ESTIMATING L^∞ NORMS BY L^{2k} NORMS FOR FUNCTIONS ON ORBITS

Alexander Barvinok

April 2002

Abstract. Let G be a compact group acting in a real vector space V. We obtain a number of inequalities relating the L^∞ norm of a matrix element of the representation of G with its L^{2k} norm for a positive integer k. As an application, we obtain approximation algorithms to find the maximum absolute value of a given multivariate polynomial over the unit sphere (in which case G is the orthogonal group) and for the assignment problem of degree d, a hard problem of combinatorial optimization generalizing the quadratic assignment problem (in which case G is the symmetric group).

1. Introduction

A general optimization problem has to do with finding the maximum (minimum) value of a real-valued function $f : X \longrightarrow \mathbb{R}$. Often, the set X is endowed with a probability measure μ and the function f possesses a certain degree of symmetry which allows one to compute the k-th moment $\int_X f^k \ d\mu$ efficiently at least for small values of k. Thus one may ask how well the k-th moment approximates the maximum value. In this paper, we describe a fairly general situation where some simple and meaningful relations between the maximum and moments can be obtained. We provide two illustrations: one, continuous, has to do with optimization of multivariate polynomials on the unit sphere with possible applications to solving systems of real polynomial equations and the other, discrete, deals with optimization on the symmetric group, namely, with the assignment problem of degree d, a hard problem of combinatorial optimization.

1991 *Mathematics Subject Classification.* 68W25, 68R05, 90C30, 90C27, 20C15.

Key words and phrases. group representations, matrix elements, multivariate polynomials, combinatorial optimization, assignment problem, polynomial equations, L^p norms.

This research was partially supported by NSF Grant DMS 9734138.

Typeset by $\mathcal{AMS}$-TEX

(1.1) The General Setting. Let G be a compact group with the Haar probability measure dg acting in a finite-dimensional real vector space V. To avoid dealing with various technical details, we assume that the representation $G \longrightarrow GL(V)$ is continuous, where the general linear group $GL(V)$ is considered in its standard topology.

Let us choose a vector $v \in V$ and a linear function $\ell : V \longrightarrow \mathbb{R}$. We consider the orbit $\{gv : g \in G\}$ of v and the resulting function $f : G \longrightarrow \mathbb{R}$ defined by

$$f(g) = \ell(gv).$$

In other words, f is a matrix element in the representation of G. We are interested in the relation between the following quantities:

The L^∞ norm of f:

$$\|f\|_\infty = \max_{g \in G} |f(g)| = \max_{g \in G} \big|\ell(gv)\big|.$$

The L^{2k} norm of f for a positive integer k:

$$\|f\|_{2k} = \left(\int_G f^{2k}(g)\ dg\right)^{\frac{1}{2k}} = \left(\int_G \ell^{2k}(gv)\ dg\right)^{\frac{1}{2k}}.$$

As we remarked earlier, for many examples in computational mathematics, the quantity $\|f\|_\infty$ is of considerable interest and is hard to compute whereas $\|f\|_{2k}$ is relatively easy to compute for moderate values of k. First, we relate $\|f\|_\infty$ and $\|f\|_2$.

Theorem 1. *Let G be a compact group acting in a finite-dimensional real vector space V and let dg be the Haar probability measure on G. Let us fix a vector v and a linear function $\ell : V \longrightarrow \mathbb{R}$ and let us define a real-valued function $f : G \longrightarrow \mathbb{R}$ by $f(g) = \ell(gv)$. Then*

$$\|f\|_2 \leq \|f\|_\infty \leq \sqrt{\dim V} \cdot \|f\|_2.$$

The bounds of Theorem 1 are generally sharp, see Section 2.1. To estimate how well $\|f\|_{2k}$ approximates $\|f\|_\infty$ for a larger k, we invoke a general construction from the representation theory, see for example, (Fulton and Harris 1991, Lecture 6).

(1.2) Tensor Power. For a positive integer k, let

$$V^{\otimes k} = \underbrace{V \otimes \ldots \otimes V}_{k \text{ times}}$$

be the k-th tensor power of V. There is a natural action of G in $V^{\otimes k}$, defined on decomposable tensors by

$$g(v_1 \otimes \ldots \otimes v_k) = gv_1 \otimes \ldots \otimes gv_k \quad \text{for} \quad g \in G.$$

There is a natural action of the symmetric group S_k permuting the components of the tensor product. Thus, for decomposable tensors, we have

$$\sigma(v_1 \otimes \ldots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(k)} \quad \text{for} \quad \sigma \in S_k.$$

The action of S_k commutes with the action of G. Let $\mathrm{Sym}_k(V)$ be the symmetric part of $V^{\otimes k}$ consisting of the tensors x such that $\sigma x = x$ for all $\sigma \in S_k$. It is known that

$$\dim \mathrm{Sym}_k(V) = \binom{\dim V + k - 1}{k},$$

since $\mathrm{Sym}_k(V)$ can be thought of as the space of all real homogeneous polynomials of degree k in $\dim V$ variables. Let

$$v^{\otimes k} = \underbrace{v \otimes \ldots \otimes v}_{k \text{ times}}$$

be the k-th tensor power of v. Thus $v^{\otimes k} \in \mathrm{Sym}_k(V)$ and $gv^{\otimes k} \in \mathrm{Sym}_k(V)$ for all $g \in G$. It turns out that how well $\|f\|_{2k}$ approximates $\|f\|_\infty$ depends on the dimension D_k of the subspace spanned by the orbit $\{gv^{\otimes k}\}$. This dimension may be different for different $v \in V$. Roughly, if D_k is small then v lies in a certain algebraic variety constructed from the action of G in V and for such v the functions f are "smoother" than for those v for which D_k is large.

Thus we obtain the following corollary of Theorem 1.

Corollary 2. *Let G be a compact group acting in a finite-dimensional real vector space V and let dg be the Haar probability measure on G. Let us fix a vector v and a linear function $\ell : V \longrightarrow \mathbb{R}$ and let us define a real-valued function $f : G \longrightarrow \mathbb{R}$ by $f(g) = \ell(gv)$. For a positive integer k, let*

$$D_k = \dim \mathrm{span}\{gv^{\otimes k} : g \in G\}$$

be the dimension of the span of the orbit of $v^{\otimes k}$ in $V^{\otimes k}$. Then

$$\|f\|_{2k} \leq \|f\|_\infty \leq \big(D_k\big)^{\frac{1}{2k}} \cdot \|f\|_{2k}.$$

Again, generally speaking, the estimates of Corollary 2 can not be improved.

H. Derksen informed the author that numbers D_k turn out to be quite useful for the constructive invariant theory, cf. (Derksen and Kraft 1997).

A straightforward estimate of $D_k \leq \dim \mathrm{Sym}_k(V)$ produces the following corollary.

Corollary 3. *Let G be a compact group acting in a finite-dimensional real vector space V and let dg be the Haar probability measure on G. Let us fix a vector v and* *a linear function $\ell : V \longrightarrow \mathbb{R}$ and let us define a real-valued function $f : G \longrightarrow \mathbb{R}$ by $f(g) = \ell(gv)$. Let k be a positive integer. Then*

$$\|f\|_{2k} \leq \|f\|_\infty \leq \binom{\dim V + k - 1}{k}^{\frac{1}{2k}} \cdot \|f\|_{2k}.$$

There are examples showing that the bounds of Corollary 3 are "almost tight". For instance, if $G = SO(n)$ is the orthogonal group acting in $V = \mathbb{R}^n$, computations of Section 3.1 show that the upper bound for $\|f\|_\infty$ is tight up to a factor of $\sqrt{2}$ (uniformly on k and n).

It follows from Corollary 3 that $\|f\|_{2k}/\|f\|_\infty \longrightarrow 1$ as long as $(\dim V)/k \longrightarrow 0$. As we remarked earlier, in many cases we are able to compute $\|f\|_{2k}$ efficiently if k is not very large. Quite often (see examples of Sections 3 and 4), we can compute $\|f\|_{2k}$ in polynomial time for any fixed k. The following estimate shows the type of bound that we can achieve if we fix k in advance.

Corollary 4. *For any $\epsilon > 0$ there exists a $k_0 = k_0(\epsilon) = O(\epsilon^{-2})$ such that for any positive integer $k > k_0$, for any compact group G acting in a real vector space V with $\dim V \geq k$, for any linear function $\ell : V \longrightarrow \mathbb{R}$, for any $v \in V$ and for the function $f(g) = \ell(gv)$, $f : G \longrightarrow \mathbb{R}$, we have*

$$\|f\|_{2k} \leq \|f\|_\infty \leq \epsilon\sqrt{\dim V} \cdot \|f\|_{2k}.$$

The paper is structured as follows. In Section 2, we prove Theorem 1 and Corollaries 2–4. In Section 3, we apply our results to the problem of finding the largest absolute value of a real homogeneous multivariate polynomial (form) on the unit sphere, in which case $G = SO(n)$, the orthogonal group. In particular, we present a simple polynomial time approximation algorithm to compute the largest absolute value on the sphere of a *fewnomial*, that is, a polynomial having only small (fixed) number of monomials. More generally, our method is applicable to forms with small (fixed) dimensions of their Newton polytopes. The results have potential applications to constructing efficient algorithms for testing the feasibility of a given system of real polynomial equations, since the latter problem can be reduced to estimating the maximum absolute value on the unit sphere of an appropriately constructed "penalty" form f.

In Section 4, we discuss an NP-hard problem of combinatorial optimization, which we call the assignment problem of degree d, in which case $G = S_n$, the symmetric group. In particular, our results lead to an approximation algorithm for finding a bijection between vertex sets of two hypergraphs H_1 and H_2, which maximizes the number edges of H_1 mapped onto the edges of H_2. For $d = 2$, the corresponding optimization problem is known as the *quadratic assignment problem*, see, for example, (Çela 1998). In the quadratic assignment problem, the objective function $f : S_n \longrightarrow \mathbb{R}$ is defined by $f(g) = \sum_{i,j} a_{ij} b_{g(i)g(j)}$, where $A = (a_{ij})$ and $B = (b_{ij})$ are two given matrices. Interpreting f as a matrix element in a representation of S_n, for any given $\epsilon > 0$, we obtain a polynomial time algorithm approximating $\|f\|_\infty = \max_{g \in S_n} |f(g)|$ within a factor of ϵn.

We use the real model for computational complexity, counting the number of arithmetic operations performed by the algorithm, see (Blum et al. 1998). Eventually, to compute $\|f\|_{2k}$ from $\|f\|_{2k}^{2k}$ we need to extract a root of degree $2k$, which we count as a single operation.

2. Proofs

In this section, we prove Theorem 1 and Corollaries 2–4. We need some standard facts from the representation theory see, for example, (Fulton and Harris 1991).

Let G be a compact group acting in a finite-dimensional real vector space V. As is known, V possesses a G-invariant scalar product $\langle \rangle$:

$$\langle u, v \rangle = \langle gu,\ gv \rangle \quad \text{for all} \quad u, v \in V \quad \text{and all} \quad g \in G.$$

We introduce the corresponding Euclidean norm:

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

The action (representation) is called *irreducible* if V contains no proper G-invariant subspaces. As is known, if G acts in a finite-dimensional real vector space V, then V can be represented as a direct sum of pairwise orthogonal (with respect to a given G-invariant scalar product) invariant subspaces V_i such that the action of G in each V_i is irreducible.

A somewhat "non-standard" feature of our construction is that we consider representations over the real rather than over the complex numbers. The same results can be obtained by using the more standard representation theory over $\mathbb{C}$, but switching from real to complex numbers and back somewhat obscures main simple ideas of the proof and makes it harder to see when the obtained bounds are sharp.

We need a "real" substitute for Schur's Lemma. It comes in the form of the following observation. Suppose that $q : V \longrightarrow \mathbb{R}$ is a G-invariant quadratic form, that is $q(gx) = q(x)$ for all $x \in V$ and all $g \in G$. We claim that the eigenspaces of q are G-invariant subspaces. A possible way to see that is to notice that the unit eigenvectors of q are precisely the critical points of the restriction $q : \mathbb{S} \longrightarrow \mathbb{R}$ where $\mathbb{S} = \big\{x : \|x\| = 1\big\}$ is the unit sphere in V.

Our first lemma is a real version of the orthogonality relations for matrix elements.

Lemma 5. *Let G be a compact group acting in a finite-dimensional real vector space V endowed with a G-invariant scalar product $\langle \rangle$. Suppose that the representation of G is irreducible and let dg be the Haar probability measure on G. Then*

$$\int_G \langle x,\ gv \rangle^2 \ dg = \frac{\|v\|^2 \cdot \|x\|^2}{\dim V} \quad \textit{for all} \quad x, v \in V.$$

Proof. Let us choose a vector $v \in V$ and let us define a quadratic form $q : V \longrightarrow \mathbb{R}$ by

$$q(x) = \int_G \langle x,\ gv \rangle^2 \ dg.$$

Clearly, $q(x)$ is G-invariant: $q(gx) = q(x)$ for all $x \in V$ and all $g \in G$. Let λ be the largest eigenvalue of q and let W be the corresponding eigenspace. Then W is an invariant subspace of V and hence $W = V$. Thus $q(x) = \lambda \|x\|^2$ for some $\lambda \geq 0$. To find λ, let us compute the trace of q.

On one hand, we have $\operatorname{tr}(q) = \lambda \dim V$. Let $q_g(x) = \langle x,\ gv \rangle^2$. Then q_g is a quadratic form of rank 1 with the non-zero eigenvalue $\|gv\|^2 = \|v\|^2$ which corresponds to an eigenvector $x = gv$. Hence $\operatorname{tr}(q_g) = \|v\|^2$ and since $q(x)$ is the average of q_g, we have $\operatorname{tr}(q) = \|v\|^2$. Thus $\lambda = \|v\|^2 / \dim V$. Hence

$$q(x) = \frac{\|v\|^2 \cdot \|x\|^2}{\dim V}$$

and the proof follows. □

Now we use that every representation is a sum of irreducible representations.

Lemma 6. *Let G be a compact group acting in a finite-dimensional real vector space V endowed with a G-invariant scalar product $\langle \rangle$. Let dg be the Haar probability measure on G. Let us fix a vector $v \in V$. Then there exists a decomposition $V = V_1 \oplus \ldots \oplus V_k$ of V into the direct sum of non-zero pairwise orthogonal invariant subspaces such that for every $x \in V$ we have*

$$\int_G \langle x,\ gv \rangle^2 \ dg = \sum_{i=1}^k \frac{\|x_i\|^2 \cdot \|v_i\|^2}{\dim V_i},$$

where x_i and v_i are the orthogonal projections onto V_i of x and v respectively.

Proof. Let us define a quadratic form $q : V \longrightarrow \mathbb{R}$ by

$$q(x) = \int_G \langle x,\ gv \rangle^2 \ dg.$$

Then q is G-invariant, $q(gx) = q(x)$ for all $g \in G$ and all $x \in V$. Thus the eigenspaces of q are G-invariant subspaces of V. Let us write every eigenspace as a direct sum of pairwise orthogonal invariant subspaces V_i such that the action of G in each V_i is irreducible. Thus we obtain the decomposition $V = V_1 \oplus \ldots \oplus V_k$ and we have

$$q(x) = \sum_{i=1}^k \lambda_i \|x_i\|^2,$$

where x_i is the orthogonal projection of x onto V_i and λ_i are non-negative numbers. To find λ_i, let us choose a non-zero $x \in V_i$. Then $\langle x, gv \rangle = \langle x, gv_i \rangle$ and by Lemma 5, we get

$$q(x) = \lambda_i ||x||^2 = \frac{||v_i||^2 \cdot ||x||^2}{\dim V_i},$$

from which

$$\lambda_i = \frac{||v_i||^2}{\dim V_i}.$$

The proof now follows. □

Remark. A decomposition $V = V_1 \oplus \ldots \oplus V_k$ of a representation into the direct sum of pairwise orthogonal irreducible components is *not* unique as long as some irreducible representation appear with a multiplicity greater than 1 (which means that the representations of G in some subspaces V_i are isomorphic). One can construct some simple examples showing that the decomposition of Lemma 6 indeed depends on v.

Now we are ready to prove Theorem 1.

Proof of Theorem 1. The inequality

$$||f||_2 \le ||f||_\infty$$

is quite standard. Let us prove that

$$||f||_\infty \le \sqrt{\dim V} \cdot ||f||_2.$$

Let e be the identity in G. We note that it suffices to prove that

$$|f(e)| = |\ell(v)| \le \sqrt{\dim V} \cdot ||f||_2,$$

because the inequality for $f(g) = \ell(gv)$ would follow then by choosing a new vector v:

$$\text{new } v := g\big(\text{old } v\big).$$

Let us introduce a G-invariant scalar product $\langle\rangle$ in V so that $\ell(x) = \langle c, x \rangle$ for some $c \in V$ and all $x \in V$. Applying Lemma 6, we obtain a decomposition $V = V_1 \oplus \ldots \oplus V_k$ of V into the direct sum of pairwise orthogonal invariant subspaces V_i such that

$$||f||_2^2 = \int_G \langle c,\ gv \rangle^2\ dg = \sum_{i=1}^k \frac{||c_i||^2 \cdot ||v_i||^2}{\dim V_i},$$

where c_i and v_i are the orthogonal projections onto V_i of c and v respectively. We have

$$f(e) = \langle c, v \rangle = \sum_{i=1}^k \langle c_i, v_i \rangle$$

and hence

$$|f(e)| \leq \sum_{i=1}^{k} |\langle c_i, v_i \rangle| \leq \sum_{i=1}^{k} \|c_i\| \cdot \|v_i\|.$$

Let

$$\alpha_i = \frac{\|c_i\| \cdot \|v_i\|}{\sqrt{\dim V_i}}.$$

Then

$$|f(e)|^2 \leq \left(\sum_{i=1}^{k} \alpha_i \sqrt{\dim V_i} \right)^2 \leq \Big(\sum_{i=1}^{k} \alpha_i^2 \Big) \Big(\sum_{i=1}^{k} \dim V_i \Big) = \big(\dim V\big) \cdot \|f\|_2^2.$$

and the proof follows. □

(2.1) When is the Upper Bound for $\|f\|_\infty$ Attained ?

Analyzing the proof of Theorem 1, it is not hard to find out when the bound $\|f\|_\infty \leq \sqrt{\dim V} \cdot \|f\|_2$ is sharp. For example, if the action of G is irreducible, it follows by Lemma 5 that the equality $\|f\|_\infty = \sqrt{\dim V} \cdot \|f\|_2$ holds if and only if $\ell(x)$ is proportional to the scalar product of x with one of the orbit vectors gv. Generally, the bound is sharp for the class of linear functions on the orbit of v as long as the orbit of v spans V. Here are some natural cases when the bound is attained.

Suppose, for example, that we have an absolutely irreducible representation ρ of G in a real vector space W (that is, the representation remains irreducible after complexification). Thus, for every $g \in G$, $\rho(g)$ is an operator in W. We interpret $\rho(g)$ as a point in the space $V = \mathrm{End}(W) = W^* \otimes W$ of all linear transformations $W \longrightarrow W$. Let $\chi(g) = \mathrm{tr}\big(\rho(g)\big)$ be the character of the representation. We think of $\chi(g)$ as of a linear function on the orbit of the identity operator $I \in \mathrm{End}(W)$ under the action $g(x) = \rho(g)x$ for all $x \in \mathrm{End}(W)$ (in other words, G acts identically in W^* and by ρ in W in the tensor product $W^* \otimes W$). Then $\|\chi\|_\infty = \dim W = \sqrt{\dim V}$. The orthogonality relations for the characters, see, for example, (Fulton and Harris 1991, Lecture 2), state that $\|\chi\|_2 = 1$ and hence the bound of Theorem 1 holds with equality.

As another example, let us consider a finite group G of cardinality $|G|$ and an arbitrary function $f : G \longrightarrow \mathbb{R}$. The function f can be thought of as a linear function on the orbit of a point in the regular representation of G. The space V in this case is the vector space of all linear functions $f : G \longrightarrow \mathbb{R}$ where G acts by shifts: $gf(x) = f(g^{-1}x)$. Let $v \in V$ be the delta-function at the identity: $v(e) = 1$ where e is the identity in G and $v(g) = 0$ for all $g \neq e$. Then f is a linear function on the orbit of v and $\dim V = |G|$. Of course, in this case, the inequality $\|f\|_\infty \leq \sqrt{|G|} \cdot \|f\|_2$ is the best we can hope as it is sharp on delta-functions.

To prove Corollary 2, we use the construction of the tensor power (see Section 1.2).

Proof of Corollary 2. Let us define a function $h : G \longrightarrow \mathbb{R}$ by

$$h(g) = f^k(g) = \ell^{\otimes k}(gv^{\otimes k}).$$

Thus h is a linear function on the orbit of $v^{\otimes k}$. Let

$$W = \operatorname{span}\{gv^{\otimes k} : \; g \in G\}$$

be the span of the orbit of $v^{\otimes k}$. Hence $\dim W = D_k$. Applying Theorem 1 to the linear function h on the orbit of $v^{\otimes k}$ in W, we get

$$\|h\|_2 \leq \|h\|_\infty \leq \sqrt{D_k} \cdot \|h\|_2.$$

Now we note that $\|h\|_\infty = \|f\|_\infty^k$ and that $\|h\|_2 = \|f\|_{2k}^k$. □

Remark. The bound $\|f\|_\infty \leq \big(D_k\big)^{\frac{1}{2k}} \|f\|_{2k}$ is rarely sharp. One example when it is sharp is provided by a generic orbit in the regular representation of a finite group G, see Section 2.1. In Section 3.1, we present a series of examples of matrix elements of $G = SO(n)$ for which the estimate is sharp up to a constant factor uniformly on k and n.

Corollary 3 follows by a general estimate of D_k.

Proof of Corollary 3. We apply Corollary 2. The orbit $\{gv^{\otimes k}\}$ lies in the symmetric part $\operatorname{Sym}_k(V)$ of the tensor product $V^{\otimes k}$ and hence

$$D_k = \dim \operatorname{span}\{gv^{\otimes k}\} \leq \dim \operatorname{Sym}_k(V) = \binom{\dim V + k - 1}{k}.$$

□

(2.2) Obtaining Sharp Estimates. As follows from Section 3.1, the upper bound for $\|f\|_\infty$ is sharp up to a constant factor uniformly on k if $G = SO(n)$, $V = \mathbb{R}^n$ and G acts in V by its defining representation.

We describe below classes of functions $f : G \longrightarrow \mathbb{R}$ for which some sharp estimates can be obtained. Let us fix a representation ρ of G in a real vector space W, let $V = \operatorname{End}(W)$ and let f be a linear function on the orbit of the identity operator $I \in V$, see Section 2.1. Then for some constant $C(\rho, k)$ we have $\|f\|_\infty \leq C(\rho, k)\|f\|_{2k}$. Let us assume that ρ is absolutely irreducible (cf. Section 2.1). In principle, the best possible value of $C(\rho, k)$ can be computed from the representation theory of G as follows. Shifting f, if necessary, we may assume that the maximum absolute value of f is attained at the identity e of G. Let us define $h : G \longrightarrow \mathbb{R}$ by

$$h(x) = \int_G f(g^{-1}xg) \; dg \quad \text{for all} \quad x \in G.$$

Then $\|h\|_\infty = \|f\|_\infty$ and $\|h\|_{2k} \leq \|f\|_{2k}$ for all positive integers k. Thus the largest ratio $\|f\|_\infty / \|f\|_{2k}$ is attained when f satisfies $f(g^{-1}xg) = f(x)$ for all $g \in G$ and all

$x \in G$, from which it follows that f is a multiple of the character $\chi(g) = \operatorname{tr}\bigl(\rho(g)\bigr)$, see Section 2.1. We observe that $\|\chi\|_\infty = \dim W$. Moreover, the orthogonality relations, see (Fulton and Harris 1991, Lecture 6), imply that $\|\chi\|_{2k}^{2k}$ is the sum of squares of multiplicities of the irreducible components of the tensor power $\rho^{\otimes k}$. Summarizing, we conclude that to compute the best possible constant $C(\rho, k)$ such that $\|f\|_\infty \leq C(\rho, k)\|f\|_{2k}$, it suffices to know how the tensor power $\rho^{\otimes k}$ decomposes into the sum of absolutely irreducible representations. This example shows also that the algebra of tensor powers is indeed relevant for estimating $\|f\|_\infty$ in terms of $\|f\|_{2k}$.

Finally, Corollary 4 follows by an estimate of the binomial coefficient.

Proof of Corollary 4. Let us choose a k_0 such that $(k!)^{1/k} > 2\epsilon^{-2}$ for all $k > k_0$. By Stirling's formula, we can choose $k_0 = O\bigl(\epsilon^{-2}\bigr)$. Then

$$\binom{\dim V + k - 1}{k}^{\frac{1}{2k}} = \left(\frac{\dim V \cdot (\dim V + 1) \cdots (\dim V + k - 1)}{k!}\right)^{\frac{1}{2k}} \leq$$
$$\left(\frac{2^k \dim^k V}{k!}\right)^{\frac{1}{2k}} = 2^{1/2}(k!)^{-1/2k} \cdot \sqrt{\dim V} \leq \epsilon\sqrt{\dim V}.$$

The proof follows by Corollary 3. □

3. Applications to Polynomials

In this section, we apply our results to approximate the maximum absolute value of a homogeneous multivariate polynomial on the unit sphere.

Let p be a homogeneous polynomial of degree d in n real variables $\xi_1, \dots, \xi_n$. Thus we can write

$$p(x) = \sum_{1 \leq i_1, \dots, i_d \leq n} \gamma_{i_1 \dots i_d} \xi_{i_1} \cdots \xi_{i_d} \quad \text{for} \quad x = (\xi_1, \dots, \xi_n)$$

where $\gamma_{i_1 \dots i_d}$ are some real numbers.

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space and let $x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ be a point. Then $V = \bigl(\mathbb{R}^n\bigr)^{\otimes d}$ can be identified with the space $\mathbb{R}^{n^d}$. The coordinates of a typical point (tensor) $X \in V$ are

$$\Bigl(X_{i_1 \dots i_d} : \quad 1 \leq i_1, \dots, i_d \leq n\Bigr)$$

and the scalar product in V is defined by

$$\bigl\langle X,\ Y \bigr\rangle = \sum_{1 \leq i_1, \dots, i_d \leq n} X_{i_1 \dots i_d} Y_{i_1 \dots i_d}.$$

For $x=(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$, the coordinates of $x^{\otimes d}$ are

$$\Big(\xi_{i_1} \cdots \xi_{i_d} \quad \text{for} \quad 1 \leq i_1, \ldots, i_d \leq n\Big).$$

Therefore, we can write

$$p(x)=\big\langle c,\ x^{\otimes d}\big\rangle \quad \text{where} \quad c=\Big(\gamma_{i_1 \ldots i_d}\Big).$$

Let $G=SO(n)$ be the group of orientation preserving orthogonal transformations of $\mathbb{R}^n$. Then G acts in V by the d-th tensor power of its defining representation in $\mathbb{R}^n$. Let us choose $w=(1,0, \ldots, 0) \in \mathbb{R}^n$. Then, for any $g \in G$, we have

$$\big\langle c,\ g w^{\otimes d}\big\rangle=p(g w)$$

and the orbit $\{g w : g \in G\}$ is the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$. Thus the values of $p(x)$, as x ranges over the unit sphere in $\mathbb{R}^n$, are the values of the linear function

$$\ell\left(g w^{\otimes d}\right)=\langle c,\ g w^{\otimes d}\rangle=\langle c,\ x^{\otimes d}\rangle$$

as g ranges over the orthogonal group $SO(n)$.

Moreover, the push-forward of the Haar probability measure dg on G is the probability measure dx on $\mathbb{S}^{n-1}$. Thus we connect the values of a polynomial on the unit sphere with the values of a linear function on the orbit of the group $G=SO(n)$.

Corollary 7. *Let p be a homogeneous polynomial of degree d in n real variables, let $\mathbb{S}^{n-1}$ be the unit sphere in $\mathbb{R}^n$ and let dx be the rotation invariant probability measure on $\mathbb{S}^{n-1}$. For a positive integer k, let us define the L^{2k} norm of p by*

$$\|p\|_{2k}=\left(\int_{\mathbb{S}^{n-1}} p^{2k}(x)\ dx\right)^{\frac{1}{2k}}$$

and let us define the L^{∞} norm by

$$\|p\|_{\infty}=\max_{x \in \mathbb{S}^{n-1}}|p(x)|.$$

Then

$$\|p\|_{2k} \leq\|p\|_{\infty} \leq\binom{kd+n-1}{kd}^{\frac{1}{2k}}\|p\|_{2k}.$$

Proof. We apply Corollary 2. Let $w=(1,0, \ldots, 0)$ be as above. Then, for $v=w^{\otimes d}$, we can write $p(g w)=\ell(g v)$ for some linear function $\ell: V \longrightarrow \mathbb{R}$ and all $g \in G$.

The dimension D_k of the span of the orbit $\{gv^{\otimes k} = gw^{\otimes kd}: \ g \in G\}$ is that of the space of homogeneous polynomials of degree kd in n variables. Hence

$$D_k = \binom{kd+n-1}{kd}.$$

We have

$$\int_G \ell^{2k}(gv) \ dg = \int_{\mathbb{S}^{n-1}} p^{2k}(x) \ dx.$$

The proof now follows. □

One way to integrate polynomials over the unit sphere is to take the sum of the integrals of the monomials. The following result is certainly known, but for the sake of completeness, we sketch its proof here.

Lemma 8. *Let $p(x) = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$ be a monomial. If at least one of α_i's is odd then*

$$\int_{\mathbb{S}^{n-1}} p(x) \ dx = 0.$$

If $\alpha_i = 2\beta_i$, where β_i are non-negative integers for $i = 1, \ldots, n$, then

$$\int_{\mathbb{S}^{n-1}} p(x) \ dx = \frac{\Gamma(n/2) \prod_{i=1}^n \Gamma(\beta_i + 1/2)}{\pi^{n/2} \Gamma\big(\beta_1 + \ldots + \beta_n + n/2\big)},$$

where dx is the Haar probability measure on $\mathbb{S}^{n-1}$.

Sketch of Proof. If α_i is odd then

$$p(\xi_1, \ldots, \xi_{i-1}, -\xi_i, \xi_{i+1}, \ldots, \xi_n) = -p(\xi_1, \ldots, \xi_{i-1}, \xi_i, \xi_{i+1}, \ldots, \xi_n)$$

and hence the average value of p over the unit sphere is 0.

Assuming that $\alpha_i = 2\beta_i$ for $i = 1, \ldots, n$, we get

$$\int_{\mathbb{R}^n} p(x) e^{-\|x\|^2} \ d\mu = \prod_{i=1}^n \int_{\mathbb{R}} \xi^{2\beta_i} e^{-\xi^2} \ d\xi = \prod_{i=1}^n \Gamma\big(\beta_1 + 1/2\big),$$

where μ is the standard Lebesgue measure in $\mathbb{R}^n$.

On the other hand, passing to the polar coordinates and using that p is homogeneous of degree $d = 2(\beta_1, + \ldots + \beta_n)$, we get

$$\int_{\mathbb{R}^n} p(x) e^{-\|x\|^2} \ d\mu = |\mathbb{S}^{n-1}| \cdot \Big(\int_{\mathbb{S}^{n-1}} p(x) \ dx\Big) \cdot \Big(\int_0^{+\infty} r^{d+n-1} e^{-r^2} \ dr\Big),$$

where $|\mathbb{S}^{n-1}| = 2\pi^{n/2}/\Gamma(n/2)$ is the Euclidean volume of the unit sphere in $\mathbb{R}^n$. The proof now follows. □

The estimates of Corollary 7 are probably not optimal (apart from the case of $k = 1$), but the following simple example shows that in some sense, they are close to being optimal.

(3.1) Powers of Linear Functions. Let p be the power of a linear function, for example, $p(x) = \xi_1^d$. Then $\|p\|_\infty = 1$ and, by Lemma 8,

$$\|p\|_{2k} = \left(\frac{\Gamma(n/2)\Gamma(kd+1/2)}{\sqrt{\pi}\Gamma(kd+n/2)} \right)^{\frac{1}{2k}}.$$

Then Corollary 7 gives us the estimate

$$\|p\|_\infty \leq \left(\frac{\Gamma(n/2)\Gamma(kd+1/2)\Gamma(kd+n)}{\sqrt{\pi}\Gamma(kd+n/2)\Gamma(n)\Gamma(kd+1)} \right)^{\frac{1}{2k}} \leq \left(\frac{\Gamma(n/2)\Gamma(kd+n)}{\Gamma(kd+n/2)\Gamma(n)} \right)^{\frac{1}{2k}}$$
$$= \left(\frac{n(n+1)\cdots(kd+n-1)}{(n/2)(n/2+1)\cdots(kd+n/2-1)} \right)^{\frac{1}{2k}} \leq 2^{\frac{d}{2}}$$

Hence, among all homogeneous polynomials of a given degree d, powers of linear functions give the largest ratio $\|f\|_\infty / \|f\|_{2k}$ up to a constant factor depending on the degree of f and independent of the number of variables n and the value of k. In particular, if $d = 1$, the upper bound of Corollary 3 exceeds the tight bound by at most a factor of $\sqrt{2}$.

This may serve as an indication that the bounds of Corollary 3 are not too bad, cf. Sections 2.1 and 2.2. G. Blekherman (2001) pointed out to the author that the powers, in general, do not provide exactly the largest ratio $\|f\|_\infty / \|f\|_{2k}$ among all polynomials of a given degree d. Such "extremal" polynomials f were computed by G. Blekherman when some of the parameters n, d and k are small.

Suppose we want to approximate $\|p\|_\infty$ by $\|p\|_{2k}$ for a sufficiently large k. Let us see what trade-off between the computational complexity and accuracy we can achieve.

(3.2) Low Degree Polynomials. Let us fix the degree d and allow the number n of variables to vary. Suppose that we are given a homogeneous polynomial p of degree d and that we want to estimate $\|p\|_\infty$. This problem is NP-hard already for $d = 4$ (one can infer it from results of (Blum et al. 1998, Part 1) and is suspected to be hard for $d = 3$.

Let m be the number of monomials in p, so $m = O(n^d)$. We observe that for any fixed k, the direct computation of $p^{2k}(x)$ and computing $\|p\|_{2k}$ via Lemma 8 has $O(m^{2k})$ complexity. One the other hand, using Corollary 7, we get that

$$\|p\|_\infty \leq C(k) n^{d/2} \|p\|_{2k} \quad \text{where} \quad C(k) = O\big(k^{-1/2}\big).$$

In other words, for any fixed $\epsilon > 0$ there is a polynomial time algorithm estimating $\|p\|_\infty$ within a factor of $\epsilon n^{d/2}$. If we want a better estimate, we have to take a larger k. Thus, for any constant $C > 1$, from Corollary 7 we get that

$$\|p\|_\infty \leq C \|p\|_{2k} \quad \text{for some} \quad k = O(n).$$

Since $p^{2k}(x)$ contains at most $\binom{2kd+n-1}{2kd}$ monomials, we can compute $\|p\|_{2k}$ by Lemma 8 in $2^{O(n)}$ time. Summarizing, for any $C > 1$ there exists a $\gamma > 0$ such that we can approximate $\|p\|_\infty$ within a factor C in $2^{\gamma n}$ time.

(3.3) Fewnomials and their Extensions. Suppose that we do not fix the degree d of p but fix instead the number m of monomials in p. Thus we can write

$$p(x)=\sum_{i=1}^{m} p_i(x),$$

where

$$p_i(x)=\gamma_i \xi_1^{\alpha_{i1}}\cdots\xi_n^{\alpha_{in}}$$

are monomials. For a positive integer k, by the multinomial expansion, we get

$$p^{2k}=\sum_{\substack{r_1,\dots,r_m\geq 0\\ r_1+\ldots+r_m=2k}} \frac{(2k)!}{r_1!\cdots r_m!} p_1^{r_1}\cdots p_m^{r_m}.$$

Thus p^{2k} contains at most $\binom{m+2k-1}{m-1}$ monomials, which is a polynomial in k when m is fixed. Using Lemma 8, we compute $\|p\|_{2k}$ in $O\big(dn(2k)^m\big)$ time. Given an $\epsilon>0$, for some $\delta>0$ let us choose an integer $k=O(\epsilon^{-1}n^{1+\delta}\ln d)$ such that

$$\frac{n-1}{2k}\ln\big(kd+1\big)<\ln(1+\epsilon).$$

Using Corollary 7 and a simple estimate

$$\binom{kd+n-1}{n-1}=\frac{(kd+1)(kd+2)\cdots(kd+n-1)}{1\cdot 2\cdots(n-1)}\leq (kd+1)^{n-1}$$

we conclude that

$$\|p\|_{2k}\leq\|p\|_\infty\leq(1+\epsilon)\cdot\|p\|_{2k}.$$

Hence as long as the number of monomials is fixed, we get a polynomial time approximation algorithm, which, for any given $\epsilon>0$ computes the maximum absolute value of a given polynomial ("fewnomial") over the unit sphere within a relative error of ϵ, in time polynomial in ϵ^{-1}, the number of variables n and the degree d of the polynomial. In fact, the only place where we have to use polynomially many in d arithmetic operations is when we compute gamma-functions (factorials) in Lemma 8. Apart from that, the running time of the algorithm is polynomial in $\ln d$.

One can define a wider class of polynomials p for which a similar polynomial time approximation scheme exists. For a monomial $x^a=\xi_1^{\alpha_1},\dots,\xi_n^{\alpha_n}$, let $a=(\alpha_1,\dots,\alpha_n)\in\mathbb{Z}^n$ be its multidegree considered as a point in $\mathbb{R}^n$. For a polynomial p, let us define the subspace $\mathcal{L}_p\subset\mathbb{R}^n$ to be the span of all a such that the monomial x^a has a non-zero coefficient in p. In other, words, $\mathcal{L}_p$ is the span of the Newton polytope of p, cf. (Khovanskii 1991). Let us fix a positive integer m and consider the class of homogeneous polynomials p such that $\dim\mathcal{L}_p\leq m$. We note that for any k, the multidegrees of monomials in p^k are points from $\mathcal{L}_p \cap \mathbb{Z}^n$ contained in the cube $0 \leq \alpha_i \leq kd$ with $d = \deg p$. In particular, the number of monomials in p^k grows polynomially in k. Using that, one can compute the monomial expansion of p^k recursively in time polynomials in k, d and n. Hence, for any $\epsilon > 0$ one can estimate $\|p\|_\infty$ within relative error ϵ it time polynomial in ϵ^{-1}, n and d.

Computing or approximating the maximum absolute value of a polynomial on the unit sphere can be used for testing whether a given system of real polynomial equations has a real solution, a difficult (NP-hard) and important problem, see for example, (Blum et al. 1998) and (Renegar 1991). Suppose that p_i: $i = 1, \ldots, s$ are given homogeneous polynomials of degree d in n variables $x = (\xi_1, \ldots, \xi_n)$ and that we would like to test whether the system

$$p_i(x) = 0 \quad \text{for} \quad i = 1, \ldots, s$$

has a real solution $x \neq 0$. Let

$$q = \sum_{i=1}^{s} p_i^2(x).$$

Thus we want to test whether

$$\min_{x \in \mathbb{S}^{n-1}} q(x) = 0.$$

Let us choose a

$$\gamma > \max_{x \in \mathbb{S}^{n-1}} q(x)$$

and let

$$p = \gamma \|x\|^{2d} - q.$$

Thus the problem reduces to checking whether

$$\max_{x \in \mathbb{S}^{n-1}} |p(x)| = \gamma.$$

If the polynomials p_i of the original system do not have too many monomials or belong to other easy cases described above, we can try to approximate $\|p\|_\infty$ by $\|p\|_{2k}$ for a reasonably large k, cf. Section 3.3. Similarly, to choose an appropriate γ, we can compute $\|q\|_{2k}$ for a sufficiently large k. The number of monomials in the system is relevant to the "topological complexity" of the set of real solutions (Khovanskii 1991), so it should not be surprising that it is also relevant to the computational complexity of the decision problem. In particular, this approach may be useful for detecting "badly unsolvable" systems (systems for which the value of $\|p\|_\infty$ is substantially smaller than γ) of fewnomial equations.

4. Applications to Combinatorial Optimization

Let us fix a number d and let $V = \left(\mathbb{R}^n\right)^{\otimes d} = \mathbb{R}^{n^d}$ be the vector space of d-dimensional arrays (tensors)

$$X = \Bigl(x_{i_1 \dots i_d} : \quad 1 \leq i_1, \dots, i_d \leq n\Bigr).$$

To simplify the notation somewhat, we denote the coordinates of X by x_I, where $I = (i_1, \dots, i_d)$.

We introduce the scalar product by

$$\langle X, Y \rangle = \sum_I x_I y_I \quad \text{for} \quad I = (1 \leq i_1, \dots, i_d \leq n).$$

Let $G = S_n$ be the symmetric group of all permutations g of the set $\{1, \dots, n\}$. We introduce the action of S_n on V by the d-th tensor power of the natural action of S_n in $\mathbb{R}^n$:

$$Y = gX \quad \text{provided} \quad x_I = y_{gI} \quad \text{where} \quad g\bigl(i_1, \dots, i_d\bigr) = \bigl(g(i_1), \dots, g(i_d)\bigr).$$

Let us choose two tensors $A, B \in V$ and let

$$f(g) = \langle B,\ gA \rangle = \sum_{1 \leq i_1, \dots, i_d \leq n} a_{i_1 \dots i_d} b_{g(i_1) \dots g(i_d)}, \qquad f : S_n \longrightarrow \mathbb{R}$$

be the corresponding matrix element.

We call the problem of optimizing f the *assignment problem of degree d*. It is straightforward for $d = 1$. For $d = 2$, it is known as the *quadratic assignment problem*, one of the most general problems of combinatorial optimization, see, for example, (Çela 1998). The quadratic assignment problem is NP-hard and viewed as one of the "hardest" among NP-hard problems, compared, for example, with the traveling salesman problem.

(4.1) An Application: Matching Edges of Hypergraphs. Recall that a *d-hypergraph* H on the set $\{1, \dots, n\}$ is a set of subsets $E \subset \{1, \dots, n\}$, called *edges* of H, such that $|E| \leq d$ for the cardinality $|E|$ of every edge E of H. A hypergraph is called *uniform* provided $|E| = d$ for every edge E of H. Let H_1 and H_2 be uniform d-hypergraphs with the set of vertices $\{1, \dots, n\}$. Let us define the *adjacency tensor* $A = (a_{i_1 \dots i_d})$ of H_1 by

$$a_{i_1 \dots i_d} = \begin{cases} 1 & \text{if } \{i_1, \dots, i_d\} \text{ is an edge of } H_1 \\ 0 & \text{otherwise.} \end{cases}$$

Let us define $B = (b_{i_1 \dots i_d})$ by:

$$b_{i_1 \dots i_d} = \begin{cases} \frac{1}{d!} & \text{if } \{i_1, \dots, i_d\} \text{ is an edge of } H_2 \\ 0 & \text{otherwise.} \end{cases}$$

A permutation g of the set $\{1, \ldots, n\}$ is interpreted as a bijection between the vertices of H_2 and the vertices of H_1 and the value of

$$f(g) = \langle B,\ gA \rangle$$

is the number of edges of H_2 mapped onto the edges of H_1. The value of $\|f\|_\infty$ is the maximum number of edges of H_1 and H_2 that can be matched by a bijection of the vertices of H_1 and H_2. If H_1 and H_2 are not uniform, we can modify B by letting

$$b_{i_1, \ldots, i_d} = \frac{k_1! \cdots k_r!}{d!}$$

provided $\{i_1, \ldots, i_d\}$ is an edge of H_2 and the multiplicities of the elements in the multiset $\{\{i_1, \ldots, i_d\}\}$ are $k_1, \ldots, k_r$, so that $k_1 + \ldots + k_r = d$. Then again the value of $\|f\|_\infty$ is equal to the maximum number of edges of H_1 and H_2 that can be matched by a bijection of the vertex sets.

One can extend this construction to oriented hypergraphs whose edges are *ordered* subsets of $\{1, \ldots, n\}$. By introducing weights on the edges of H_1 and H_2 we can introduce "prices" for matching (or mismatching) particular edges.

Applying Corollary 3, we get the inequality

$$\|f\|_{2k} \leq \|f\|_\infty \leq \binom{n^d + k - 1}{k}^{\frac{1}{2k}} \|f\|_{2k}$$

for the function f of a general assignment problem of degree d.

In various special cases, the bound can be somewhat improved by using Corollary 2. For example, if the coordinates of A (or B) are 0's and 1's, one can prove that

$$\|f\|_{2k} \leq \|f\|_\infty \leq D^{\frac{1}{2k}}(n, d, k) \cdot \|f\|_{2k} \quad \text{where} \quad D(n, d, k) = \sum_{j=1}^{k} \binom{n^d}{j}.$$

We claim that for small (fixed) values of k the value of $\|f\|_{2k}$ can be computed relatively easily (in polynomial time). First, we observe that computation of $\|f\|_{2k}$ reduces to computation of the average of a matrix element for larger tensors.

Lemma 9. *Let us fix two tensors $A = (a_I)$ and $B = (b_I)$ for $I = \big(1 \leq i_1, \ldots, i_d \leq n\big)$. For a positive integer m (in particular, for $m = 2k$), let us define tensors $X = A^{\otimes m}$ and $Y = B^{\otimes m}$ as follows:*

$$X = (x_J) \quad \textit{and} \quad Y = (y_J) \quad \textit{where} \quad J = \big(1 \leq j_1, \ldots, j_{dm} \leq n\big)$$

and where

$$x_J = a_{I_1} \cdots a_{I_m} \quad \textit{and} \quad y_J = b_{I_1} \cdots b_{I_m} \quad \textit{provided} \quad J = (I_1, \ldots, I_m).$$

Then

$$\frac{1}{n!}\sum_{g\in S_n}\langle B,\ gA\rangle^m=\frac{1}{n!}\sum_{g\in S_n}\langle Y,\ gX\rangle.$$

Proof. The proof follows by observation that

$$\langle B,\ gA\rangle^m=\langle B^{\otimes m},\ gA^{\otimes m}\rangle=\langle Y,\ gX\rangle.$$

□

Next, we show how to compute the average.

Lemma 10. *Let us fix a positive integer l (in particular, $l=md=2kd$). For a partition $\Sigma=\{\Sigma_1,\dots,\Sigma_r\}$ of the set $\{1,\dots,l\}$ into non-empty disjoint subsets, we say that a sequence $I=(i_1,\dots,i_l)$ has type Σ if for each Σ_p the indices $i_j: j\in\Sigma_p$ are all equal and if for each pair of subsets Σ_p and Σ_q the indices $i_j: j\in\Sigma_p$ and $i_j: j\in\Sigma_q$ are different.*

Let $X=(x_I)$ and $Y=(y_I)$, $I=(1\le i_1,\dots,i_l\le n)$ be tensors (in particular, we can have $X=A^{\otimes m}=A^{\otimes 2k}$ and $Y=B^{\otimes m}=B^{\otimes 2k}$).

Let us define the tensors $\overline{X}=(\overline{x}_I)$ and $\overline{Y}=(\overline{y}_I)$, $I=(1\le i_1,\dots,i_l\le n)$ by

$$\overline{x}_I=\frac{(n-r)!}{n!}\sum_{J:\ type\ J=\ type\ I}x_J\quad provided\quad type\ I=(\Sigma_1,\dots,\Sigma_r)$$

and

$$\overline{y}_I=\frac{(n-r)!}{n!}\sum_{J:\ type\ J=type\ I}y_J\quad provided\quad type\ I=(\Sigma_1,\dots,\Sigma_r).$$

Then

$$\frac{1}{n!}\sum_{g\in S_n}\langle Y,\ gX\rangle=\langle\overline{Y},\ \overline{X}\rangle.$$

Proof. The two index sets $I=(i_1,\dots,i_l)$ and $J=(j_1,\dots,j_l)$ belong to the same orbit of the action $I\longmapsto gI$ of S_n if and only if they have the same type $\{\Sigma_1,\dots,\Sigma_r\}$. Moreover, the stabilizer of I consists of $(n-r)!$ permutations. Hence

$$\overline{X}=\frac{1}{n!}\sum_{g\in S_n}gX\quad\text{and}\quad\overline{Y}=\frac{1}{n!}\sum_{g\in S_n}gY.$$

We have

$$\frac{1}{n!}\sum_{g\in S_n}\langle Y,\ gX\rangle=\left\langle\ \frac{1}{n!}\sum_{g\in S_n}gY,\quad\frac{1}{n!}\sum_{g\in S_n}gX\ \right\rangle$$

and the proof follows. □

Combining Lemmas 9 and 10, we observe that as long as d and k are fixed, we can compute $\|f\|_{2k}$ in $O(n^{2kd})$ time, that is, in polynomial in n time.

In particular, from Corollary 4, we conclude that for any fixed d and for any fixed $\epsilon>0$ there exists a polynomial in n algorithm for estimating $\|f\|_\infty$ within a factor of $\epsilon n^{d/2}$. This result seems to be new already for $d=2$, cf. (Arkin et al. 2001).

(4.2) Finding a Permutation. So far we have shown how to approximate $\|f\|_\infty$ by $\|f\|_{2k}$ but we did not discuss how to find a particular permutation g which gives the value of $|f(g)|$ close to $\|f\|_\infty$. In fact, it is not hard to construct a permutation $g \in S_n$ for which $|f(g)| \geq \|f\|_{2k}$ and hence $|f(g)|$ approximates $\|f\|_\infty$ within a factor of $\epsilon n^{d/2}$ at the cost of some extra work, which still results in a polynomial time algorithm when k is fixed. The idea is to use the "method of conditional probabilities", see for example, (Motwani and Raghavan 1995, Section 5.6). We split the symmetric group S_n into the union of cosets $S_j = \{g : g(1) = j\}$ and then compute the average value of f^{2k} over each coset separately (this would require some straightforward modification of Lemma 10). Then a coset should be chosen which gives the largest average. Thus we have determined $g(1) = j$ and we proceed to determine $g(2), \ldots, g(n)$ successively.

References

E.M. Arkin, R. Hassin and M. Sviridenko, *Approximating the maximum quadratic assignment problem*, Inform. Process. Lett. **77** (2001), 13–16.

G. Blekherman, Personal communication (2001).

L. Blum, F. Cucker, M. Shub and S. Smale, *Complexity and Real Computation. With a foreword by Richard M. Karp*, Springer, New York, 1998.

E. Çela, *The Quadratic Assignment problem. Theory and Algorithms*, Combinatorial Optimization, vol. 1, Kluwer, Dordrecht, 1998.

H. Derksen and H. Kraft, *Constructive invariant theory*, Algèbre non commutative, groupes quantiques et invariants (Reims, 1995), Sémin. Congr., vol. 2, Soc. Math. France, Paris, 1997, pp. 221–244.

W. Fulton and J. Harris, *Representation Theory. A First Course*, Graduate Texts in Mathematics, vol. 129, Springer, New York, 1991.

A.G. Khovanskii, *Fewnomials*, Translations of Mathematical Monographs, vol. 88, 1991; Translated from the Russian by Smilka Zdravkovska, Amer. Math. Soc., Providence, RI.

R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge Univ. Press, Cambridge, 1995.

J. Renegar, *Computational complexity of solving real algebraic formulae*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), Math. Soc. Japan, Tokyo, 1991, pp. 1595–1606.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109
E-mail address: barvinok@umich.edu