Minimax Theory and its Applications
Volume 4 (2019), No. 2, 329–339

An Elementary Proof for the Exact Relaxation for Rank One Moment Matrices in Multi-Polynomial SOS Relaxation

Shenglong Hu
Department of Mathematics, School of Science, Hangzhou Dianzi University,
Hangzhou 310018, China
shenglonghu@hdu.edu.cn

Received: October 7, 2018
Accepted: February 18, 2019

We present an elementary proof for the fact that an optimal rank one moment matrix in the multi-polynomial SOS relaxation gives an exact relaxation. This fact is a fundamental result in multi-polynomial SOS relaxation method for the class of multi-polynomial optimization problems. The multi-polynomial SOS relaxation method is designed by exploring the special structures of the class of multi-polynomial optimization problems, which has the advantage for giving an SDP with size about half of that for the classical SOS relaxation in the general formulation.

Keywords: Multi-polynomial, multi-form, SOS, exact relaxation.

2010 Mathematics Subject Classification: 15A18, 15A42, 15A69.

1. Introduction

Polynomial optimization problems (POPs) have been finding numerous applications in diverse areas, see [5, 6, 7, 8, 10, 11] and references therein. In some of the applications (e.g., tensor computations), the variables are naturally grouped into several sets, see [1, 4, 7, 8, 9] and references therein. Thus, it is naturally and beneficially to view the corresponding polynomials as *multi-polynomials*, rather than polynomials in the general formulation ignoring their special group-variable structures. Therefore, the class of polynomial optimization problems involving multi-polynomials should be studied in their distinct way. We will call this class of problems as *multi-polynomial optimization problems* (MPOPs). Accordingly, theory and methods for studying POPs should be further developed for MPOPs.

Hopefully, this short article serves two primary purposes. The first is formulating MPOPs explicitly and presenting the *multi-polynomial SOS relaxation* method for solving them, which originates from Nie [8], giving an invitation to draw people's particular attention to this class of polynomial optimization problems. We want to remark that MPOPs are prevalent in numerical multilinear algebra and applications related [1, 4]. The second is presenting an elementary proof for

ISSN 2199-1413 (printed), ISSN 2199-1421 (electronic) / $ 2.50

the cornerstone fact in multi-polynomial SOS relaxation that if the optimal matrix is rank one, then the relaxation is exact. This fact for the classical SOS relaxation method can be proved using flat extensions of truncated moment matrices [6]. Although MPOP is a special POP, this result does not follow directly from the existing one.

The rest of the article is organized as follows. An appetizer on why multi-polynomial SOS relaxation is superior to the classical SOS relaxation is given in Section 2. Section 3 confirms that all MPOPs can be studied within multi-forms. Section 4 presents the multi-polynomial SOS relaxation. Section 5 presents characterizations for the Veronese-Segre varieties, preparing for the proof of the main result. Section 6 establishes the bridge between the results in Sections 5 and moment matrices, and then gives the main result. Section 7 is a short conclusion to this article.

2. Multi-polynomial optimization problems

Given positive integers d and n, a form (a.k.a. homogeneous polynomial) f of degree d in n variables $\{x_1, \ldots, x_n\}$ is a polynomial with each monomial being of fixed degree d, i.e.,

$$f(\lambda \mathbf{x}) = \lambda^d f(\mathbf{x}) \text{ for all } \lambda \in \mathbb{R} \text{ and } \mathbf{x} \in \mathbb{R}^n.$$

Given positive integers $d_1, \ldots, d_s, n_1, \ldots, n_s$ and

$$\mathbf{x} := (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_s},$$

then a *multi-form* f of degree d_i with respect to the variables $\mathbf{x}^{(i)}$ is a polynomial with each monomial being of fixed degree d_i for the group variables $\{x_1^{(i)}, \ldots, x_{n_i}^{(i)}\}$,

$$f(\lambda_1 \mathbf{x}^{(1)}, \ldots, \lambda_s \mathbf{x}^{(s)}) = \prod_{i=1}^{s} \lambda_i^{d_i} f(\mathbf{x}) \text{ for all } \lambda \in \mathbb{R}^s \text{ and } \mathbf{x} \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_s}.$$

Definitely, we can treat a multi-form as a form in the whole variables $\mathbf{x}$ with degree $\sum_{i=1}^s d_i$. However, the particular group-variable structure of the multi-form is lost when doing so.

If we de-homogenize each group of the variables $\mathbf{x}^{(i)}$'s we get a polynomial with s grouped variables. Likewise, polynomials of this kind will be called *multi-polynomials*. Multi-polynomials arise in many applications, see [1, 4, 8, 9] and references therein. In these applications, it is more beneficial to treat them as multi-polynomials other than polynomials in the general formulation. We will show in the following one advantage when optimizing with such polynomials.

Consider the polynomial optimization problem

$$f_{\min} := \min f(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}), \quad \text{s.t.} \quad \mathbf{x} \in K \subseteq \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_s}, \tag{1}$$

where $K \subseteq \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_s}$ is a semi-algebraic set defined by polynomial equations and inequalities, e.g.,

$$K := \left\{ \mathbf{x} : \begin{array}{l} g_1(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}) = \cdots = g_p(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}) = 0, \\ g_j(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}) \leq 0 \ \text{ for all } j = p+1, \ldots, p+q \end{array} \right\}, \tag{2}$$

for some multi-polynomials $f, g_1, \ldots, g_{p+q}$.

Given a multi-polynomial h, let $\deg_{\mathrm{i}}(h)$ be the degree of the polynomial h with respect to the group variables $\mathbf{x}^{(i)}$ for all $i \in \{1, \ldots, s\}$. For all $i \in \{1, \ldots, s\}$, let d_i be the smallest nonnegative integer such that

$$\max\{\deg_{\mathrm{i}}(f), \deg_{\mathrm{i}}(g_1), \ldots, \deg_{\mathrm{i}}(g_{p+q})\} \leq 2d_i.$$

As a polynomial optimization problem, the classical SOS relaxation method can be applied to problem (1) directly (cf. [5, 10, 11]). The *primal approach* is by relaxing the truncated monomial vector of the variables $\mathbf{x}$ up to the degree $2d := 2\sum_{i=1}^{s} d_i$ by the *spectrahedral*

$$\{\mathbf{y} \colon M(\mathbf{y}) \succeq \mathbf{0}\}.$$

A well-understood theory and sophisticated methods on SOS relaxations for POPs have been establishing during the years [6]. The main difficulty is then moved to solving the resulting semidefinite programming problems (SDPs). The matrix size and the number of equality constraints in the resulting SDP determine the solvability of all existing SDP solvers. Unfortunately, all the SDP problems coming from SOS relaxations for POPs have the number of equality constraints in the magnitude $\mathcal{O}(m^2)$ with m the dimension of the moment matrix $M(\mathbf{y})$, since approximately $\mathcal{O}(m^2)$ linear equalities are needed to characterize the *Gram matrices*. While, the dimension of the matrix for the above method to problem (1) is

$$m := \binom{N+d}{d}, \text{ with } N = n_1 + \cdots + n_s. \tag{3}$$

The current major interior point method based SDP solvers, including SeDuMi [12], SDPA [2], SDPT3 [13], are all limited to matrices of moderately large dimensions, saying hundreds. Getting back to our MPOPs, or more concretely on bi-quadratic optimization problem [7] (i.e., $s = 2$ and $d_1 = d_2 = 1$), this method solves problems up to dimension $(15, 15)$, i.e., bi-quadratic polynomials with both groups being less than fifteen variables. Even with the much improved SDP solver SDPNAL [14], which can solve SDP problems of matrix sizes being two thousands, the sizes of the MPOPs are still limited. Taking also the bi-quadratic case as an example, it solves problems up to dimension $(30, 30)$.

This general method treats the s groups of variables $\mathbf{x}^{(i)}$ as a whole. As we can see, the natural group structure of the variables $\mathbf{x}$ and those in the polynomials f and g_i's are not explored. A better way is treating each group variables $\mathbf{x}^{(i)}$ independently. Since only the monomials

$$\{\mathbf{x}^{\alpha} \colon |\alpha^{(i)}| \leq 2d_i \text{ for } i = 1, \ldots, s\}$$

are involved in all the multi-polynomials, we only need to consider a moment vector corresponding to these monomials. In this article, we will present an SOS relaxation method, termed *multi-polynomial SOS relaxation*, by relaxing the moment vector mentioned through the spectrahedral (cf. Section 4)

$$\{\mathbf{z}\colon \mathcal{M}(\mathbf{z}) \succeq \mathbf{0}\}$$

with the matrix size being

$$\binom{n_1+d_1}{d_1}\cdots\binom{n_s+d_s}{d_s}. \tag{4}$$

The number in (4) is much smaller than that in (3). For the bi-quadratic case, this method with SDPNAL [14] can then solve problems up to dimension $(45, 45)$. We remark that the difficulty of solving (1) *is not linear* with respect to the dimensions $n_1, \ldots, n_s$. Actually, the improvement from $(30, 30)$ to $(45, 45)$ is much bigger than the intuition from 30 to 45.

Given the power of the multi-polynomial SOS relaxation approach, theoretical aspects parallel to the classical SOS relaxations should be addressed. The first fundamental result in the classical SOS relaxations for POPs is

If the optimal marix $M(\mathbf{y})$ is rank one, then the relaxation is exact.

This is very helpful to certify the optimality and happens typically in numerical computations, which distinguishes this method from the other global/local techniques for solving (1). Likewise, a first question for the multi-polynomial SOS relaxation is

Is the relaxation exact, if the optimal marix $\mathcal{M}(\mathbf{z})$ is rank one?

The answer is affirmative. We will give a detailed proof for this important result in the sequel.

3. Multi-form optimization

Consider the following MPOP

$$\begin{aligned} f_{\min} := \min\ & f(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}) \\ \text{s.t.}\ \ & g_0(\mathbf{x}) = 1, \\ & g_i(\mathbf{x}) = 0,\ \text{for all } i = 1, \ldots, p, \\ & g_i(\mathbf{x}) \le 0,\ \text{for all } i = p+1, \ldots, p+q, \end{aligned} \tag{5}$$

where f and g_i's are *multi-forms* with

$$2d_i = \deg_{\mathrm{i}}(f) = \deg_{\mathrm{i}}(g_j) \text{ for all } j = 0, \ldots, p+q, \text{ for all } i = 1, \ldots, s.$$

The problem is general, including problem (1). Given an MPOP as

$$\begin{aligned} f_{\min} := & \min f(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(s)}) \\ \text{s.t.} \;\; & g_i(\mathbf{x}) = 0, \text{ for all } i = 1, \ldots, p, \\ & g_i(\mathbf{x}) \leq 0, \text{ for all } i = p+1, \ldots, p+q. \end{aligned} \tag{6}$$

If it is not in the form as (5), then define $\mathbf{y}^{(i)} := \mathbf{x}^{(i)}/x_0^{(i)}$ and

$$g_0(\mathbf{x}) := \prod_{i=1}^{s} (x_0^{(i)})^{2d_i}.$$

It is easy to see that $\tilde{f} := g_0 f(\mathbf{y})$, g_0, $\tilde{g}_1 := g_0 g_1(\mathbf{y}), \ldots, \; \tilde{g}_{p+q} := g_0 g_{p+q}(\mathbf{y})$ are all multi-forms in the group variables $\tilde{\mathbf{x}} := (\tilde{\mathbf{x}}^{(1)}, ..., \tilde{\mathbf{x}}^{(s)})$ with $\tilde{\mathbf{x}}^{(i)} := (x_0^{(i)}, x_1^{(i)}, ..., x_{n_i}^{(i)})$ for all $i \in \{1, \ldots, s\}$. Consider the following instance of (5)

$$\begin{aligned} \tilde{f}_{\min} := & \min \tilde{f}(\tilde{\mathbf{x}}^{(1)}, \ldots, \tilde{\mathbf{x}}^{(s)}) \\ \text{s.t.} \;\; & g_0(\tilde{\mathbf{x}}) = 1, \\ & \tilde{g}_i(\tilde{\mathbf{x}}) = 0, \text{ for all } i = 1, \ldots, p, \\ & \tilde{g}_i(\tilde{\mathbf{x}}) \leq 0, \text{ for all } i = p+1, \ldots, p+q. \end{aligned} \tag{7}$$

In the following, we show that (6) and (7) are equivalent.

If $\mathbf{x}$ is a feasible point of (6) with objective function value $f(\mathbf{x})$, then with $x_0^{(i)} = 1$ for all $i = 1, \ldots, s$, $\tilde{\mathbf{x}}$ is a feasible point of (7) with the same objective function value.

Conversely, suppose that $\tilde{\mathbf{x}}$ is a feasible point of (7) with objective function value $\tilde{f}(\tilde{\mathbf{x}})$. By the constraint $g_0(\tilde{\mathbf{x}}) = 1$, we conclude that $x_0^{(i)} \neq 0$ for all $i \in \{1, \ldots, s\}$. Since each function is a multi-form with degree $2d_i$ with respect to $\tilde{\mathbf{x}}^{(i)}$, we can assume that $x_0^{(i)} > 0$ for all $i \in \{1, \ldots, s\}$. Define

$$\tilde{\mathbf{y}}^{(i)} := \tilde{\mathbf{x}}^{(i)}/x_0^{(i)} \;\text{ for all } i = 1, \ldots, s.$$

It is easy to see that
$$\tilde{f}(\tilde{\mathbf{y}}) = \frac{\tilde{f}(\tilde{\mathbf{x}})}{g_0(\tilde{\mathbf{x}})} = \tilde{f}(\tilde{\mathbf{x}}),$$

and
$$\tilde{g}_i(\tilde{\mathbf{y}}) = \frac{\tilde{g}_i(\tilde{\mathbf{x}})}{g_0(\tilde{\mathbf{x}})} = \tilde{g}_i(\tilde{\mathbf{x}}) \;\text{ for all } i = 0, 1, \ldots, p+q.$$

Thus, $\tilde{\mathbf{y}}$ is also a feasible point of (7) with the same objective function value as $\tilde{\mathbf{x}}$. On the other side, $y_0^{(i)} = 1$ for each $i \in \{1, \ldots, s\}$. Thus,

$$\tilde{f}(\tilde{\mathbf{y}}) = f(\mathbf{y}) \text{ and } \tilde{g}_i(\tilde{\mathbf{y}}) = g_i(\mathbf{y}) \;\text{ for all } i = 1, \ldots, p+q.$$

Consequently, a feasible point of (6) with the same objective function value as $\tilde{\mathbf{x}}$ is constructed. Thus, the equivalence follows.

4. SOS-relaxation

The promised multi-polynomial SOS relaxation will be given for (5). The multi-forms structure will simplify the notation considerably.

Given a vector $\mathbf{z} \in \mathbb{R}^n$ and a positive integer r, we let $\mathbf{z}^{[r]}$ be the vector of monomials of degree r in lexicographic order

$$\mathbf{z}^{[r]} := (z_1^r, z_1^{r-1}z_2, \ldots, z_1^{r-1}z_n, \ldots, z_2^r, \ldots, z_n^r)^\mathsf{T}. \tag{8}$$

It is easy to see that the length of the vector $\mathbf{z}^{[r]}$ is

$$\nu(r,n) := \binom{n+r-1}{r}.$$

For the group variables $\mathbf{x} = (\mathbf{x}^{(1)}, ..., \mathbf{x}^{(s)})$ and a vector of integers $\mathbf{d} := (d_1, ..., d_s)$, define $\mathbf{x}^{[\mathbf{d}]}$ as the Kronecker product of $\left(\mathbf{x}^{(i)}\right)^{[d_i]}$'s, i.e.,

$$\mathbf{x}^{[\mathbf{d}]} := \left(\mathbf{x}^{(1)}\right)^{[d_1]} \otimes \cdots \otimes \left(\mathbf{x}^{(s)}\right)^{[d_s]}.$$

Let $\mathbf{n} := (n_1, \ldots, n_s)$. Thus, the length of $\mathbf{x}^{[\mathbf{d}]}$ is $\nu(\mathbf{d}, \mathbf{n}) := \prod_{i=1}^{s} \nu(d_i, n_i)$.

For $\kappa \in \mathbb{N}^n$, define $|\kappa| := \sum_{i=1}^n \kappa_i$ and $\mathbf{z}^\kappa := \prod_{i=1}^{n} z_i^{\kappa_i}$.

For $\alpha \in \mathbb{N}^{n_1} \times \cdots \times \mathbb{N}^{n_s}$, define $\mathbf{x}^\alpha := \prod_{i=1}^{s} \left(\mathbf{x}^{(i)}\right)^{\alpha^{(i)}}$.

Let $\Lambda(\mathbf{d}) := \{\alpha \in \mathbb{N}^{n_1} \times \cdots \times \mathbb{N}^{n_s} : |\alpha^{(i)}| = d_i \text{ for all } i = 1, \ldots, s\}$.

Then there exists a set of 0/1 symmetric matrices A_α such that

$$\left(\mathbf{x}^{[\mathbf{d}]}\right)^\mathsf{T} \mathbf{x}^{[\mathbf{d}]} = \sum_{\alpha \in \Lambda(2\mathbf{d})} A_\alpha \mathbf{x}^\alpha.$$

Given a vector $\mathbf{y} \in \mathbb{R}^{\nu(2\mathbf{d},\mathbf{n})}$, indexed by $\Lambda(2\mathbf{d})$, we can therefore formulate the moment matrix $\mathcal{M}(\mathbf{y})$ as

$$\mathcal{M}(\mathbf{y}) := \sum_{\alpha \in \Lambda(2\mathbf{d})} y_\alpha A_\alpha. \tag{9}$$

With these preparations, a natural SOS relaxation of problem (5) is

$$\begin{aligned} f_{sos} := \min & \sum_{\alpha \in \Lambda(2\mathbf{d})} f_\alpha y_\alpha, \\ \text{s.t.} & \sum_{\alpha \in \Lambda(2\mathbf{d})} (g_0)_\alpha y_\alpha = 1, \quad \sum_{\alpha \in \Lambda(2\mathbf{d})} (g_i)_\alpha y_\alpha = 0, \text{ for all } i = 1, \ldots, p, \\ & \sum_{\alpha \in \Lambda(2\mathbf{d})} (g_i)_\alpha y_\alpha \leq 0, \text{ for all } i = p+1, \ldots, p+q, \quad \mathcal{M}(\mathbf{y}) \succeq \mathbf{0}, \end{aligned} \tag{10}$$

where (f_α) and $\big((g_i)_\alpha\big)$'s are the coefficient vectors of the multi-forms f and g_i's under the monomial vector $\mathbf{x}^{[\mathbf{2d}]}$. For the sake of simplicity, (10) is a simplification of the usual SOS-relaxation using localizing matrices.

Obviously, if there is an optimal solution $\mathbf{y}^*$ of (10) with $\mathbf{y}^* = \mathbf{x}_*^{[\mathbf{2d}]}$ for some vector $\mathbf{x}_*$, then the relaxation is exact and $\mathbf{x}_*$ is an optimizer of the problem (5). We will show in Theorem 6.3 that this case is characterized by the optimal matrix $\mathcal{M}(\mathbf{y}^*)$ being rank one. To that end, characterizations for Veronese-Segre varieties will be given in Section 5.

5. Veronese-Segre varieties

Let n be a positive integer and $\mathbb{P}^{n-1}$ be the projective space of $\mathbb{R}^n$. The mapping $\mathrm{v_r} : \mathbb{P}^{n-1} \to \mathbb{P}^{\nu(r,n)-1}$ giving by

$$\mathrm{v_r}(\langle \mathbf{z}\rangle) = \langle \mathbf{z}^{[r]}\rangle, \tag{11}$$

where $\langle \mathbf{z}\rangle := \langle z_1 : \cdots : z_n\rangle \in \mathbb{P}^{n-1}$ is the *homogeneous coordinates* of the point $\langle \mathbf{z}\rangle$. This mapping is well known as the *r-th Veronese embedding*, see [3]. The image $V_r(\mathbb{P}^{n-1})$ of $\mathrm{v_r}$ on $\mathbb{P}^{n-1}$ is known as the *r-th Veronese variety.*

Let $\mathbf{n} = (n_1, \ldots, n_s)$ with positive n_i's. The *Segre embedding* is the mapping $\mathrm{s}_{\mathbf{n}} : \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_s-1} \to \mathbb{P}^{n_1\cdots n_s-1}$ giving by

$$\mathrm{s}_{\mathbf{n}}(\langle \mathbf{x}\rangle) = \langle \mathbf{x}^{(1)} \otimes \cdots \otimes \mathbf{x}^{(s)}\rangle. \tag{12}$$

The image of $\mathrm{s}_{\mathbf{n}}$ on $\mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_s-1}$ is known as the *Segre variety.*

A combination of the Veronese and Segre mappings is the *Veronese-Segre mapping* $\mathrm{vs}_{\mathbf{d},\mathbf{n}}$ giving by

$$\mathrm{vs}_{\mathbf{d},\mathbf{n}}(\langle \mathbf{x}\rangle) := \mathrm{s}_{(\nu(\mathrm{d_1,n_1}),\ldots,\nu(\mathrm{d_s,n_s}))}(\langle \big(\mathbf{x}^{(1)}\big)^{[d_1]} : \cdots : \big(\mathbf{x}^{(s)}\big)^{[d_s]}\rangle) = \langle \mathbf{x}^{[\mathbf{d}]}\rangle. \tag{13}$$

The image of the Veronese-Segre mapping $\mathrm{vs}_{\mathbf{d},\mathbf{n}}$ on $\mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_s-1}$ is denoted as $V_{\mathbf{d},\mathbf{n}}$, called the *Veronese-Segre variety.*

The 2-nd Veronese variety is the set of rank one symmetric real matrices. It is widely known that the set of rank one matrices is determined by the vanishing of its two by two minors. Surprisedly, this is also true for higher order Veronese varieties.

Lemma 5.1. *The Veronese variety $V_r(\mathbb{P}^{n-1})$ is given by*

$$\left\{ \langle \mathbf{y}\rangle \in \mathbb{P}^{\nu(r,n)-1} : \begin{array}{l} y_\alpha y_\beta - y_\gamma y_\eta = 0 \textit{ for all } \alpha+\beta = \gamma+\eta \\ \textit{with } \alpha,\beta,\gamma,\eta \in \Lambda(r) \end{array} \right\}. \tag{14}$$

Proof. Obviously, the Veronese variety $V_r(\mathbb{P}^{n-1})$ is contained in the set described by (14). In the following, we show that a vector in the set (14) also belongs to $V_r(\mathbb{P}^{n-1})$.

Let $\langle \mathbf{y}\rangle$ be a vector in the set (14). Then there exists a $\alpha \in \Lambda(r)$ such that $y_\alpha \neq 0$. Suppose without loss of generality that $\alpha_1 > 0$. Define

$$\langle \mathbf{z}\rangle := \langle y_\alpha : y_{\alpha-\mathbf{e}_1+\mathbf{e}_2} : \cdots : y_{\alpha-\mathbf{e}_1+\mathbf{e}_n}\rangle,$$

where $\mathbf{e}_i$ is the i-th standard basis vector for all $i \in \{1,\dots,n\}$. We claim that $\langle \mathbf{z}\rangle \in \mathbb{P}^{n-1}$ is well-defined and it is independent of choices of α. Actually, if $y_\beta \neq 0$ is chosen with $\beta_i > 0$, then the vector

$$\langle \mathbf{u}\rangle := \langle y_{\beta-\mathbf{e}_i+\mathbf{e}_1} : \cdots : y_\beta : \cdots : y_{\beta-\mathbf{e}_i+\mathbf{e}_n}\rangle$$

is indeed the same vector $\langle \mathbf{z}\rangle$, since

$$(\alpha - \mathbf{e}_1 + \mathbf{e}_j) + (\beta - \mathbf{e}_i + \mathbf{e}_k) = (\alpha - \mathbf{e}_1 + \mathbf{e}_k) + (\beta - \mathbf{e}_i + \mathbf{e}_j)$$

which implies by (14) $z_j u_k - z_k u_j = 0$ for all $j,k \in \{1,\dots,n\}$.

Consequently, the above process actually gives a regular mapping which is an inverse of the Veronese mapping. The proof is then complete. □

Proposition 5.2. *The Veronese-Segre variety $V_{\mathbf{d},\mathbf{n}}$ is given by*

$$\left\{ \langle \mathbf{y}\rangle \in \mathbb{P}^{\nu(\mathbf{d},\mathbf{n})-1} : \begin{array}{l} y_\alpha y_\beta - y_\gamma y_\eta = 0 \\ \textit{for all } \alpha+\beta=\gamma+\eta \textit{ with } \alpha,\beta,\gamma,\eta \in \Lambda(\mathbf{d}) \end{array} \right\}. \tag{15}$$

Proof. We proceed by mathematical induction on the length s of the vector $\mathbf{n}$. The case $s = 1$ follows from Lemma 5.1. In the following, we assume that the conclusion is true for $s-1$.

Let $\bar{\mathbf{d}} := (d_1,\dots,d_{s-1})$ and $\bar{\mathbf{n}} := (n_1,\dots,n_{s-1})$.

From (15), we first have that

$$y_{(\alpha^{(1)},\alpha^{(2)})}y_{(\beta^{(1)},\beta^{(2)})} - y_{(\beta^{(1)},\alpha^{(2)})}y_{(\alpha^{(1)},\beta^{(2)})}00$$

for any $\alpha^{(1)},\beta^{(1)} \in \Lambda(\bar{\mathbf{d}})$ and $\alpha^{(2)},\beta^{(2)} \in \Lambda(n_s)$. Thus, the $\nu(\bar{\mathbf{d}},\bar{\mathbf{n}}) \times \nu(d_s,n_s)$ matrix M with elements $m_{\alpha,\beta} := y_{(\alpha,\beta)}$ for $\alpha \in \Lambda(\bar{\mathbf{d}})$ and $\beta \in \Lambda(n_s)$ is a rank one matrix. Consequently, we conclude that $\mathbf{y} = \mathbf{u} \otimes \mathbf{v}$ for some $\mathbf{u} \in \mathbb{R}^{\nu(\bar{\mathbf{d}},\bar{\mathbf{n}})}$ and $\mathbf{v} \in \mathbb{R}^{\nu(d_s,n_s)}$.

By assumption, $\mathbf{y} \neq \mathbf{0}$. Then both $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$.

Suppose that $v_\beta \neq 0$ for some $\beta \in \Lambda(n_s)$. We then have from (15) that

$$y_{(\alpha,\beta)}y_{(\gamma,\beta)} - y_{(\kappa,\beta)}y_{(\eta,\beta)} = 0$$

for all $\alpha,\gamma,\kappa,\eta \in \Lambda(\bar{\mathbf{d}})$ such that $\alpha+\gamma = \kappa+\eta$. On the other hand,

$$y_{(\alpha,\beta)} = u_\alpha v_\beta \quad \text{for all } \alpha \in \Lambda(\bar{\mathbf{d}}).$$

Therefore, $u_\alpha u_\gamma - u_\kappa u_\eta = 0$ for all $\alpha,\gamma,\kappa,\eta \in \Lambda(\bar{\mathbf{d}})$ such that $\alpha+\gamma = \kappa+\eta$.

By the inductive hypothesis, we can conclude that $\langle \mathbf{u}\rangle \in V_{\bar{\mathbf{d}},\bar{\mathbf{n}}}$.

By symmetry we can conclude that $\langle \mathbf{v}\rangle \in V_{d_s}(\mathbb{P}^{n_s-1})$. The result follows. □

6. Duality and matrix representations

In this section, we will represent the quadratic relations in (15) by symmetric matrices, for linking this with the SDP problems of SOS-relaxations.

Let $\mathcal{S}^n$ be the space of $n\times n$ symmetric real matrices. There is a natural identification of $\mathcal{S}^n$ and the space of $\mathbb{R}[\mathbf{z}]_2$, real quadratic forms in n variables $\{z_1,\dots,z_n\}$. The dual of $\mathbb{R}[\mathbf{z}]$ is denoted as $\mathbb{R}[\partial]_2$. The dual pairing is given by differentiation, and it is denoted simply by $\bullet$. Since the dual of $\mathcal{S}^n$ can be identified as $\mathcal{S}^n$ itself, for convenience, we also identify the dual space $\mathbb{R}[\partial]_2$ as $\mathcal{S}^n$. Thus, a given matrix $A\in\mathcal{S}^n$ naturally gives both a quadratic form and also a quadratic relation.

It follows from (9) that the matrices in the set

$$\mathcal{A} := \{A_\alpha \colon \alpha\in\Lambda(2\mathbf{d})\}$$

are mutually orthogonal. Let $\xi(\mathbf{d},\mathbf{n}) := \frac{\nu(\mathbf{d},\mathbf{n})^2+\nu(\mathbf{d},\mathbf{n})}{2} - \nu(2\mathbf{d},\mathbf{n})$ and the set of symmetric matrices

$$\mathcal{C} := \{C_i \colon i=1,\dots,\xi(\mathbf{d},\mathbf{n})\}$$

be an orthogonal basis for the orthogonal complement of the linear subspace $\mathcal{M}_{\mathcal{A}}$ in $\mathcal{S}^{\nu(\mathbf{d},\mathbf{n})}$ generated by the set $\mathcal{A}$. It is easy to see that each matrix in the linear subspace $\mathcal{M}_{\mathcal{A}}$ is a moment matrix and vice versa. Dually, a symmetric matrix B is a moment matrix if and only if

$$\langle B, C_i\rangle = 0 \text{ for all } i\in\{1,\dots,\xi(\mathbf{d},\mathbf{n})\}.$$

Lemma 6.1. *Under the identification of $\mathcal{S}^{\nu(\mathbf{d},\mathbf{n})}$ and $\mathbb{R}_2[\mathbf{y}]$ with $\mathbf{y}\in\mathbb{R}^{\nu(\mathbf{d},\mathbf{n})}$, the linear subspace of $\mathcal{S}^{\nu(\mathbf{d},\mathbf{n})}$ generated by $\mathcal{C}$ is the linear subspace of $\mathbb{R}_2[\mathbf{y}]$ generated by*

$$\{y_\alpha y_\beta - y_\gamma y_\eta \colon\ \alpha+\beta=\gamma+\eta \text{ with } \alpha,\beta,\gamma,\eta\in\Lambda(\mathbf{d})\}. \tag{16}$$

Proof. It sufficient to show that the matrix representations of the quadratic forms in (16) forms a basis of the orthogonal complement of the linear subspace $\mathcal{M}_{\mathcal{A}}$ spanned by $\mathcal{A}$.

First note that each matrix A_α with $\alpha\in\Lambda(2\mathbf{d})$ is a matrix of 0/1's. Second, we have that (cf. (9))

$$\sum_{\alpha\in\Lambda(2\mathbf{d})} A_\alpha = E,$$

the matrix of all ones. Therefore, we have a natural decomposition of $\mathcal{S}^{\nu(\mathbf{d},\mathbf{n})}$ into a direct sum of $\nu(2\mathbf{d},\mathbf{n})$ linear subspaces of real symmetric matrices. Each linear subspace of the direct sum is characterized by a matrix A_α, with the linear subspace being defined when all the entries vary at all the positions of A_α with 1's. This linear subspace will be denoted as $\mathcal{L}_\alpha$. Thus,

$$\mathcal{S}^{\nu(\mathbf{d},\mathbf{n})} = \oplus_{\alpha\in\Lambda(2\mathbf{d})}\mathcal{L}_\alpha.$$

Inside each $\mathcal{L}_\alpha$, there is an orthogonal decomposition as

$$\mathcal{L}_\alpha = \mathbb{R}A_\alpha \oplus \mathcal{Q}_\alpha,$$

where $\mathcal{Q}_\alpha$ is the orthogonal complement of the one dimensional line $\mathbb{R}A_\alpha$ inside $\mathcal{L}_\alpha$. Since $\mathcal{Q}_\alpha$ is inside $\mathcal{L}_\alpha$, every index (β,γ) of its entry should satisfy

$$\beta + \gamma = \alpha.$$

Since A_α is a 0/1 matrix, immediately one see that the matrix representations of the quadratic forms

$$\{y_\beta y_\gamma - y_\eta y_\tau : \ \alpha + \gamma = \eta + \tau = \alpha\}$$

form a basis for $\mathcal{Q}_\alpha$. Consequently, the matrix representations in (16) form a basis for the linear subspace

$$\oplus_{\alpha\in\Lambda(2\mathbf{d})}\mathcal{Q}_\alpha,$$

which is the complement of the linear subspace $\mathcal{M}_{\mathcal{A}}$. The result then follows. $\square$

By the duality, we have the next result.

Proposition 6.2. *A vector $\mathbf{y} \in \mathbb{R}^{\nu(\mathbf{d},\mathbf{n})}$ is in Veronese-Segre variety $V_{\mathbf{d},\mathbf{n}}$ if and only if $\mathbf{y}^\mathsf{T} C_i \mathbf{y} = 0$ for all $i \in \{1,\dots,\xi(\mathbf{d},\mathbf{n})\}$.*

Theorem 6.3. *A vector $\mathbf{y} \in \mathbb{R}^{\nu(2\mathbf{d},\mathbf{n})}$ is a monomial vector $\mathbf{x}^{[2\mathbf{d}]}$ for some vector $\mathbf{x}$ if and only if the moment matrix $\mathcal{M}(\mathbf{y})$ is of rank one.*

Proof. The necessity is obvious from the definitions. For the sufficiency, suppose that $\mathcal{M}(\mathbf{y})$ is of rank one. Let $\mathcal{M}(\mathbf{y}) = \mathbf{q}^\mathsf{T}\mathbf{q}$ for some vector $\mathbf{q} \in \mathbb{R}^{\nu(\mathbf{d},\mathbf{n})}$. We have

$$\mathbf{q}^\mathsf{T}\mathbf{q} = \sum_{\alpha\in\Lambda(2\mathbf{d})} y_\alpha A_\alpha.$$

Consequently, $\mathbf{q}^\mathsf{T} C_i \mathbf{q} = 0$ for all $i \in \{1,\dots,\xi(\mathbf{d},\mathbf{n})\}$. It then follows from Proposition 6.2 that $\mathbf{q} \in V_{\mathbf{d},\mathbf{n}}$, i.e., $\mathbf{q} = \mathbf{x}^{[\mathbf{d}]}$ for some $\mathbf{x}$. Thus, $\mathbf{y} = \mathbf{x}^{[2\mathbf{d}]}$. $\square$

Theorem 6.3 shows that the relaxation (10) is exact if there is a rank one optimal matrix $\mathcal{M}(\mathbf{y}^*)$.

7. Conclusions

Theorem 6.3 can be proved in a higher level by toric varieties, as claimed in [11]. The elementary proof in this article should provide a more clear understanding for non-experts in algebraic geometry, especially in the numerical optimization community. More importantly, this article suggests that questions on multi-polynomial SOS relaxation method should be investigated.

Acknowledgement. This work is partially supported by National Science Foundation of China (Grant No. 11771328), Young Elite Scientists Sponsorship Program by Tianjin, the Natural Science Foundation of Zhejiang Province, China

(Grant No. LD19A010002), and Innovation Research Foundation of Tianjin University (Grant Nos. 2017XZC-0084 and 2017XRG-0015).

References

[1] L. De Lathauwer, B. De Moor, J. Vandewalle: *On the best rank-1 and rank-$(R_1, R_2, \dots, R_N)$ approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl. 21 (2000) 1324–1342.

[2] K. Fujisawa, Y. Futakata, M. Kojima, S. Matsuyama, S. Nakamura, K. Nakata, M. Yamashita: SDPA-M (SemiDefinite Programming Algorithm in MATLAB), http://homepage.mac.com/klabtitech/sdpa-homepage/download.html.

[3] R. Hartshorne: *Algebraic Geometry*, Springer, New York (1977).

[4] T. G. Kolda, B. W. Bader: *Tensor decompositions and applications*, SIAM Review 51(3) (2009) 455–500.

[5] J. Lasserre: *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. 11 (2001) 796–817.

[6] M. Laurent: *Sums of squares, moment matrices and optimization over polynomials*, in: *Emerging Applications of Algebraic Geometry*, M. Putinar and S. Sullivant (eds.), IMA Vol. Math. Appl. 149, Springer, New York (2009) 157–270.

[7] C. Ling, J. Nie, L. Qi, Y. Ye: *Biquadratic optimization over unit spheres and semidefinite programming relaxations*, SIAM J. Optim. 20 (2009) 1286–1310.

[8] J. Nie: *Sums of squares methods for minimizing polynomial forms over spheres and hypersurfaces*, Front. Math. China 7 (2012) 321–346.

[9] J. Nie, L. Wang: *Semidefinite relaxations for best rank-1 tensor approximations*, SIAM J. Matrix Analysis Appl. 35 (2014) 1155–1179.

[10] P. Parrilo: *Semidefinite programming relaxations for semialgebraic problems*, Math. Program, Ser. B 96 (2003) 293–320.

[11] P. Parrilo, B. Sturmfels: *Minimizing polynomial functions*, in: Proc. DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science, S. Basu and L. Gonzalez-Vega (eds.), American Mathematical Society, Providence (2003) 83–100.

[12] J. F. Sturm: *SeDuMi 1.02: A Matlab toolbox for optimization over symmetric cones*, Optim. Methods Software 11&12 (1999) 625–653.

[13] K. C. Toh, M. J. Todd, R. H. Tutuncu: *SDPT3: A Matlab software package for semidefinite programming*, Optim. Methods Software 11 (1999) 545–581.

[14] X. Y. Zhao, D. F. Sun, K. C. Toh: *A Newton-CG augmented Lagrangian method for semidefinite programming*, SIAM J. Optimization 20 (2010) 1737–1765.