$ where Q is a subpolytope of P and P has a large symmetry group. This allows us to relate $\rho(Q_X)$ and $\rho(P_n)$, where $P_n \subset \mathbb{R}^{n^2}$ is the Minkowski sum of n regular $(n-1)$-dimensional coordinate simplices. In Section 4 we estimate $\rho(\Delta)$ for the coordinate simplex Δ in $\mathbb{R}^n$. In Section 5 we complete the proof of Theorem 1.2. Finally, in Section 6 we discuss connections between approximate counting and random optimization.

2. Measure Concentration

We define an inner metric on S^{d-1} by $\operatorname{dist}(x,y) = \arccos\langle x,y\rangle$ for $x,y \in S^{d-1}$. In other words, $0 \le \operatorname{dist}(x,y) \le \pi$ is the angle between x and y.

Let $A \subset S^{d-1}$ be a closed set and let $u \in S^{d-1}$ be a point. Let us denote

$$\operatorname{dist}(u,A) = \min\big\{\operatorname{dist}(u,x) : x \in A\big\}$$

the distance from u to A. For an $r > 0$ we denote

$$A(r) = \big\{u \in S^{d-1} : \operatorname{dist}(u,A) \le r\big\}$$

the r-neighborhood of A. For $y \in S^{d-1}$ and $r > 0$ we denote by

$$C(y,r) = \big\{u \in S^{d-1} : \operatorname{dist}(y,u) \le r\big\}$$

the spherical cap centered in y and of the radius r. In particular, $C(y,\pi) = S^{d-1}$ and $C(y,\pi/2)$ is a hemisphere. We are going to use the following fact, known as the isoperimetric inequality on the sphere (see Section 2 of [5]).

(2.1) Theorem. *Let $A \subset S^{d-1}$ be a closed set and let $C(x,r) \subset S^{d-1}$ be a spherical cap such that $\mu\big(C(x,r)\big) = \mu(A)$. Then for any $\epsilon > 0$ one has $\mu\big(A(\epsilon)\big) \ge \mu\big(C(x,r+\epsilon)\big)$.*

□

Next, we are going to use some estimates of the spherical volumes (see Section 2 of [5]).

(2.2) Lemma. *Let $C(x,\pi/2)$ be a hemisphere in S^{d-1}. Then*

$$\mu\Big(C(x,\pi/2+\epsilon)\Big) \ge 1 - \sqrt{\frac{\pi}{8}}\exp\Big\{-\frac{\epsilon^2(d-2)}{2}\Big\};$$

$$\mu\Big(C(x,\pi/2-\epsilon)\Big) \le \sqrt{\frac{\pi}{8}}\exp\Big\{-\frac{\epsilon^2(d-2)}{2}\Big\}$$

for any $\epsilon > 0$.

□

Theorem 2.1 and Lemma 2.2 imply the famous Levy's Lemma (see Section 2 of [5]).

(2.3) Corollary. *Let $f : S^{d-1} \longrightarrow \mathbb{R}$ be a continuous function and M_f be its median. Let $A = \big\{x \in S^{d-1} : f(x) = M_f\big\}$. Then for any $\epsilon > 0$ one has*

$$\mu\big(A(\epsilon)\big) \ge 1 - \sqrt{\frac{\pi}{2}}\exp\Big\{-\frac{\epsilon^2(d-2)}{2}\Big\}.$$

□

Corollary 2.3 applied to the support function of a polytope gives us the following result.

(2.4) Theorem. *Let $P \subset \mathbb{R}^d$ be a polytope with v vertices such that $\|x\| \le R$ for some R and any $x \in P$. Then*

$$\mu\Big\{u \in S^{d-1} : |h_P(u) - \rho(P)| > \epsilon\Big\} \le \sqrt{\frac{\pi}{2}}\exp\Big\{-\frac{\epsilon^2(d-2)}{2R^2}\Big\}; \tag{2.4.1}$$

$$\Big|\int_{S^{d-1}} h_P(u)\,du - \rho(P)\Big| \le R\frac{\sqrt{\ln(d-2)}+2}{\sqrt{d-2}}; \tag{2.4.2}$$

(2.4.3) $$\rho(P) \le R\sqrt{\frac{2\ln 2v}{d-2}}.$$

Proof. First, we observe that $\|h_P(x) - h_P(y)\| \le R\|x-y\|$ for any $x, y \in \mathbb{R}^d$. Let $A = \{x \in S^{d-1} : h_P(x) = \rho(P)\}$. If $u \in A(\epsilon/R)$ then for some $x \in A$ we have $\|x-u\| \le \operatorname{dist}(x,u) \le \epsilon/R$ and thus $|h_P(u) - \rho(P)| \le \epsilon$. Now (2.4.1) follows by Corollary 2.3.

Let us choose $\epsilon = R\sqrt{\dfrac{\ln(d-2)}{d-2}}$ in (2.4.1). Then $|h_P(u) - \rho(P)| > \epsilon$ on the set of measure not greater than $\dfrac{2}{\sqrt{d-2}}$. Since $|h_P(u)| \le R$ for any $u \in S^{d-1}$ and $\rho(P) \ge 0$, the proof of (2.4.2) follows.

Let us choose $\epsilon = \sqrt{\dfrac{2\ln 2v}{d-2}}$ and $r = \dfrac{\pi}{2} - \epsilon$. Let A be the set of radial projections of non-zero vertices of P onto the unit sphere S^{d-1}. The set $A(r)$ is a union of at most v spherical caps of the radius r and hence by Lemma 2.2 we get

$$\mu\big(A(r)\big) \le v\exp\Big\{-\frac{\epsilon^2(d-2)}{2}\Big\} \le \frac{1}{2}.$$

For each $u \in S^{d-1} \setminus A(r)$ and each vertex x of P we have $\langle u, x\rangle \le R\cos\delta = R\sin\epsilon \le R\epsilon$. Since $h_P(u) = \langle u, x\rangle$ for some vertex x of P we conclude that $h_P(u) \le R\epsilon$ for at least half of $u \in S^{d-1}$, so (2.4.3) follows. □

3. Group Action

In this section we review some results from [1]. For the sake of completeness we present full proofs here.

(3.1) Lemma. *Let $V \subset S^{d-1}$ be a finite set. Suppose that for any two points $x, y \in V$ there exists an isometry $g_{x,y}$ of the sphere S^{d-1} such that $g_{x,y}(V) = V$ and $g_{x,y}(x) = y$. Then for any subset $U \subset V$ and any $r \ge 0$ one has*

$$\mu\Big(U(r)\Big) \ge \frac{|U|}{|V|}\mu\Big(V(r)\Big).$$

Proof. We use Voronoi diagrams on the sphere. For every point $x \in V$ let us define the Voronoi cell:

$$K(x,r) = \Big\{u \in V(r) : \ \operatorname{dist}(u,x) \le \operatorname{dist}(u,y) \text{ for any } y \in V\Big\}.$$

It is seen that $V(r)$ is represented as a union of closed subsets $\{K(x,r) : x \in V\}$ with pairwise disjoint interiors. Furthermore, for any two $x, y \in V$ the isometry $g_{x,y}$ maps $K(x,r)$

onto $K(y,r)$. Therefore $\mu\big(K(x,r)\big) = \mu\big(K(y,r)\big)$ and hence $\mu\big(K(x,r)\big) = \mu\big(V(r)\big)/|V|$ for each $x \in V$. Since $U(r) \subset \bigcup_{y\in U} K(y,r)$ we get that

$$\mu\big(U(r)\big) \geq \sum_{y\in U} \mu\big(K(y,r)\big) = \frac{|U|}{|V|}\mu\big(V(r)\big).$$

□

(3.2) Theorem. *Let $P \subset \mathbb{R}^d$ be a polytope with the vertex set V. Suppose that for any two vertices $x, y \in V$ there is an orthogonal transformation $g_{x,y}$ of $\mathbb{R}^d$ such that $g_{x,y}(V) = V$ and $g_{x,y}(x) = y$. Let $U \subset V$ and $Q = \mathrm{conv}\{U\} \subset P$ be the polytope with the vertex set U. Then for $\alpha = |V|/|U|$ and $R = \max\{\|x\| : x \in P\}$ one has*

$$\rho(P) - R\sqrt{\frac{2\ln 2\alpha}{d-2}} \leq \rho(Q) \leq \rho(P).$$

Proof. Since $h_Q(u) \leq h_P(u)$ for any $u \in S^{d-1}$, we have $\rho(Q) \leq \rho(P)$. Without loss of generality we may assume that $U \subset V \subset S^{d-1}$ and thus $\rho(P) \leq R = 1$. Let $r = \arccos \rho(P)$. Then $\mu\Big(V(r)\Big) = \dfrac{1}{2}$ and $h_P(u) \geq \rho(P)$ for any $u \in V(r)$. By Lemma 3.1 we get that $\mu\Big(U(r)\Big) \geq \dfrac{|U|}{2|V|} = \dfrac{1}{2\alpha}$. Let $C(y, \pi/2 - \epsilon)$ be a spherical cap such that $\mu\big(C(y,\pi/2-\epsilon)\big) = \mu\big(U(r)\big)$. Applying Lemma 2.2 we get that

$$\frac{1}{2\alpha} \leq \sqrt{\frac{\pi}{8}}\exp\Big\{-\frac{(d-2)\epsilon^2}{2}\Big\} \qquad \text{and hence} \qquad \epsilon \leq \sqrt{\frac{2\ln 2\alpha}{d-2}}.$$

By Theorem 2.1 $\mu\Big(U(r+\epsilon)\Big) \geq \mu\Big(C(y, \dfrac{\pi}{2})\Big) = \dfrac{1}{2}$. Furthermore, for every $u \in U(r+\epsilon)$ we have that $\mathrm{dist}\big(u, U(r)\big) \leq \epsilon$ and therefore $h_Q(u) \geq \rho(P) - \epsilon$ since the Lipschitz constant of h_Q does not exceed $R = 1$. Hence $h_Q(u) \geq \rho(P) - \epsilon$ for at least half of $u \in S^{d-1}$. Therefore $\rho(Q) \geq \rho(P) - \epsilon$ and the proof follows. □

4. Coordinate Simplex

In this section we estimate the average value of the support function for the coordinate simplex. Using such simplices as building blocks, we'll prove our main result in the next section.

(4.1) Theorem. *Let* $f(x) = \max\{x_1, \ldots, x_n\}$ *for* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. *Then*

$$\int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\ dx = \sqrt{\frac{\ln n}{\pi}}\Big(1 + o(1)\Big)$$

as $n \longrightarrow +\infty$.

Proof. Let us choose any $\epsilon > 0$. We are going to prove that for all sufficiently large n one has

$$(1+\epsilon)^2\sqrt{\frac{\ln n}{\pi}} \geq \int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\ dx \geq \frac{1}{(1+\epsilon)^2}\sqrt{\frac{\ln n}{\pi}}. \tag{4.1.1}$$

For a Borel set $X \subset \mathbb{R}^n$ we let $\nu(X) = \int_X e^{-\pi\|x\|^2}\ dx$. Clearly, ν is a probability measure on $\mathbb{R}^n$.

Let us prove the lower bound of (4.1.1) first. We observe that $f(x) < 0$ if and only if x belongs to the negative orthant $\mathbb{R}^n_- = \big\{(x_1, \ldots, x_n) : x_i < 0 :\ i = 1, \ldots, n\big\}$. Furthermore, the contribution of the negative part of f is asymptotically negligible:

$$\Big|\int_{\mathbb{R}^n_-} f(x)e^{-\pi\|x\|^2}\ dx\Big| \leq \int_{\mathbb{R}^n_-} \Big(|x_1| + \ldots + |x_n|\Big)e^{-\pi(x_1^2+\ldots+x_n^2)}\ dx = \frac{n}{\pi 2^n} = o(1)$$

as $n \longrightarrow +\infty$. Let us choose a positive a. Then we have

$$\int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\ dx \geq a \cdot \nu\big\{x \in \mathbb{R}^n : f(x) \geq a\big\} + o(1)$$

$$= a\Big(1 - \nu\big\{x \in \mathbb{R}^n : f(x) < a\big\}\Big) + o(1).$$

We observe that $f(x) < a$ if and only if $x_i < a$ for $i = 1, \ldots, n$ and hence

$$\nu\big\{x \in \mathbb{R}^n : f(x) < a\big\} = \Big(\int_{-\infty}^{a} e^{-\pi x^2}\ dx\Big)^n = \Big(1 - \int_a^{+\infty} e^{-\pi x^2}\ dx\Big)^n$$

$$\leq \Big(1 - \int_a^{(1+\epsilon)a} e^{-\pi x^2}\ dx\Big)^n \leq \big(1 - \epsilon a e^{-\pi(1+\epsilon)^2 a^2}\big)^n.$$

Therefore

$$\int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\ dx \geq a\Big(1 - \big(1 - \epsilon a e^{-\pi(1+\epsilon)^2 a^2}\big)^n\Big) + o(1).$$

Let us choose $a = a_n = \dfrac{1}{1+\epsilon}\sqrt{\dfrac{\ln n}{\pi}}$. Then $\big(1 - \epsilon a_n e^{-\pi(1+\epsilon)^2 a^2}\big)^n = \big(1 - \epsilon a_n/n\big)^n \longrightarrow 0$ as $n \longrightarrow +\infty$. Therefore for all sufficiently large n we have

$$\int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\ dx \geq \frac{1}{1+\epsilon}a_n$$

and the lower bound in (4.1.1) is proven.

Let us prove the upper bound. Let us choose an $a > 1$. Then we have

$$\int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\,dx \le \sum_{k=0}^{+\infty}(k+1)a\cdot\nu\{x : ka \le f(x) \le (k+1)a\}$$

$$\le a\sum_{k=0}^{+\infty}(k+1)\nu\{x : f(x) \ge ka\} \le a + a\sum_{k=1}^{+\infty}(k+1)\big(1-\nu\{x : f(x) < ka\}\big).$$

Now

$$\nu\{x \in \mathbb{R}^n : f(x) < ka\} = \Big(\int_{-\infty}^{ka} e^{-\pi x^2}\,dx\Big)^n = \Big(1 - \int_{ka}^{+\infty} e^{-\pi x^2}\,dx\Big)^n.$$

Let us choose $a = a_n = (1+\epsilon)\sqrt{\dfrac{\ln n}{\pi}}$. Thus for $k \ge 1$ we have

$$\int_{ka}^{+\infty} e^{-\pi x^2}\,dx \le \int_{ka}^{+\infty} 2\pi x e^{-\pi x^2}\,dx \le e^{-\pi k^2 a^2} = n^{-(1+\epsilon)^2k^2}$$

and

$$\nu\{x \in \mathbb{R}^n : f(x) < ka\} \ge \big(1 - n^{-(1+\epsilon)^2k^2}\big)^n \ge 1 - n^{-\epsilon k^2}$$

for all sufficiently large n. Summarizing, we get that for all sufficiently large n

$$\int_{\mathbb{R}^n} f(x)e^{-\pi\|x\|^2}\,dx \le (1+\epsilon)\sqrt{\frac{\ln n}{\pi}}\sum_{k=0}^{+\infty}(k+1)n^{-\epsilon k^2} \le (1+\epsilon)^2\sqrt{\frac{\ln n}{\pi}}$$

and the upper bound in (4.1.1) is proven. Our result readily follows by (4.1.1). □

5. Proof of the Main Result

In this section we complete the proof of Theorem 1.2. Our plan is the following. Let us introduce a polytope $P_n \subset \mathbb{R}^{n^2}$ as the set of solutions $x = (x_{ij})$ to the system of n linear equations and n^2 inequalities

$$\sum_{j=1}^{n} x_{ij} = 1 : \quad i = 1,\ldots,n \qquad \text{and} \qquad x_{ij} \ge 0 : \quad i,j = 1,\ldots,n.$$

Let us consider $\mathbb{R}^{n^2}$ as the direct sum of n copies of $\mathbb{R}^n$, so $\mathbb{R}^{n^2} = \mathbb{R}^n_1 \oplus \ldots \oplus \mathbb{R}^n_n$ with $(x_{k1},\ldots,x_{kn})$ being the coordinates in the k-th summand $\mathbb{R}^n_k$. It is seen that the polytope P_n can be represented as the Minkowski sum of the $(n-1)$-dimensional simplices $\Delta_k \subset \mathbb{R}^n_k$, $k = 1,\ldots,n$

$$\Delta_k = \Big\{(x_{kj}) : \; x_{k1} + \ldots + x_{kn} = 1 \quad \text{and} \quad x_{kj} \ge 0 : j = 1,\ldots,n\Big\}.$$

Using Theorem 4.1 and (2.4.2) we estimate $\rho(P_n)$. Then using Theorem 3.2 and (2.4.3) we estimate $\rho(Q_X)$.

(5.1) Lemma.

$$\int_{S^{n^2-1}} h_{P_n}(u)\ du = \sqrt{2\ln n}\Big(1+o(1)\Big)$$

as $n \longrightarrow +\infty$.

Proof. We use the representation of $P_n = \Delta_1 + \ldots + \Delta_n$ as the Minkowski sum of simplices Δ_k. Thus for $x = (x_{ij}) \in \mathbb{R}^{n^2}$ we have

$$h_{P_n}(x) = \sum_{k=1}^{n} h_{\Delta_k}(x) = \sum_{k=1}^{n} \max\big\{x_{k1}, \ldots, x_{kn}\big\}.$$

Since $h_{P_n}(x)$ is a homogeneous function of degree 1, applying the standard trick of passing to polar coordinates, we get

$$\int_{S^{n^2-1}} h_{P_n}(u)\ du = \frac{\sqrt{\pi}\,\Gamma\Big(\frac{n^2}{2}\Big)}{\Gamma\Big(\frac{n^2+1}{2}\Big)} \int_{\mathbb{R}^{n^2}} h_{P_n}(x) e^{-\pi\|x\|^2}\ dx.$$

On the other hand, using Theorem 4.1 we get

$$\int_{\mathbb{R}^{n^2}} h_{P_n}(x) e^{-\pi\|x\|^2}\ dx = \sum_{k=1}^{n} \int_{\mathbb{R}^{n^2}} \max\big\{x_{k1}, \ldots, x_{kn}\big\} e^{-\pi\|x\|^2}\ dx$$

$$= \sum_{k=1}^{n} \int_{\mathbb{R}^n_k} \max\big\{x_{k1}, \ldots, x_{kn}\big\} e^{-\pi(x_{k1}^2 + \ldots + x_{kn}^2)}\ dx_{k1} \ldots\ dx_{kn} = n\sqrt{\frac{\ln n}{\pi}}\Big(1+o(1)\Big).$$

Stirling's formula implies that

$$\lim_{n\longrightarrow\infty} \frac{n\,\Gamma\Big(\frac{n^2}{2}\Big)}{\Gamma\Big(\frac{n^2+1}{2}\Big)} = \sqrt{2}$$

and the proof follows. □

(5.2) Corollary. *We have*

$$\rho(P_n) = \sqrt{2\ln n}\,(1+o(1))$$

as $n \longrightarrow +\infty$.

Proof. Follows by Lemma 5.1 and (2.4.2) with $R = \sqrt{n}$ and $d = n^2$. □

Proof of Theorem 1.2. Let us prove (1.2.1). Let $V_n \subset \mathbb{R}^{n^2}$ be the vertex set of P_n. We observe that V_n consists of 0-1 matrices x_{ij} such that each row contains precisely one 1. Therefore $|V_n| = n^n$ and for every two vertices $x, y \in V_n$ there exists an orthogonal transformation $g_{x,y}$ of $\mathbb{R}^{n^2}$ such that $g_{x,y}(x) = y$ (one can choose $g_{x,y}$ to be a permutation of the coordinates in $\mathbb{R}^{n^2}$). Let $U_X = \{\pi(\sigma) : \sigma \in X\}$ be the vertex set of Q_X. Then $U_X \subset V_n$ and we may apply Theorem 3.2 with $R = \sqrt{n}$ and $\alpha = n^n/|X| = \exp\{\delta(X) n \ln n\}$. Thus we have

$$\rho(Q_X) \geq \rho(P_n) - \sqrt{2\delta(X)\ln n}\,(1 + o(1))$$

and the lower bound for $\rho(Q_X)$ in (1.2.1) follows by Corollary 5.2. The upper bound in (1.2.1) follows by (2.4.3) with $v = |X|$, $R = \sqrt{n}$ and $d = n^2$.
Part (1.2.2) follows by (2.4.1) with $R = \sqrt{n}$ and $d = n^2$. □

6. Approximate Counting and Random Optimization

Theorem 1.2 has obvious applications to the average case analysis of Combinatorial Optimization problems (cf. Example 1.3). What seems to be more interesting is that Theorem 1.2 can be used for estimating the cardinality of an implicitly given subset $X \subset S_n$ by solving an optimization problem on X with a randomly chosen goal function. We are going to describe a *randomized algorithm* for approximate enumeration. Our algorithm uses a procedure that samples a random point from the uniform distribution on the unit sphere in Euclidean space. It is known that there is a polynomial time algorithm for simulating this distribution from the standard Bernoulli distribution, see, for example, [4]. By introducing randomness, we allow a certain probability that our algorithm does not produce the correct answer. However, we'll make sure that this probability quickly fades to 0 as $n \longrightarrow +\infty$.

Our algorithm is the following. We sample a random point $u \in S^{n^2-1}$ from the uniform distribution on the sphere, solve the optimization problem

$$(6.1) \qquad \text{Given} \quad u \in S^{n^2-1}, \qquad \text{compute} \quad h_{Q_X}(u) = \max\Big\{\sum_{i=1}^{n} u_{i\sigma(i)} : \sigma \in X\Big\}$$

at this point u and then write the estimate

$$(6.2) \quad \alpha_X^2(u) + o(1) \leq \frac{\ln|X|}{n\ln n} \leq 2\alpha_X(u) - \alpha_X^2(u) + o(1), \qquad \text{where} \quad \alpha_X(u) = \frac{h_{Q_X}(u)}{\sqrt{2\ln n}}.$$

As above, "$o(1)$" stands for a function depending on n alone which tends to 0 as n grows to infinity. Indeed, from (1.2.2) we deduce that with the probability at least $1 - O\big(e^{-\sqrt{n}}\big)$ the solution to (6.1) at a randomly chosen point $u = (u_{ij}) \in S^{n^2-1}$ approximates $\rho(Q_X)$ within at most $n^{-1/4}$ error (let us call such a $u \in S^{n^2-1}$ *typical*). Then the inequalities (6.2) for a typical u can be obtained by inverting (1.2.1).

We note that the gap between the upper and lower bound in (6.2) is small when $\alpha_X(u)$ is close to 0 or 1 and is the largest if $\alpha_X(u) = 1/2$.

Sometimes the problem (6.1) is easy to solve and so we get an easy way to approximate the cardinality of $|X|$ via (6.2) Even when the problem (6.1) is difficult, we can still get a lower bound for $|X|$ by choosing a permutation $\sigma \in X$ and using the number $\sum_{i=1}^{n} u_{i\sigma(i)}$ as a lower bound for $h_{Q_X}(u)$. Theorem 1.2 implies that for a typical $u \in S^{n^2-1}$ there is a $\sigma \in X$ such that

$$\frac{1}{\sqrt{2\ln n}} \sum_{i=1}^{n} u_{i\sigma(i)} \geq 1 - \sqrt{\delta(X)} + o(1).$$

If we are lucky we can *guess* such a σ and hence by (6.2) we'll have a probabilistic *certificate* that $\ln|X|$ is at least $\big(1 - \sqrt{\delta(X)} + o(1)\big)^2 n \ln n$. If the problem (6.1) is easy we can *find* such a certificate efficiently. We discuss two examples below.

(6.3) Estimating the permanent. Let $A = (a_{ij})$ be an $n \times n$ matrix with 0-1 entries. The number $\operatorname{per} A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}$ is called the *permanent* of A. Computing the permanent is a #P-complete problem and even estimating it seems to be difficult if no special property of A is assumed. Let us interpret A as the adjacency matrix of a directed graph G. Let $X \subset S_n$ be the set of permutations σ such that $a_{i\sigma(i)} = 1$ for each $i = 1, \ldots, n$. Thus $\operatorname{per} A = |X|$ and the permutations from X are in one-to-one correspondence with *cycle covers* of G. Then (6.1) is the problem of computing the largest weight of a cycle cover in G provided each edge (i, j) of G has given weight u_{ij}. This problem admits an algorithm of $O(n^3)$ complexity (see, for example, [6]). Hence in this example we get a randomized polynomial time algorithm for estimating $\operatorname{per} A = |X|$ via (6.2).

(6.4) Counting Hamiltonian circuits. Let G be a directed graph with n vertices. An ordering $1, \ldots, n$ of its vertices is called a *Hamiltonian circuit* iff $(1,2), (2,3), \ldots, (n-1, n)$ and $(n, 1)$ are edges of G. Let $\operatorname{ham}(G)$ be the number of Hamiltonian circuits in G. It is a #P-complete problem to compute $\operatorname{ham}(G)$ and even to check whether $\operatorname{ham}(G) > 0$ is an NP-complete problem. Let $X \subset S_n$ be the set of all one-cycle permutations σ such that $(i, \sigma(i))$ is an edge of G. Then $\operatorname{ham}(G) = |X|$. In this example the problem (6.1) of finding the maximal weight of a Hamiltonian circuit in a weighted graph is NP-hard. However, we can try to guess a Hamiltonian circuit with a sufficiently large weight. It follows that if $\operatorname{ham}(G) = \exp\{(1-\delta) n \ln n\}$ then there exists a short *randomized proof* that $\operatorname{ham}(G)$ is at least $\exp\{\big(1 - \sqrt{\delta} + o(1)\big)^2 n \ln n\}$. The proof consists of generating a random weighting $u \in S^{n^2-1}$ and then demonstrating a Hamiltonian circuit in G with a $\sqrt{2\ln n}\big(1 - \sqrt{\delta} + o(1)\big)$ weight (such a circuit exists almost surely as $n \longrightarrow +\infty$). Of course, to *find* such a proof is difficult but its very existence seems to be of interest.

Acknowledgment

The first author was supported by the Alfred P. Sloan Research fellowship and by NSF Grant DMS 9501129.

References

1. A.I. Barvinok, Integral geometry of higher-dimensional polytopes and the average case in combinatorial optimization, *Proceedings of the 36th Annual Symposium on the Foundations of Computer Science (FOCS'95)*, IEEE Computer Society Press, 1995, pp. 275-283.
2. V.A. Emelichev, M.M. Kovalev and M.K. Kravtsov, *Polytopes, Graphs and Optimization*, Cambridge University Press, New York, 1984.
3. M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.
4. D.E. Knuth, *The Art of Computer Programming. Vol.2: Seminumerical Algorithms*, second edition, Addison-Wesley, Reading, MA, 1981.
5. V.D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces*, Lecture Notes in Mathematics, **1200**, Springer-Verlag, 1986.
6. C.H. Papadimitriou, K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1982.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109.