DECIDING POSITIVITY OF MULTISYMMETRIC POLYNOMIALS

PAUL GÖRLACH, CORDIAN RIENER, AND TILLMANN WEISER

Abstract. The question how to certify non-negativity of a polynomial function lies at the heart of Real Algebra and also has important applications to Optimization. In this article we investigate the question of non-negativity in the context of multisymmetric polynomials. In this setting we generalize the characterization of non-negative symmetric polynomials given in [18, 22] by adapting the method of proof developed in [19]. One particular case where our results can be applied is the question of certifying that a (multi-)symmetric polynomial defines a convex function. As a direct corollary of our main result we are able to derive that in the case of (multi-)symmetric polynomials of a fixed degree testing for convexity can be done in a time which is polynomial in the number of variables. This is in sharp contrast to the general case, where it is known that testing for convexity is NP-hard already in the case of quartic polynomials [1].

1. Introduction

A real polynomial is called positive (non-negative) if its evaluation on every real point is positive (non-negative). The study of this property of polynomials functions is one of the aspects that separates Real Algebraic Geometry from Algebraic Geometry over algebraically closed fields. Indeed, Real Algebraic Geometry developed building on Hilbert’s problem of characterizing non-negative polynomials via sums of squares. On the complexity side, it is know that the problem to algorithmically decide whether a given polynomial assumes only positive or non-negative values is NP-hard in general (see for example [15, 6]) and is essential for example to understand global optimization of polynomial functions. Besides the general results some authors have studied particular cases of polynomials which are invariant under group actions, for example by permuting the variables. In particular, symmetric polynomials exhibit some interesting properties that behave differently over real closed and algebraically closed fields. For example, in [3] the authors show, that the equivariant cohomology of a complex variety which is defined by symmetric polynomials of a given maximal degree is bounded by a quantity which (for a large number of variables) only depends on this maximal degree. In contrast to that there are examples of real varieties where this does not hold, i.e., this quantity actually grows with the number of variables. A seminal characterization of symmetric polynomials that are positive is due to Procesi [16]. Building on work by Harris [11], Timofte [22] was able to provide a characterization of symmetric non-negative functions that can be used to algorithmically certify non-negativity: He could establish that a symmetric polynomial of degree $2d$ is non-negative if and only if it is non-negative on all points with at most $d$ distinct components. This observation leads to a number of interesting consequences. For example, it provides an essential part to
the description of the asymptotical behavior of the cone of symmetric non-negative forms of a given degree when the number of variables grows [5]. Algorithmically, this result allows to show that the complexity of deciding non-negativity of a symmetric function with a fixed degree only grows polynomially in the number of variables. Following the work of Timofte, the second author was able to provide short proofs of this characterization of symmetric non-negative polynomials [18, 19].

Contributions: In this article we extend the previous results on symmetric polynomials to arrive at a similar characterization of multisymmetric polynomial functions that assume only positive (non-negative) values. The class of multisymmetric polynomials naturally generalizes the symmetric polynomials. They can be thought of as functions which are invariant under simultaneously permuting \(k\)-tupels of variables. Similar to the case of symmetric polynomials, we are able to show that when the degree of such a polynomial is sufficiently smaller than the number of variables, also for these multisymmetric polynomials non-negativity can be checked on a lower-dimensional subset consisting of points whose orbit length is not maximal. As in the case of the usual action of the symmetric group, these points lie on linear subspaces.

Our main result is Theorem 14 which bounds the dimension of subspaces one has to consider to decide non-negativity of a multisymmetric polynomial. Besides this general bound the idea of the proof can be adjusted in particular situations to derive stronger bounds. We give several other bounds to illustrate this.

As an application of our results we investigate the question of deciding, if a given symmetric or multisymmetric polynomial defines a convex function. It is straightforward to observe that the question of convexity of a \(k\)-symmetric polynomial leads to the question of certifying whether a \(2k\)-symmetric polynomial is non-negative. Consequently, we show in Theorems 26 and 28 that our results on non-negativity of multisymmetric polynomials imply in particular that for (multi-)symmetric polynomials of a fixed degree the complexity of deciding convexity does not grow exponentially in the number of variables.

The article is structured as follows. In the next section we will give a brief introduction to the theory of multisymmetric polynomials and provide a relation between \(k\)-multisymmetric polynomials and \(k\)-variats polynomials which will play a crucial role in our arguments. With these preliminaries at hand we are able to state and prove the main result in Section 3. The last section is then devoted to the application of the non-negativity result to the problem of deciding convexity. This section also includes some refinements of our main results which apply to this setting.

2. Multisymmetric Polynomials

For \(n \in \mathbb{N}\) let \(S_n\) denote the symmetric group on \(n\) elements which acts on the \(n\)-dimensional vector space \(V := \mathbb{R}^n\) by permuting coordinates. For every \(k \in \mathbb{N}\) we consider the diagonal action of \(S_n\) on the vector space

\[ V^k := \bigoplus_{i=1}^{k} V \]
of \( k \)-tuples of vectors from \( V \). The action of \( S_n \) extends to the ring \( \mathbb{R}[V^k] \), which is just the polynomial ring \( \mathbb{R}[X_{11}, \ldots, X_{nk}] \) after identifying \( X_{11}, \ldots, X_{nk} \) with the standard basis of the \( i \)-th direct summand of \( V^k \). It is convenient to think of these variables as an \( n \times k \) array as indicated in the following notation.

**Notation 1.** Throughout this paper let

\[
X := \begin{pmatrix}
X_{11} & X_{12} & \cdots & X_{1k} \\
X_{21} & X_{22} & \cdots & X_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
X_{n1} & X_{n2} & \cdots & X_{nk}
\end{pmatrix}
\]

denote an \( n \times k \) array of variables. By \( X_i \) and \( X_{j} \) we denote the \( i \)-th row of \( X \) and \( j \)-th column of \( X \), respectively.

Following this notation, \( S_n \) permutes the rows of \( X \). The invariant ring of the polynomial ring \( \mathbb{R}[V^k] = \mathbb{R}[X_{11}, \ldots, X_{nk}] \) with respect to this action is the algebra of \( k \)-(multi-)symmetric polynomials, denoted by \( \mathbb{R}[V^k]^{S_n} \). Alternatively, \( \mathbb{R}[V^k]^{S_n} \) can be thought of as the \( n \)-fold symmetric product of the polynomial \( \mathbb{R} \)-algebra in \( k \) variables, which is a classically studied object (see [20, 13]). Note that for \( k = 1 \) this is the algebra of symmetric polynomials, which is a polynomial ring. In the case \( k > 1, n > 1 \) the symmetric group \( S_n \) is not operating as a finite reflection group, thus the algebra of \( k \)-symmetric polynomials is no longer a polynomial ring. However, for all \( k \) it is finitely generated as \( \mathbb{R} \)-algebra [7, 9]. For our purposes the representation in terms of so called multisymmetric power sums will be crucial.

**Definition 2.** Given a polynomial \( f \in \mathbb{R}[V^k] \), we denote its symmetrization, that is, the sum over its \( S_n \)-orbit, by \( \text{sym}(f) \in \mathbb{R}[V^k]^{S_n} \). For all \( \alpha \in \mathbb{N}^k \) we define the (multisymmetric) power sum

\[
p_\alpha := \text{sym}(X_1^\alpha),
\]

where \( X_1^\alpha := X_{11}^{\alpha_1} \cdots X_{1k}^{\alpha_k} \).

This is is a generalization of the power sum polynomials used in [19], where the symmetric case \( k = 1 \) is considered. In that case, the first \( n \) power sum polynomials form an algebraically independent set generating the \( \mathbb{R} \)-algebra of symmetric polynomials. In the \( k \)-symmetric case, this statement does not generalize, i.e., the generators are no longer algebraically independent.

In the following, we study the \( k \)-symmetric power sums and their connection to \( k \)-symmetric functions.

**Definition 3.** Let \( w := (w_1, \ldots, w_k) \) be a \( k \)-tuple of positive integers. We consider the grading on the \( \mathbb{R} \)-algebra \( \mathbb{R}[Y_1, \ldots, Y_k] \) given by defining \( Y_j \) to be homogeneous of degree \( w_j \). This grading also induces a grading on the \( \mathbb{R} \)-algebra \( \mathbb{R}[V^k] \) by the algebra-homomorphism \( \varphi : \mathbb{R}[V^k] \to \mathbb{R}[Y_1, \ldots, Y_k] \), \( X_{ij} \mapsto Y_j \). Alternatively, the latter grading is given by defining each \( X_{ij} \) to be homogeneous of degree \( w_j \).

The degree \( \text{deg}_w(f) \) of an element \( f \in \mathbb{R}[Y_1, \ldots, Y_k] \) (resp. \( f \in \mathbb{R}[V^k] \)) with respect to the above grading is called the \( w \)-weighted degree, or simply \( w \)-degree, of \( f \).
Alternatively, one can define the $w$-degree on the monomials of $\mathbb{R}[Y_1, \ldots, Y_k]$ and $\mathbb{R}[V^k]$ by $\deg_w(Y^\alpha) := \sum_{j=1}^k w_j \alpha_j$ and $\deg_w(X_1^{\alpha(1)} \cdots X_n^{\alpha(n)}) := \sum_{i=1}^n \sum_{j=1}^k w_j \alpha_j^{(i)}$, respectively, where $\alpha, \alpha^{(1)}, \ldots, \alpha^{(n)} \in \mathbb{N}$. After that, one extends this definition to a polynomial $f$ by taking the maximal $w$-degree of all monomials of $f$. Note that we retrieve the usual degree from this definition by setting all weights $w_j$ equal to 1.

**Example 4.** Fix $n \in \mathbb{N}$ rather large and consider for all (nonzero) parameters $\gamma \in \mathbb{R}^t$ the 2-symmetric polynomial given by

$$f(X) := \gamma_1 \sum_i X_{i1}^4 + \gamma_2 \left( \sum_i \sum_j X_{i1} X_{j1} \right)^2 - \gamma_3 \left( \sum_i \sum_j X_{i1} X_{j1} \right) \left( \sum_i X_{i1} X_{i2} \right) - \gamma_4 \left( \sum_i X_{i1}^3 \right) \left( \sum_i X_{i2} \right) - \gamma_5 \left( \sum_i X_{i1} \right) \left( \sum_i X_{i2} \right)^2 + \gamma_6 \left( \sum_i X_{i1} \right)^2 + \gamma_7 \left( \sum_i X_{i1} X_{i2} \right),$$

where all sums go from 1 to $n$. We have $\deg_{g(1,1)}(f) = \deg(f) = 4$. Recognizing that the exponents of the column $X_2$ are small compared to the exponents of $X_1$ one could give more weight to the second column, e.g. by considering the $(3,5)$-degree: $\deg_{g(3,5)}(f) = 14$.

**Notation 5.** Fix some weights $w$. For a polynomial $f \in \mathbb{R}[V^k]$ we define $M_f \subset \mathbb{R}[V^k]$ to be the set of monomials of $f$. By construction, $\varphi : \mathbb{R}[V^k] \to \mathbb{R}[Y_1, \ldots, Y_k]$, $X_{ij} \mapsto Y_j$ is a morphism of graded $\mathbb{R}$-algebras, considering the gradings as above. Thus,

$$\deg_w(f) = \max\{ \deg_w(m) \mid m \in \varphi(M_f) \} = \max\{ w^T \alpha \mid \alpha \in E_f \},$$

where $E_f \subset \mathbb{N}^k$ is the set of exponent tuples of all monomials in $\varphi(M_f)$. Note however, that in general $\deg_w(f) \neq \deg_w(\varphi(f))$, as $\varphi(M_f)$ may differ from $M_{\varphi(f)}$.

Let $\deg_w(f) = d$. Then the interpretation above gives rise to a very useful view on the $w$-degree as a hyperplane defining a simplex $\{ y \in \mathbb{R}^k \mid w^T y \leq d \}$ enclosing $E_f$.

Applied to Example 4, where $k = 2$, we can consider Figure 4. Note that $\varphi(M_f) = \{ Y_1^3, Y_1^2 Y_2, Y_2^3, Y_1 Y_2^2, Y_1 Y_2 \}$ and hence, $E_f = \{(4,0), (3,1), (2,0), (1,2), (1,1)\}$. Both choices of the weights $w$ in Example 4 define a triangle enclosing $E_f$.

In the following Theorem, we see that the lattice points contained in the simplex which is given by some weights $w$ have an emerging meaning when we consider the according power sums.

**Theorem 6.** Let $w = (w_1, \ldots, w_k)$ be a $k$-tuple of positive integers and let $d \in \mathbb{N}$. The $\mathbb{R}$-algebra generated by all power sums $p_\alpha$ with $|\alpha|_w := w^T \alpha \leq d$ contains all $k$-symmetric polynomials of $w$-degree at most $d$.

**Proof.** Reviewing the proof of [3, Thm. 1.2] yields the assertion:

It is enough to show that the first power sums generate all $k$-symmetric monomial functions

$$m_{\alpha^{(1)}, \ldots, \alpha^{(r)}} := \text{sym}(X_1^{\alpha^{(1)}} \cdots X_n^{\alpha^{(r)}})$$
of $w$-degree at most $d$, where $\ell \in \{1, \ldots, n\}$. For $\alpha^{(0)}, \ldots, \alpha^{(\ell)} \in \mathbb{N}^k$ and some positive integer $c$ the following equality holds:

$$c \cdot m_{\alpha^{(0)}, \ldots, \alpha^{(\ell)}} = p_{\alpha^{(0)}} m_{\alpha^{(1)}, \ldots, \alpha^{(\ell)}} - \sum_{i=1}^{\ell} m_{\alpha^{(1)}, \ldots, \alpha^{(i)} + \alpha^{(0)}, \ldots, \alpha^{(\ell)}}.$$ 

Since $\deg_w(m_{\alpha^{(0)}, \ldots, \alpha^{(\ell)}}) = \sum_{j=0}^{\ell} |\alpha^{(j)}|_w \leq d$, there are only polynomials of $w$-degree equal or less than $d$ on the right-hand side of this identity. Hence, we can write a monomial function with $\ell + 1$ exponent tuples as combination of a power sum of $w$-degree at most $d$ and some monomial functions with $\ell$ exponent tuples, both of $w$-degree at most $d$. Noticing that a monomial function with only one exponent tuple is a power sum ends the proof. □

Reconsider Figure 1. Taking the lattice points contained in the simplices given by $w_1$ and $w_2$, respectively, we get two sets of power sums, each of them sufficient to describe $f$ as a polynomial expression in its elements. In fact, the proof of Theorem 6 shows even more.

**Remark 7.** Let $f \in \mathbb{R}[V^k]^S_n$. For all $k$-tuples $w$ let $d_w := \deg_w(f)$. Then $f$ can be written as a polynomial expression in the power sums associated to the lattice points that are contained in the intersection $\bigcap_w \{y \in \mathbb{R}^k_+ \mid w^T y \leq d_w\}$.

The reason why we are interested in a different way of describing $k$-symmetric polynomials is illuminated by the following observation. For all $\alpha \in \mathbb{N}^k$ the partial derivative with respect to $X_{ij}$ of $p_\alpha$ is a polynomial in the variables $X_i$. Moreover, for every $j \in \{1, \ldots, k\}$ there exists one polynomial $r_j \in \mathbb{R}[Y_1, \ldots, Y_k]$ such that

$$\partial_{ij} p_\alpha = r_j(X_i).$$

Note that $r_j$ does not depend on $i$. This makes the derivatives of power sums easier to handle than the derivatives of the usual generating set, that is the set of monomials of $w$-degree at most $d$. We immediately conclude the following result for linear combinations of power sums:
Proposition 8. Let \( w = (w_1, \ldots, w_k) \) be a \( k \)-tuple of positive integers. Denote \( N_d := \{ \alpha \in \mathbb{N}^k \mid |\alpha|_w \leq d \} \) and let \( u \in \mathbb{R}^{|N_d|} \). Then there are \( k \) polynomials \( \tilde{q}_1, \ldots, \tilde{q}_k \in \mathbb{R}[Y_1, \ldots, Y_k] \) such that for all \( 1 \leq i \leq n, 1 \leq j \leq k \) we have

\[
\partial_{ij} \left( \sum_{\alpha \in N_d} u_\alpha p_\alpha \right) = \tilde{q}_j(X_i)
\]

and \( \tilde{q}_j \) is of \( w \)-degree at most \( d - w_j \) for each \( j \in \{1, \ldots, k\} \).

We will use this fact in the proof of Theorem 14.

3. Positivity of multisymmetric polynomials

Definition 9. For a positive integer \( m \) let \( A_m \) denote the subset of \( \mathbb{R}^{n \times k} \) consisting of all points \( x = (x_{ij}) \) with at most \( m \) distinct rows:

\[ A_m := \left\{ x \in \mathbb{R}^{n \times k} \mid \# \{ x_1, \ldots, x_n \} \leq m \right\}. \]

For \( f \in \mathbb{R}[V^k] \) or, more generally, \( f \in C^0(V^k) \) we define \( \kappa(f) \) to be the smallest positive integer such that

\[ \min_{x \in B_r} f(x) = \min_{x \in B_r \cap A_m} f(x) \]

holds for all \( r \geq 0 \), where \( B_r := \{ x \in \mathbb{R}^{n \times k} \mid \sum_{i=1}^n \sum_{j=1}^k x_{ij}^2 = r \} \).

Note that \( \kappa(f) \leq n \) in any case. If \( f \in \mathbb{R}[V] \), \( k = 1 \), i.e., the case of a symmetric polynomials, it is known that \( \kappa(f) \leq \max\{2, \left\lfloor \deg f \right\rfloor \} \) (see [22, 18]). In Theorem 14 below we will show a bound for a \( k \)-symmetric polynomial \( f \in \mathbb{R}[V^k]^{S_n} \) in terms of the (weighted) degree of \( f \). We observe that \( \kappa \) is lower semi-continuous:

Proposition 10. Let \( (f_\ell)_{\ell \in \mathbb{N}} \subset C^0(V^k) \) be a sequence of continuous functions converging pointwise to some \( f \in C^0(V^k) \). Then

\[ \kappa(f) \leq \liminf_{\ell \to \infty} \kappa(f_\ell). \]

Proof. Note, that by compactness of \( B_r \) the sequence \( f_\ell|_{B_r} \) converges uniformly to \( f|_{B_r} \). This implies that for all \( m \in \mathbb{N}, \min_{x \in B_r} f_\ell(x) \) and \( \min_{x \in B_r \cap A_m} f_\ell(x) \) converge to \( \min_{x \in B_r} f(x) \) and \( \min_{x \in B_r \cap A_m} f(x) \), respectively. \( \square \)

By definition of \( \kappa(f) \) we immediately see:

Proposition 11. Let \( f \in \mathbb{R}[V^k] \).

(i) If \( f \geq 0 \) on \( A_{\kappa(f)} \), then \( f \geq 0 \) on \( \mathbb{R}^{n \times k} \).

(ii) If \( f > 0 \) on \( A_{\kappa(f)} \), then \( f > 0 \) on \( \mathbb{R}^{n \times k} \).

(iii) If \( f \neq 0 \) on \( A_{\kappa(f)} \), then \( f \neq 0 \) on \( \mathbb{R}^{n \times k} \).

Definition 12. Let \( n \in \mathbb{N} \). A tuple \( \lambda := (\lambda_1, \ldots, \lambda_\ell) \) of \( \ell \) positive integers such that \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_\ell \) and \( n = \lambda_1 + \ldots + \lambda_\ell \) is called an \( \ell \)-partition of \( n \).
The following Proposition gives a rough estimate on the number of $\ell$-partitions for a fixed natural number $n$.

**Proposition 13.** Let $n \in \mathbb{N}$. Then for every $\ell \in \mathbb{N}$ the number of $\ell$-partitions of $n$ is bounded by $n^\ell$.

Algorithmically checking for the global properties in Proposition 11 is an instance of a decision problem. As mentioned, it is known that deciding global positivity of a multivariate polynomial is NP-hard in general (see for example [15]). Our aim is to exploit the structure of $k$-symmetric polynomials $f$ by bounding $\kappa(f)$.

Observe, that $A_m$ is a union of $km$-dimensional subspaces, each of which corresponds to one particular way of assigning the $m$ distinct rows. It follows that modulo the action of $S_n$ on the rows each of these choices is uniquely represented by an $m$-partition of $n$. Therefore the statements in Proposition 11 amount to saying that all of the mentioned global properties can be checked by verifying them on each of the $k \cdot \kappa(f)$-dimensional subspaces corresponding to the various $\kappa(f)$-partitions of $n$. The number of such partitions can be bounded by $n^{\kappa(f)}$, so if $\kappa(f)$ can be bounded by a quantity independent of $n$ for a certain family of polynomials $f$, this implies that the complexity of testing for those global properties for $k$-symmetric polynomials grows only polynomially in the number of variables – in the contrast to the general case.

Our main result is now presented in the following Theorem.

**Theorem 14.** Let $w = (w_1, \ldots, w_k)$ be a $k$-tupel of positive integers. Let $d \geq \max\{2w_j \mid 1 \leq j \leq k\}$ and let $f \in \mathbb{R}[V^k]^{S_n}$ be a $k$-symmetric polynomial of $w$-weighted degree at most $d$. Then

$$\kappa(f) \leq \prod_{j=1}^k \left\lfloor \frac{d}{w_j} \right\rfloor.$$

**Proof.** Set $d_j := \lfloor d/w_j \rfloor$ and $\mu := \prod_{j=1}^k d_j$. Define $g := \sum_{i=1}^n \sum_{j=1}^k X_{ij}^{d_j+1}$. By Proposition 11 it suffices to show that $\kappa(f) \leq \mu$ for all $\varepsilon > 0$, where $f_\varepsilon := f + \varepsilon g$.

Fix $r > 0$ and let $\varepsilon > 0$. Consider a point $z^* \in B_r$ where $f_\varepsilon|_{B_r}$ is minimized. We have to show $z^* \in A_\mu$.

Henceforth, we fix an order on $N_d = \{\alpha \in \mathbb{N}^k \mid |\alpha|_w \leq d\}$ and on $I_{n,k} := \{1, \ldots, n\} \times \{1, \ldots, k\}$, respectively. Define $p := \sum_{i=1}^n \sum_{j=1}^k X_{ij}^2$ and denote

$$\nabla : \mathbb{R}[V^k] \to \mathbb{R}[V^k]^{1 \times nk}, h \mapsto (\partial_{11} h, \ldots, \partial_{nk} h).$$

Since $z^*$ is a minimum point of $f_\varepsilon$ on $B_r = \{x \in \mathbb{R}^{n \times k} \mid p(x) = r\}$ and $\nabla p(z^*)$ is not zero, there exists a Lagrange multiplier $\lambda \in \mathbb{R}$ such that

$$\nabla f_\varepsilon(z^*) + \lambda \nabla p(z^*) = 0.$$

Since $f, p \in \mathbb{R}[V^k]^{S_n}$, by Theorem 6 there are polynomials $F, P \in \mathbb{R}[(Z_\alpha)_{\alpha \in N_d}]$ such that $f = F((p_\alpha)_{\alpha \in N_d})$ and $p = P((p_\alpha)_{\alpha \in N_d})$. We define the matrix polynomial
Let $c \in \mathbb{R}^{[N_d]}$ be such that $c_\alpha := p_\alpha(z^*)$. Then we can write (3.1) as
\[ \varepsilon \nabla g(z^*) + (\nabla F(c) + \lambda \nabla P(c)) M(z^*) = 0. \]
In other words, if we set $u := (\nabla F(c) + \lambda \nabla P(c))^T \in \mathbb{R}^{[N_d]}$, then $z^*$ is a solution to the system of polynomial equations
\[ \varepsilon \nabla g(X) + u^T M(X) = 0. \]
By Proposition 11 there are polynomials $\tilde{q}_1, \ldots, \tilde{q}_k \in \mathbb{R}[Y_1, \ldots, Y_k]$ such that the $(i, j)$-th entry of $u^T M(X)$ is $\tilde{q}_j(X_i)$ and $\deg_w(\tilde{q}_j) \leq d_j - w_j$. Thus, (3.2) can be rewritten as $\tilde{q}_j(X_i) = 0$ for all $(i, j) \in I_{n,k}$, where
\[ q_j := \varepsilon (d_j + 1) Y_j^{d_j} + \tilde{q}_j \in \mathbb{R}[Y_1, \ldots, Y_k] \quad \text{for all } j \in \{1, \ldots, k\}. \]
In other words, each row $z_1^*, \ldots, z_k^*$ of $z^*$ is a point of the complex zero set
\[ \mathcal{V}(q_1, \ldots, q_k) = \{ y \in \mathbb{C}^k \mid q_1(y) = 0, \ldots, q_k(y) = 0 \} \subset \mathbb{C}^k. \]
However, the coordinate ring $\mathbb{C}[Y_1, \ldots, Y_k]/(q_1, \ldots, q_k)$ is a $\mathbb{C}$-vector space of dimension at most $\prod_{j=1}^k d_j = \mu$. Indeed, because of $\deg_w(Y_j^{d_j}) > \deg_w(\tilde{q}_j)$ all monomials $Y^\alpha$ with $\alpha_j > d_j$ for some $1 \leq j \leq k$ can be rewritten as a sum of monomials of smaller $w$-degree. Hence, $\# \mathcal{V}(q_1, \ldots, q_k) \leq \mu$, so at most $\mu$ of the rows $z_1^*, \ldots, z_k^*$ can be different, that is, $z^* \in A_\mu$. \hfill $\square$

**Remark 15.** Note that the proof of Theorem 14 given above works in exactly the same manner for a more general setting: If $f \in C^1(V^k)$ and there exists $F \in C^1(\mathbb{R}^{[N_d]})$ such that $f = F((p_\alpha)_{\alpha \in N_d})$, then we get the same bound for $\kappa(f)$. Additionally, we can further extend this bound by Proposition 11 to functions of the form $f = F((p_\alpha)_{\alpha \in N_d})$ with $F \in C^0(\mathbb{R}^{[N_d]})$.

**Example 16.** We consider the case $k = 2$. The family of polynomials
\[ f_1^{m, \ell} := p_{(1,2)}^m p_{(1,1)}^2 - p_{(0,3)}^\ell \quad (m, \ell \in \mathbb{N}) \]
is not bounded with respect to any weighted degree. However, all $f_1^{m, \ell}$ can be written in terms of the power sum polynomials of (usual) degree at most 3. Thus, by the preceding remark, we still get the bound $\kappa(f_1^{m, \ell}) \leq 9$ for all $m, \ell \in \mathbb{N}$. Even for the rational function $f_2 := (p_{(3,0)} - 2p_{(1,2)}p_{(2,1)}) / (p_{(2,0)}^2 + 1)$ and the functions $f_3^m(x) := |\exp(f_2(x)) - 2| + f_1^{m,m}(x)$ ($m \in \mathbb{N}$) we get that the $\kappa$-value is at most 9.

**Remark 17.** Note that in the proof of Theorem 14 the bound for $\kappa(f)$ is given by an upper bound on the number of complex solutions of a certain system of polynomial equations. Since in fact we are only interested in the number of real solutions, $\mu$ might be chosen significantly smaller depending on the specific representation $F$. If for instance, $f$ is a polynomial with only sparse monomials but of high degree, it can be more appropriate to use results similar to Khovanskii’s fewnomial bound [21 Thm. 4.2] which then could give a much better estimation for $\kappa(f)$.
For a very rough estimation of $\kappa(f)$ we can just use the usual degree and get the following upper bound without any effort.

**Corollary 18.** If $f \in \mathbb{R}[V^k]^{S_n}$ is a $k$-symmetric polynomial of degree $d \geq 2$, then

$$\kappa(f) \leq d^k.$$ 

However, Theorem 14 is much stronger. Recall that, using Notation 5, the condition that $f$ is of $w$-degree at most $d$ can be expressed as $E_f \subset \{y \in \mathbb{R}^k_{\geq 0} \mid w^Ty \leq d\}$ and note that the hyperplane $\{y \in \mathbb{R}^k \mid w^Ty = d\}$ defining this enclosing simplex intersects the $j$-th coordinate axis at the point $(d/w_j)e_j$ (where $e_1, \ldots, e_k$ is the standard basis of $\mathbb{R}^k$). This gives the following geometric reformulation of the Theorem 14.

**Theorem 19.** Let $f \in \mathbb{R}[V^k]^{S_n}$ be a $k$-symmetric polynomial and let $a_1, \ldots, a_k \in \mathbb{Q}_{\geq 2}$ such that $E_f$ is contained in the simplex

$$\Delta(a_1, \ldots, a_k) := \text{conv}(0, a_1e_1, \ldots, a_ke_k) \subset \mathbb{R}^k.$$

Then $\kappa(f)$ is bounded by $\prod_{j=1}^k |a_j|$, i.e. the number of lattice points in $[0, a_1 - 1] \times \cdots \times [0, a_k - 1]$.

Note that this number of lattice points is approximately $k! \text{vol}(\Delta(a_1, \ldots, a_k))$. Therefore, finding a good bound for $\kappa(f)$ via Theorem 19 roughly amounts to finding the smallest simplex $\Delta(a_1, \ldots, a_k)$ enclosing $E_f$.

**Example 20** (Example 4 continued). Reconsider Figure 1. The triangle described by $w^{(2)} = (3, 5)$ minimizes the number of lattice points. Counting the lattice points in the rectangle drawn in Figure 2 gives the upper bound $\kappa(f) \leq 8$. This bound holds for all $n$ and for all choices of the parameters in Example 4.

In the case $k > 2$ drawing a picture and fitting the right simplex might not be that easy. However, we can provide a further bound on $\kappa(f)$. Note that this bound is better than the bound given in Corollary 18 in the case that the degree of some columns $X_j$ is much smaller than the others.
Corollary 21. Let $k \geq 2$ and let $f \in \mathbb{R}[V^k]^{S_n}$ be a $k$-symmetric polynomial such that each column $X_j$ occurs only with degree at most $d_j \geq 1$. Then
\[
\kappa(f) \leq k^k \prod_{j=1}^{k} d_j.
\]

Proof. Note that $E_f \subset [0, d_1] \times \cdots \times [0, d_k] \subset \Delta(kd_1, \ldots, kd_k)$ and use Theorem 19 to deduce the claim. \qed

4. Deciding Convexity of multisymmetric polynomials

In this section we apply Theorem 19 to the problem of algorithmically deciding (strict) convexity of a polynomial of fixed degree. Already for general quartic polynomials deciding convexity is an NP-hard problem (see [1]). However, we will see that in the case of symmetric (or, more generally, $k$-symmetric) polynomials of a bounded degree, a convexity test can be provided whose complexity is mainly determined by the degree in the sense that for a fixed degree it is polynomial in the number of variables. Recall that convexity of a function is defined as follows.

Definition 22. Let $C \subset \mathbb{R}^m$ be a convex set and $f : C \rightarrow \mathbb{R}$ be a real valued function. Then
\begin{enumerate}
  \item $f$ is called convex if
  \[
  \forall x_1, x_2 \in C, \forall t \in [0, 1] : f(t x_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2),
  \]
  \item $f$ is called strictly convex if
  \[
  \forall x_1 \neq x_2 \in C, \forall t \in (0, 1) : f(t x_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2).
  \]
\end{enumerate}

We remark that convexity is of particular interest also for the question of deciding if a polynomial is non-negative.

Proposition 23. Let $f \in \mathbb{R}[V^k]^{S_n}$. If $f$ is convex, then $\kappa(f) = 1$.

Proof. Let $\xi \in V^k$. Then
\[
\kappa(f) = \sum_{\sigma \in S_n} \frac{1}{n!} f(\sigma(\xi)) \geq f \left( \sum_{\sigma \in S_n} \frac{1}{n!} \sigma(\xi) \right) := \zeta.
\]

Since $\zeta \in A_1$, the statement follows. \qed

Using the classical definition above we remark the following characterizations which will allow the use of our main theorem.

Proposition 24. Let $f : \mathbb{R}^m \rightarrow \mathbb{R}$ be a polynomial function.
Proposition 25. Let \( f \in \mathbb{R}[V^k]^S_n \) be a \( k \)-symmetric polynomial. For \( i,j \in \{1, \ldots, k\} \) we consider \( \tau_{ij} : \mathbb{R}^k \to \mathbb{R}^k, x \mapsto x - e_i - e_j \) (where \( e_1, \ldots, e_k \) is the standard basis of \( \mathbb{R}^k \)).
\( \mathbb{R}^k \). Then we have

\[ E_{gf} \subset \bigcup_{i=1}^{k} \bigcup_{j=1}^{k} (\tau_{ij}(E_f) \cap \mathbb{R}^k_{\geq 0}) \times \{ e_i + e_j \} \subset \mathbb{R}^k \times \mathbb{R}^k. \]

In particular,

\[ E_{gf} \subset H_f \times \Delta(2, \ldots, 2), \]

where \( H_f := \bigcup_{i,j} \tau_{ij}(E_f) \cap \mathbb{R}^k_{\geq 0}. \)

**Proof.** This immediately follows by examining the definition of \( g_f \): Just note that if \( h \) is a second partial derivative of \( f \) with respect to variables in columns \( i \) and \( j \), then \( E_h \subset \tau_{ij}(E_f) \cap \mathbb{R}^k_{\geq 0}. \) \( \square \)

The above observations now yield the following Theorem.

**Theorem 26.** Let \( f \in \mathbb{R}[V^k]^S_n \) be a \( k \)-symmetric polynomial and let \( a_1, \ldots, a_k \in \mathbb{Q}_{\geq 1} \) be such that \( \Delta(a_1, \ldots, a_k) \) is a simplex enclosing \( H_f \) (as defined in Proposition 25). Then

\[ \kappa(g_f) \leq 3^k k! \prod_{j=1}^{k} [2a_j]. \]

In particular, \( \kappa(g_f) \leq 6^k k! \text{vol}(\Delta(a_1, \ldots, a_k)). \)

**Proof.** This follows directly from Theorem 19 with the preceding Proposition by noting that there exist \( \varepsilon, \delta \in (0, 1) \) such that \( \Delta(a_1, \ldots, a_k) \times \Delta(2, \ldots, 2) \) is contained in \( \Delta(2a_1 + \delta, \ldots, 2a_k + \delta, 4 - \varepsilon, \ldots, 4 - \varepsilon) \) and such that \( [2a_j + \delta] = [2a_j] \) for all \( j \in \{1, \ldots, k\}. \) \( \square \)

**Example 27** (Example 4 continued). Consider again the family of 2-symmetric polynomials \( f \) which was studied in Example 4. By reading \( E_f \) from Figure 1 we can easily construct \( H_f \) (see Proposition 24). We need to fit a \( w \)-degree line such that the originating triangle encloses \( H_f \). With a view to Theorem 26 we should choose a triangle \( \Delta(a_1, a_2) \) that minimizes \( [2a_1] \cdot [2a_2] \). With the choice indicated in Figure 3 we get \( a_1 < 2.5 \) and \( a_2 < 2. \) Hence, \( \kappa(g_f) \leq 108. \) However, it turns out that the bound can be improved by taking a closer look at \( E_{gf} \) using (4.1). By solving a small optimization problem we are able to find that \( E_{gf} \subset \Delta(9 - \varepsilon, 3 - \varepsilon, 4 - \varepsilon, 2 - \varepsilon) \) for a small \( \varepsilon > 0. \) From this observation one can deduce the better bound \( \kappa(f) \leq 48. \)

Apart from this geometrical view we provide a formulation in terms of weighted degrees. Note that this formulation is a bit weaker than Theorem 26 but it might still be useful as we need less information about the function in question in order to calculate the resulting bound.

**Theorem 28.** Let \( w = (w_1, \ldots, w_k) \) be a \( k \)-tupel of positive integers. Let \( d \geq 2 \min_{1 \leq j \leq k} w_j + \max_{1 \leq j \leq k} w_j \) and let \( f \in \mathbb{R}[V^k]^S_n \) be a \( k \)-symmetric polynomial of
Choosing all weights equal to 1 we get the following Corollary,

**Corollary 29.** Let \( f \in \mathbb{R}[V^k]^{S_n} \) be a \( k \)-symmetric polynomial of degree \( d \geq 3 \). Then

\[
\kappa(g_f) \leq 6^k(d - 2)^k.
\]

The following Corollary is just a reformulation of our results, which we include to emphasize the results in the case of convexity of symmetric polynomials.

**Corollary 30.** Let \( f \in \mathbb{R}[V]^{S_n} \) be a symmetric polynomial of degree \( d \geq 3 \). Then \( f \) is a convex function if and only if it is convex on each of the subspaces of points with at most \( 6(d - 2) \) distinct coordinates.

**Acknowledgement**

The research presented in this article was carried out during the first and third author’s stay at the Aalto Science Institute during the summer internship program 2014. The authors would like to thank the Aalto Science Institute for the hospitality and support.

**References**

[1] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, and J. N. Tsitsiklis, \( \text{Np-hardness of deciding convexity of quartic polynomials and related problems} \), Mathematical Programming, 137 (2013), pp. 453–476.

[2] S. Basu, R. Pollack, and M.-F. Roy, On the combinatorial and algebraic complexity of quantifier elimination, Journal of the ACM (JACM), 43 (1996), pp. 1002–1045.
[3] S. Basu and C. Riener, *Bounding the equivariant Betti numbers and computing the generalized euler-poincaré characteristic of symmetric semi-algebraic sets*, arXiv preprint arXiv:1312.6582, (2013).

[4] E. Becker, V. Powers, and T. Wormann, *Deciding positivity of real polynomials*, Contemporary Mathematics, 253 (2000), pp. 19–24.

[5] G. Blekherman and C. Riener, *Symmetric nonnegative forms and sums of squares*, arXiv preprint arXiv:1205.3102, (2012).

[6] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*, (1997).

[7] E. Briand, *When is the algebra of multisymmetric polynomials generated by the elementary multisymmetric polynomials?*, Contributions to Algebra and Geometry, 45 (2004), pp. 353–368.

[8] M.-D. Choi, T.-Y. Lam, and B. Reznick, *Even symmetric sextics*, Mathematische Zeitschrift, 115 (1997), pp. 559–580.

[9] J. Dalbec, *Multisymmetric functions*, Beiträge Algebra Geom, 40 (1999), pp. 27–51.

[10] J. Damon, *A global weighted version of Bezout’s theorem*, in The Arnoldfest: Proceedings of a Conference in Honour of VI Arnold for His Sixtieth Birthday, vol. 24, Fields Communications Series, 1999, pp. 115–129.

[11] W. R. Harris, *Real even symmetric ternary forms*, Journal of Algebra, 222 (1999), pp. 204–245.

[12] F. Junker, *Die Relationen, welche zwischen den elementaren symmetrischen Functionen bestehen*, Mathematische Annalen, 38 (1891), pp. 91–114.

[13] ——, *über symmetrische Functionen von mehreren Reihen von Veränderlichen*, Mathematische Annalen, 43 (1893), pp. 225–270.

[14] ——, *Die symmetrischen Functionen und die Relationen zwischen den Elementarfunctionen derselben*, Mathematische Annalen, 45 (1894), pp. 1–84.

[15] K. G. Murty and S. N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Mathematical programming, 39 (1987), pp. 117–129.

[16] C. Procesi, *Positive symmetric functions*, Advances in Mathematics, 29 (1978), pp. 219–225.

[17] B. Reznick, *Forms derived from the arithmetic-geometric inequality*, Mathematische Annalen, 283 (1989), pp. 431–464.

[18] C. Riener, *On the degree and half-degree principle for symmetric polynomials*, Journal of Pure and Applied Algebra, 216 (2012), pp. 850–856.

[19] ——, *Symmetric semi-algebraic sets and non-negativity of symmetric polynomials*, arXiv preprint arXiv:1409.0699, (2013).

[20] L. Schläfli, *Über die Resultante eines Systemes mehrerer algebraischen Gleichungen: ein Beitrag zur Theorie der Elimination*, Denkschr. Kais. Akad. Wiss. Math.-Natur. Kl, 4 (1852), pp. 9–112.

[21] F. Sottile, *Real solutions to equations from geometry*, vol. 57, American Mathematical Soc., 2011.

[22] V. Timofte, *On the positivity of symmetric polynomial functions. Part i: General results*, Journal of Mathematical Analysis and Applications, 284 (2003), pp. 174–190.

[23] ——, *On the positivity of symmetric polynomial functions. Part ii: Lattice general results and positivity criteria for degrees 4 and 5*, Journal of mathematical analysis and applications, 304 (2005), pp. 652–667.

[24] F. Vaccario, *The ring of multisymmetric functions*, Ann. Inst. Fourier, Grenoble, 55 (2005), pp. 717–731.
Mathematisches Institut der Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany

E-mail address: goerlach@uni-bonn.de

Aalto Science Institute, Aalto University, PO Box 11000, FI-00076 Aalto, Finland

E-mail address: cordian.riener@aalto.fi

Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany.

E-mail address: tillmann.weisser@uni-konstanz.de