数理解析研究所講究録
第 1773 巻 2012 年 36-46

A PERTURBATION THEOREM ON POLYNOMIAL OPTIMIZATION AND ITS EXTENSIONS

Masakazu Muramatsu, and Hayato Waki

Department of Communication Engineering and Informatics,
The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585 JAPAN.

Abstract

We prove a property of positive polynomials on a compact set with a small perturbation. When applied to a POP, the property implies that the optimal value of the corresponding SDP relaxation with sufficiently large relaxation order is bounded from below by $f^{*}-\epsilon$ and from above by $f^{*}+\epsilon(n+1)$, where f^{*} is the optimal value of the POP. This may help us to understand the funny phenomena of SDP relaxations for polynomial optimization problems observed in [4, 17, 18].

1 Introduction

1.1 Lasserre's SDP relaxation for POP

We consider the POP:

$$\text{minimize } \ f(x) \ \text{ subject to } \ f_{i}(x)\geq 0\ (i=1,\ldots,m), \tag{1}$$

where $f,\ f_{1},\ldots,f_{m}:\mathbb{R}^{n}\rightarrow \mathbb{R}$ are polynomials. The feasible region is denoted by $K=\{x\in \mathbb{R}^{n}\mid f_{j}(x)\geq 0(j=1,\ldots,m)\}$. Then it is easy to see that the optimal value f^{*} can be represented as

$$f^{*}=\sup\{\rho\mid f(x)-\rho\geq 0\ (\forall x\in K)\}.$$

We briefly describe the framework of the SDP realxation method for POP (1) proposed by Lasserre [8]. See also [13].

We denote the set of polynomials and sums of squares by $\mathbb{R}[x]$ and Σ, respectively. $\mathbb{R}[x]_{r}$ is the set of polynomials whose degree is less than or equals to r. We let $\Sigma_{r}=\Sigma\cap \mathbb{R}[x]_{2r}$. We define the quadratic module generated by $f_{1},\ldots,f_{m}$ as

$$M(f_{1},\ldots,f_{m})=\left\{\sigma_{0}+\sum_{j=1}^{m}\sigma_{j}f_{j}\mid\sigma_{0},\ldots,\sigma_{m}\in\Sigma\right\}.$$

The truncated quadratic module whose degree is less than or equal to $2r$ is defined by

$$M_{r}(f_{1},\ldots,f_{m})=\left\{\sigma_{0}+\sum_{i=1}^{m}\sigma_{j}f_{j}\mid\sigma_{0}\in\Sigma_{r},\sigma_{j}\in\Sigma_{r_{j}}(j=1,\ldots,m)\right\},$$

where $r_j = r - \lceil \deg f_j/2 \rceil$ for $j = 1, \dots, m$.

Substituting the condition that $f(x) - \rho$ is nonnegative by another relaxed condition that the polynomial is contained in $M_r(f_1, \dots, f_m)$, we obtain the following SOS relaxation problem:

$$\rho_r = \sup\{\, \rho \mid f(x) - \rho \in M_r(f_1, \dots, f_m) \,\}. \tag{2}$$

Lasserre[8] showed that $\rho_r \to f^*$ as $r \to \infty$ if $M(f_1, \dots, f_m)$ is Archimedean. See [12, 14] for the definition of Archimedean. In particluar, we point out that when $M(f_1, \dots, f_m)$ is Archimedean, then K is compact.

The problem (2) can be encoded to an SDP problem. Note that we can express a sum of squares $\sigma \in \Sigma_r$ by using a positive semidefinite matrix $X \in \mathbb{S}_+^{s(r)}$ as $\sigma(x) = u_r(x)^T X u_r(x)$, where $s(r) = \binom{n+r}{n}$ and $u_r(x)$ is a monomial vector which contains all the monomials in n variables up to degree r. By using this relation, the containment by $M_r(f_1, \dots, f_m)$ in (2), i.e.,

$$f - \rho = \sigma + \sum_{j=1}^{m} \sigma_j f_j,$$

can be transformed to equality constraints between semidefinite matrix variables corresponding to σ and σ_j's.

Note that, in this paper, we do not assume that K is compact nor that $M(f_1, \dots, f_m)$ is Archmedean. Still the framework of Lasserre's SDP relaxation can be applied to (1), althogh the good theoretical convergence property is lost.

1.2 Problems in the SDP relaxation for POP

Because POP is NP-hard, solving POP practically is sometimes extremely difficult. The SDP relaxation method described above also has some difficulty. The major difficulty consists in the size of the SDP relaxation problem (2). In fact, (2) contains $\binom{n+2r}{n}$ variables and $s(r) \times s(r)$ matrix. When n and/or r get larger, solving (2) is just impossible.

To overcome this difficulty, several techniques using sparsity of polynomials are proposed. See, e.g., [6, 15]. Based on the fact that most of the practical POPs are sparse in some sense, these techniques exploit special sparse structure of POPs to reduce the size and the number of variables of the SDP (2).

Another problem of the SDP relaxation is that (2) is often ill-posed. In [4, 17, 18], strange behaviors of SDP solvers are reported. Among them is that an SDP solver returns an 'optimal' value of (2) which is significantly different from the true optimal value without reporting no numerical errors. Even more strange is that the returned value by the SDP solver is nothing but the real optimal value of the POP (1). This is a 'super-accurate' property of the SDP relaxation for POP.

1.3 Contribution of this paper

We assume that there exists an optimal solution x^* of (1). Let

$$\begin{aligned} b &= \max(1, \max\{\, |x_i^*| \mid i = 1, \dots, n \,\}) \\ B &= [-b, b]^n. \end{aligned}$$

Obviously $x^* \in B$. We define:

$$\begin{aligned}
\bar{K} &= B \cap K \\
R_j &= \max\{\, |f_j(x)| \mid x \in B \,\} \ \ (j = 1, \ldots, m) \\
R &= \sum_{j=1}^{m} R_j.
\end{aligned}$$

Define also, for a positive integer r,

$$\begin{aligned}
\psi_r(x) &= -\sum_{j=1}^{m} f_j(x) \left(1 - \frac{f_j(x)}{R_j}\right)^{2r}, \\
\Theta_r(x) &= 1 + \sum_{i=1}^{n} x_i^{2r}, \\
\Theta_{r,b}(x) &= 1 + \sum_{i=1}^{n} \left(\frac{x_i}{b}\right)^{2r}.
\end{aligned}$$

We will prove the following theorem.

Theorem 1 *Suppose that for $\rho \in \mathbb{R}$, $f(x) - \rho > 0$ for every $x \in \bar{K}$, i.e., ρ is an lower bound of f^*.*

i. Then there exists $\tilde{r} \in \mathbb{N}$ such that for all $r \geq \tilde{r}$, $f - \rho + \psi_{\tilde{r}}$ is positive over B.

ii. In addition, for any $\epsilon > 0$, there exists a positive integer $\hat{r}$ such that, for every $r \geq \hat{r}$,

$$f - \rho + \epsilon\Theta_{r,b} + \psi_{\tilde{r}} \in \Sigma.$$

We remark that $\hat{r}$ depends on ρ and ϵ, while $\tilde{r}$ depends on ρ, but not ϵ.

The implication of this theorem is twofold. First, it elucidates the super-accurate property of the SDP relaxation for POPs. Notice that by construction, $-\psi_{\tilde{r}}(x) \in M_{\bar{r}}(f_1, \ldots, f_m)$ where $\bar{r} = \tilde{r} \max_j(\deg(f_j))$. Now assume that in (2), $r \geq \bar{r}$. Then, for any lower bound $\bar{\rho}$ of f^*, Theorem 1 means that $f - \bar{\rho} + \epsilon\Theta_{r,b} \in M_r(f_1, \ldots, f_m)$ for arbitrary small $\epsilon > 0$ and sufficiently large r. Such small perturbation is inevitably introduced everywhere in the floating point arithmetics which is used by the interior-point methods for solving the SDP relaxation problems. Note that we chose an artibrary lower bound of f^*, and in (2), the lower bound is being maximized. Therefore, we may obtain f^* due to the implicit perturbation introduced by the floating point arithmetics.

Second, we can use the result to construct new sparse SDP relaxations for POP (1). A naive idea is that we use (1) as is. Note that $-\psi_{\tilde{r}}(x)$ contains only monomials whose exponents are contained in

$$\cup_{j=1}^{m} \left(\mathcal{F}_j + \underbrace{\tilde{\mathcal{F}}_j + \cdots + \tilde{\mathcal{F}}_j}_{2\tilde{r}} \right),$$

where $\mathcal{F}_j$ is the *support* of the polynomial f_j, *i.e.*, the set of exponents of monomials with nonzero coefficients in f_j, and $\tilde{\mathcal{F}}_j = \mathcal{F}_j \cup \{0\}$. To state the idea more precisely, we introduce a notation. For a finite set $\mathcal{F} \subseteq \mathbb{N}^n$ and a positive integer r, we denote $r\mathcal{F} = \underbrace{\mathcal{F} + \cdots + \mathcal{F}}_{r}$ and

$$\Sigma(\mathcal{F}) = \left\{ \sum_{k=1}^{q} g_k(x)^2 \mid \mathrm{supp}(g_k) \subseteq \mathcal{F} \right\},$$

where $\mathrm{supp}(g_k)$ is the support of g_k. Note that $\Sigma(\mathcal{F})$ is the set of sums of squares of polynomials whose supports are contained by $\mathcal{F}$.

Now, fix an admissible error $\epsilon > 0$ and $\tilde{r}$ as in Theorem 1, and consider:

$$\hat{\rho}(\epsilon, \tilde{r}, r) = \sup \left\{ \rho \;\middle|\; f - \rho + \epsilon\Theta_{r,b} - \sum_{j=1}^{m} f_j \sigma_j = \sigma_0, \sigma_0 \in \Sigma_r, \sigma_j \in \Sigma(\tilde{r}\tilde{\mathcal{F}}_j) \right\} \tag{3}$$

for some $r \geq \tilde{r}$. Due to Theorem 1, (3) has a solution for a sufficiently large r.

Theorem 2 *For any $\epsilon > 0$, there exists $r \in \mathbb{N}$ such that $f^* - \epsilon \leq \hat{\rho}(\epsilon, \tilde{r}, r) \leq f^* + \epsilon(n+1)$.*

Proof: We apply Theorem 1 to POP (1) with $\rho = f^* - \epsilon$. Then for any $\epsilon > 0$, there exist $\hat{r}, \tilde{r} \in \mathbb{N}$ such that for every $r \geq \hat{r}$, $f - (f^* - \epsilon) + \epsilon\Theta_{r,b} + \psi_{\tilde{r}} \in \Sigma$. We choose a positive integer $r \geq \hat{r}$ which satisfies

$$r \geq \max\{\lceil \deg(f)/2 \rceil, \lceil (\tilde{r} + 1/2) \deg(f_1) \rceil, \ldots, \lceil (\tilde{r} + 1/2) \deg(f_m) \rceil\}. \tag{4}$$

Then there exists $\tilde{\sigma}_0 \in \Sigma_r$ such that $f - (f^* - \epsilon) + \epsilon\Theta_{r,b} + \psi_{\tilde{r}} = \tilde{\sigma}_0$ because the degree of the polynomial in the left hand side is equal to $2r$. We denote $\tilde{\sigma}_j := (1 - f_j/R_j)^{2\tilde{r}}$ for all j. The triplet $(f^* - \epsilon, \tilde{\sigma}_0, \tilde{\sigma}_j)$ is feasible in (3) because $(1 - f_j/R_j)^{2\tilde{r}} \in \Sigma(\tilde{r}\tilde{\mathcal{F}}_j)$. Therefore, we have $f^* - \epsilon \leq \hat{\rho}(\epsilon, \tilde{r}, r)$.

We prove that $\hat{\rho}(\epsilon, \tilde{r}, r) \leq f^* + \epsilon(n+1)$. We choose r as in (4) and consider the following POP:

$$\tilde{f} := \inf_{x \in \mathbb{R}^n} \{ f(x) + \epsilon\Theta_{r,b}(x) | \, f_1(x) \geq 0, \ldots, f_m(x) \geq 0 \}. \tag{5}$$

Apply Lasserre's SDP relaxation with relaxation order r to (5), we obtain the following SOS relaxation problem:

$$\hat{\rho}(\epsilon, r) := \sup \left\{ \rho \;\middle|\; f - \rho + \epsilon\Theta_{r,b} = \sigma_0 + \sum_{j=1}^{m} f_j \sigma_j, \sigma_0 \in \Sigma_r, \sigma_j \in \Sigma_{r_j} \right\}, \tag{6}$$

where $r_j := r - \lceil \deg(f_j)/2 \rceil$ for $j = 1, \ldots, m$. Then we have $\hat{\rho}(\epsilon, r) \geq \hat{\rho}(\epsilon, \tilde{r}, r)$ because $\Sigma(\tilde{r}\tilde{\mathcal{F}}_j) \subseteq \Sigma_{r_j}$ for all j. Indeed, it follows from (4) and definition of r_j that $r_j \geq \tilde{r}\deg(f_j)$, and thus $\Sigma(\tilde{r}\tilde{\mathcal{F}}_j) \subseteq \Sigma_{r_j}$.

The optimal solution x^* of POP (1) is feasible in (5) and the objective value is $f^* + \Theta_{r,b}(x^*)$. We have $f^* + \Theta_{r,b}(x^*) \geq \hat{\rho}(\epsilon, r)$ because (3) is the relaxation problem of (1).

In addition, it follows from $x^* \in B$ that $n+1 \geq \Theta_{r,b}(x^*)$, and thus $\hat{\rho}(\epsilon, \tilde{r}, r) \leq \hat{\rho}(\epsilon, r) \leq f^* + \epsilon(n+1)$. $\square$

In the above sparse relaxation (3), we have only to consider positive semidefinite matrices whose rows and columns correspond to $\tilde{r}\tilde{\mathcal{F}}_j$ for f_j. In contrast in Lasserre's SDP relaxation, we have to consider the whole set of monomials whose degree is less than or equals to r_j for each polynomial f_j. Only σ_0 is large; it contains all the set of monomials whose degree is less than or equals to r. However, because the other polynomials do not contain most of the monomials of σ_0, such monomials may be safely eliminated to reduce the size of σ_0 using the technique proposed in [6]. As a result, our sparse relaxation reduces the size of the matrix significantly if each $|\mathcal{F}_j|$ is small enough. We note that in most of the practical cases, in fact this is true.

Finally, we consider the case where the feasible region K is empty. If $-1 \in M(f_1, \ldots, f_m)$, then the feasible region K is empty. Indeed, there exist sums of square polynomials $\sigma_0, \ldots, \sigma_m$ such that $-1 = \sigma_0(x) + \sum_{j=1}^m \sigma_j(x) f_j(x)$. If K is non empty, then we obtain a contradiction by substituting $\tilde{x} \in K$ into this identity.

However, the converse does not hold in general. For instance, let $f_1 = x, f_2 = y, f_3 = -1 - xy$. Then K is empty, but $-1 \notin M(f_1, f_2, f_3)$. We can prove this fact by using a discussion in [14, Example 6.3.1].

The following result is directly obtained as a corollary of Theorem 1. We omit the proof.

Theorem 3 *We assume that K is empty. Then for any $\epsilon > 0$, there exists $\hat{r} \in \mathbb{N}$ such that for every $r \geq \hat{r}$, $-1 + \epsilon\Theta_r \in M(f_1, \ldots, f_m)$.*

The rest of this paper is organized as follows. In the next section, we prove Theorem 1. In Section 3, we give two extensions of Theorem 1. The former is a sparse version, and the latter, a symmetric cone version.

2 A Proof of Theorem 1

Lemma 4 *For any $\epsilon > 0$, there exists $\bar{r}$ such that for all $r \geq \bar{r}$ and $x \in B$, $\psi_r(x) \geq -\epsilon$ holds.*

Proof : We have

$$\begin{aligned}
\psi_r(x) &= -\sum_{j=1}^m f_j(x)\left(1 - \frac{f_j(x)}{R_j}\right)^{2r} \\
&= -\sum_{j:f_j(x)>0} f_j(x)\left(1 - \frac{f_j(x)}{R_j}\right)^{2r} - \sum_{j:f_j(x)<0} f_j(x)\left(1 - \frac{f_j(x)}{R_j}\right)^{2r} \\
&\geq -\sum_{j:f_j(x)>0} R_j \frac{f_j(x)}{R_j}\left(1 - \frac{f_j(x)}{R_j}\right)^{2r}.
\end{aligned}$$

Note that, for any $0 \le \lambda \le 1$,

$$\lambda(1-\lambda)^{2r} \le \left(\frac{2r}{2r+1}\right)^{2r} \frac{1}{2r+1} \le \frac{1}{2r+1} \le \frac{1}{2r},$$

and that if $r \ge R/(2\epsilon)$, then $\lambda(1-\lambda)^{2r} \le \epsilon/R$ on $\lambda \in [0,1]$. Therefore, for such r, we can further evaluate $\psi_r(x)$ as

$$\psi_r(x) \ge - \sum_{j:f_j(x)>0} R_j \frac{\epsilon}{R} \ge -\epsilon,$$

because $x \in B$ and $f_j(x) > 0$ imply $0 < f_j(x)/R_j \le 1$. This completes the proof. $\square$

Lemma 5 *Let $\tilde{x} \in B \setminus \bar{K}$, and $\kappa > 0$ be given. Then there exist $\tilde{\delta} > 0$ and $\tilde{r}$ such that for all $r \ge \tilde{r}$ and for any $x \in B(\tilde{x}, \tilde{\delta}) \cap B$, $\psi_r(x) \ge \kappa$ holds.*

Proof: For every $x \in B$, we have

$$\begin{aligned}
\psi_r(x) &= - \sum_{j:f_j(x)>0} f_j(x) \left(1 - \frac{f_j(x)}{R_j}\right)^{2r} - \sum_{j:f_j(x)<0} f_j(x) \left(1 - \frac{f_j(x)}{R_j}\right)^{2r} \\
&\ge - \sum_{j:f_j(x)>0} f_j(x) - \sum_{j:f_j(x)<0} f_j(x) \left(1 - \frac{f_j(x)}{R_j}\right)^{2r} \\
&\ge -R - \sum_{j:f_j(x)<0} f_j(x) \left(1 - \frac{f_j(x)}{R_j}\right)^{2r}.
\end{aligned}$$

Since $\tilde{x} \in B \setminus \bar{K}$, the minimum of $f_j(\tilde{x})$ over $j = 1, \ldots, m$ is negative. The continuity of polynomials implies that there exist $\delta > 0$ and $\bar{\lambda} < 0$ such that if $x \in B(\tilde{x}, \delta)$, then $\min_j f_j(x) \le \bar{\lambda}$. Then we can further evaluate $\psi_r(x)$ as

$$\psi_r(x) \ge -R - \bar{\lambda} \left(1 - \frac{\bar{\lambda}}{R}\right)^{2r}$$

for every $x \in B(\tilde{x}, \delta) \cap B$. Because $\bar{\lambda} < 0$ and $1 - \bar{\lambda}/R > 1$, there exists a positive integer $\tilde{r}$ such that $\psi_r(x) \ge \kappa$ for every $r \ge \tilde{r}$ and $x \in B(\tilde{x}, \delta) \cap B$. $\square$

Proof of (i) of Theorem 1 : Let $\tilde{x}^r$ be a minimizer of $f - \rho + \psi_r$ on B. We show the lemma by proving that there exists a positive integer $\tilde{r}$ such that $f(\tilde{x}^r) - \rho + \psi_r(\tilde{x}^r) > 0$ for every $r \ge \tilde{r}$.

Suppose to the contrary that for any $\tilde{r} > 0$, there exists r such that $f(\tilde{x}^r) - \rho + \psi_r(\tilde{x}^r) \le 0$. Because $\tilde{r}$ is arbitrary, the set $L = \{ r \mid f(\tilde{x}^r) - \rho + \psi_r(\tilde{x}^r) \le 0 \}$ is infinite. Since $\{ \tilde{x}^r \mid r \in L \} \subseteq B$, we can take an accumulation point $\tilde{x}^* \in B$ of $\{ \tilde{x}^r \mid r \in L \}$ and a subsequence $\{ \tilde{x}^r \mid r \in L' \}$ converging to $\tilde{x}^*$.

In the following, we will prove there exist a positive integer $\tilde{r}$ and a positive number $\tilde{\delta}$ such that $f(x) - \rho + \psi_r(x) > 0$ for every $x \in B(\tilde{x}^*, \tilde{\delta}) \cap B$ and $r \ge \tilde{r}$. Because

$\tilde{x}^r \in B(\tilde{x}^*, \tilde{\delta}) \cap B$ for sufficiently large $r \in L'$, this contradicts that $\tilde{x}^*$ is an accumulation point of $\{\, \tilde{x}^r \,|\, r \in L' \,\}$, establishing the lemma.

We first consider the case where $\tilde{x}^* \in \bar{K}$. Since $\bar{K}$ is compact, we can take $\epsilon > 0$ such that $f(x) - \rho \geq \epsilon$ for every $x \in \bar{K}$. Then, there exists a positive number $\tilde{\delta} > 0$ such that $f(x) - \rho \geq \epsilon/2$ for every $x \in B(\tilde{x}^*, \tilde{\delta})$. On the other hand, Lemma 4 implies that there exists $\tilde{r} > 0$ such that $\psi_r(x) \geq -\epsilon/4$ for every $r \geq \tilde{r}$ and $x \in B$. Therefore if $r \geq \tilde{r}$ and $x \in B(\tilde{x}^*, \tilde{\delta}) \cap B$, then $f(x) - \rho + \psi_r(x) \geq \epsilon/4 > 0$.

Next we consider the case where $\tilde{x}^* \in B \setminus \bar{K}$. Let $\kappa^* = -\inf\{\, f(x) - \rho \,|\, x \in B \,\} + 1$, which is finite because B is compact. Then Lemma 5 implies that there exist a positive number $\tilde{\delta}$ and a positive integer $\tilde{r}$ such that $\psi_r(x) \geq \kappa^*$ for every $x \in B(\tilde{x}^*, \tilde{\delta}) \cap B$ and $r \geq \tilde{r}$. For such x and r, we have $f(x) - \rho + \psi_r(x) \geq 1 > 0$. This completes the proof. □

Finally, to prove (ii) of Theorem 1, we need the following lemma estabilished by Lasserre and Netzer [11].

Lemma 6 *(Corollary 3.3 of [11]) Let $f \in \mathbb{R}[x]$ be a polynomial nonnegative on $[-1,1]^n$. For arbitrary $\epsilon > 0$, there exists some $\hat{r}$ such that for every $r \geq \hat{r}$, the polynomial $f + \epsilon\Theta_r$ is an SOS.*

Proof of (ii) of Theorem 1 : We have already proved that there exist $\tilde{r}$ such that for all $r \geq \tilde{r}$, $f(x) - \rho + \psi_{\tilde{r}}(x) > 0$ for every $x \in B = [-b, b]^n$. If we put $g(y) = f(by) - \rho + \psi_{\tilde{r}}(by)$, then $g(y) > 0$ over $[-1,1]^n$. Now Lemma 6 shows that for arbitrary $\epsilon > 0$, there exists $\hat{r}$ such that for every $r \geq \hat{r}$, $g(y) + \epsilon\Theta_r(y)$ is an SOS. Putting $by = x$, we conclude that $f + \epsilon\Theta_{r,b} + \psi_{\tilde{r}}$ is also an SOS. □

3 Extensions

In this section, we give two extensions of Theorem 1. The first extension is for POP with correlative sparsity. The second one is for POP over symmetric cones.

3.1 Extension to POP with correlative sparsity

In [15], the authors introduced the correlative sparsity for POP (1), proposed a sparse SDP relaxation that exploits the correlative sparsity, and demonstrated that the sparse SDP relaxation outperforms Lasserre's SDP relaxation. The sparse SDP relaxation is implemented in [16] and its source code is freely available.

We give the definition of the correlative sparsity for POP (1). For this, we use $n \times n$ symbolic symmetric matrix R, whose element is either 0 or $\star$ representing a nonzero value. We assign either 0 or $\star$ as follows:

$$R_{k,\ell} = \begin{cases} \star & \text{if } k = \ell, \\ \star & \text{if } \alpha_k \geq 1 \text{ and } \alpha_\ell \geq 1 \text{ for some } \alpha \in \mathcal{F}, \\ \star & \text{if } x_k \text{ and } x_\ell \text{ are involved in the polynomial } f_j \text{ for some } j = 1, \ldots, m, \\ 0 & \text{o.w.} \end{cases}$$

POP (1) is said to be correlatively sparse if the matrix R is sparse.

We give the detail of the sparse SDP relaxation proposed in [15]. We induce the undirected graph $G(V, E)$ from R. Here $V := \{1, \ldots, n\}$ and $E := \{(k, \ell) \mid R_{k,\ell} = \star\}$. After applying the chordal extension to $G(V, E)$, we generate all maximal cliques $C_1, \ldots, C_p$ of the extension $G(V, \tilde{E})$ with $E \subseteq \tilde{E}$. See [2, 15] and references therein for the detail of the chordal extension. For a finite set $C \subseteq \mathbb{N}$, x_C denotes the subvector which consists of x_i $(i \in C)$. For all $f_1, \ldots, f_m$ in POP (1), F_j denotes the set of indices whose variables are involved in f_j, *i.e.*, $F_j := \{i \in \{1, \ldots, n\} \mid \alpha_i \geq 1 \text{ for some } \alpha \in \mathcal{F}_j\}$. For a finite set $C \subseteq \mathbb{N}$, the sets $\Sigma_{r,C}$ and $\Sigma_{\infty,C}$ denote the subsets of Σ_r as follows:

$$\begin{aligned}
\Sigma_{r,C} &:= \left\{ \sum_{k=1}^{q} g_k(x)^2 \,\middle|\, \forall k = 1, \ldots, q, g_k \in \mathbb{R}[x_C]_r \right\}, \\
\Sigma_{\infty,C} &:= \bigcup_{r \geq 0} \Sigma_{r,C}.
\end{aligned}$$

Note that if $C = \{1, \ldots, n\}$, then we have $\Sigma_{r,C} = \Sigma_r$ and $\Sigma_{\infty,C} = \Sigma$. The sparse SDP relaxation problem with relaxation order r for POP (1) is obtained from the following SOS relaxation problem:

$$\rho_r^{\text{sparse}} := \sup \left\{ \rho \,\middle|\, \begin{array}{l} f - \rho = \sum_{h=1}^{p} \sigma_{0,h} + \sum_{j=1}^{m} \sigma_j f_j, \\ \sigma_{0,h} \in \Sigma_{r,C_h} \ (h = 1, \ldots, p), \sigma_j \in \Sigma_{r_j, D_j} \ (j = 1, \ldots, m) \end{array} \right\}, \tag{7}$$

where D_j is the union of some of the maximal cliques $C_1, \ldots, C_p$ such that $F_j \subseteq C_h$ and $r_j = r - \lceil \deg(f_j)/2 \rceil$ for $j = 1, \ldots, m$.

It should be noted that another sparse SDP relaxation is proposed in [3, 10, 12] and the asymptotic convergence is proved. In contrast, the convergence of the sparse SDP relaxation (7) is not shown although (7) is smaller than the SDP problem obtained by using the sparse SDP relaxation proposed in [3, 10, 12].

We give an extension of Theorem 1 into POP with correlative sparsity. If $C_1, \ldots, C_p \subseteq \{1, \ldots, n\}$ satisfy the following property, we refer this property as *the running intersection property* (RIP):

$$\forall h \in \{1, \ldots, p-1\}, \exists t \in \{1, \ldots, p\} \text{ such that } C_{h+1} \cap (C_1 \cup \cdots \cup C_h) \subsetneq C_t.$$

For $C_1, \ldots, C_p \subseteq \{1, \ldots, n\}$, we define sets $J_1, \ldots, J_p$ as follows:

$$J_h := \{j \in \{1, \ldots, m\} \,|\, f_j \in \mathbb{R}[x_{C_h}]\}.$$

Clearly, we have $\cup_{h=1}^{p} J_h = \{1, \ldots, m\}$. In addition, we define

$$\begin{aligned}
\psi_{r,h}(x) &:= -\sum_{j \in J_h} \left(1 - \frac{f_j(x)}{R_j}\right)^{2r}, \\
\Theta_{r,h,b}(x) &:= 1 + \sum_{i \in C_h} \left(\frac{x_i}{b}\right)^{2r}
\end{aligned}$$

for $h = 1, \ldots, p$.

By a similar proof of the theorem on convergence of the sparse SDP relaxation given in [3], we can establish the correlatively sparse case of Theorem 1. Indeed, we can obtain the theorem by using [3, Lemma 4] and Theorem 1.

Theorem 7 *Assume that nonempty sets $C_1,\dots,C_p \subseteq \{1,\dots,n\}$ satisfy (RIP) and we can decompose f into $f = \hat{f}_1+\cdots+\hat{f}_p$ with $\hat{f}_h \in \mathbb{R}[x_{C_h}]$ $(h=1,\dots,p)$. Under assumptions in Theorem 1, there exists $\tilde{r}\in\mathbb{N}$ such that for all $r\geq\tilde{r}$, $f-\rho+\sum_{h=1}^{p}\psi_{r,h}$ is positive over $B=[-b,b]^n$. In addition, for any $\epsilon>0$, there exists $\hat{r}\in\mathbb{N}$ such that for every $r\geq\hat{r}$,*

$$f-\rho+\epsilon\sum_{h=1}^{p}\Theta_{r,h,b}+\sum_{h=1}^{p}\psi_{\tilde{r},h}\in\Sigma_{\infty,C_1}+\cdots+\Sigma_{\infty,C_p}.$$

We remark that it could follow from Theorem 5 in [3] that Theorem 7 holds without the polynomial $\epsilon\sum_{h=1}^{p}\Theta_{r,h,b}$ if we assumed in Theorem 7 that all quadratic modules generated by f_j $(j\in C_h)$ for all $h=1,\dots,p$ are Archimedean. To prove Theorem 7, we describe Lemma 4 in [3].

Lemma 8 *([3, Lemma 4]) Assume that we decompose f into $f=\hat{f}_1+\cdots+\hat{f}_p$ with $\hat{f}_h\in\mathbb{R}[x_{C_h}]$ and $f>0$ on K. Then for any bounded set $B\subseteq\mathbb{R}^n$, there exist $\lambda\in(0,1]$, $r\in\mathbb{N}$ and $g_h\in\mathbb{R}[x_{C_h}]$ with $g_h>0$ on B such that*

$$f=\sum_{h=1}^{p}\sum_{j\in J_h}(1-\lambda f_j)^{2r}f_j+\sum_{h=1}^{p}g_h.$$

Proof of Theorem 7 : We choose $\bar{K}$ as B in Lemma 8 because $\bar{K}$ is compact. Applying Theorem 1 into g_h in Lemma 8, we obtain the desired result. □

3.2 Extension to POP with symmetric cones

In this subsection, we extend Theorem 1 into POP over symmetric cones, *i.e.*,

$$f^* := \inf_{x\in\mathbb{R}^n}\{f(x)\mid G(x)\in\mathcal{E}_+\}, \tag{8}$$

where $f\in\mathbb{R}[x]$, $\mathcal{E}_+$ is a symmetric cone associated with an N-dimensional Euclidean Jordan algebra $\mathcal{E}$, and G is $\mathcal{E}$-valued polynomial in x. The feasible region K of POP (8) is $\{x\in\mathbb{R}^n\mid G(x)\in\mathcal{E}_+\}$. Note that if $\mathcal{E}$ is $\mathbb{R}^m$ and $\mathcal{E}_+$ is the nonnegative orthant $\mathbb{R}^m_+$, then (8) is identical to (1). In addition, because $n\times n$ symmetric positive semidefinite cone $\mathbb{S}^n_+$ is a symmetric cone, the bilinear matrix inequalities can be formulated as (8).

To construct ψ_r for (8), we introduce some notation and symbols. The product and inner product of $x,y\in\mathcal{E}$ are, respectively, $x\circ y$ and $x\bullet y$. Let e be the identity element in the Jordan algebra $\mathcal{E}$. For any $x\in\mathcal{E}$, we have $e\circ x=x\circ e=x$. We can define eigenvalues for all elements in the Jordan algebra $\mathcal{E}$ as well as square matrices. See [1] for the detail. We construct ψ_r for (8) as follows:

$$\begin{aligned} M &:= \sup\left\{\text{maximum absolute eigenvalue of } G(x)\mid x\in\bar{K}\right\}, \\ \psi_r(x) &:= -G(x)\bullet\left(e-\frac{G(x)}{M}\right)^{2r}, \end{aligned} \tag{9}$$

where we define $x^k:=x^{k-1}\circ x$ for $k\in\mathbb{N}$ and $x\in\mathcal{E}$.

Lemma 4 in [7] shows that ψ_r defined in (9) has the same properties as in Lemmas 4 and 5. Therefore, we can extend Theorem 1 into POP (8).

Theorem 9 *For a given ρ, we assume that $f(x) - \rho > 0$ for every $x \in \bar{K}$. Then there exists $\tilde{r} \in \mathbb{N}$ such that for all $r \geq \tilde{r}$, $f - \rho + \psi_r$ is positive over B. In addition, for any $\epsilon > 0$, there exists $\hat{r} \in \mathbb{N}$ such that for every $r \geq \hat{r}$,*

$$f - \rho + \epsilon\Theta_{r,b} + \psi_{\tilde{r}} \in \Sigma.$$

References

[1] J. Faraut and A. Korányi: *Analysis on symmetic cones.* (Oxford University Press, New York, 1994)

[2] M. Fukuda, M. Kojima, K. Murota and K. Nakata: "Exploiting sparsity in semidefinite programming via matrix completion I: General framework". *SIAM Journal on Optimization,* 11 (2000), 647–674.

[3] D. Grimm, T. Netzer and M. Schweighofer: "A note on the representation of positive polynomials with structured sparsity". *Archiv der Mathematik,* 89 (2007), 399–403.

[4] H. Henrion, J. B. Lasserre: "Detecting global optimality and extracting solutions in GloptiPoly", in H. Henrion and A. Garulli (eds.) Positive Polynomials in Control. Lecture Notes on Control and Information Sciences, vol. 312, Springer, Berlin (2005).

[5] S. Kim, M. Kojima and H. Waki: "Generalized Lagrangian Duals and Sums of Squares Relaxations of Sparse Polynomial Optimization Problems". *SIAM Journal on Optimization,* 15 (2005), 697–719.

[6] M. Kojima, S. Kim and H. Waki: "Sparsity in sums of squares of polynomials". *Mathematical Programming,* 103 (2005), 45–62.

[7] M. Kojima and M. Muramatsu: "An extension of sums of squares relaxations to polynomial optimization problems over symmetric cones". *Mathematical Programming,* 110 (2007), 315–336.

[8] J. B. Lasserre: Global optimization with polynomials and the problems of moments. *SIAM Journal on Optimization,* 11 (2001), 796–817.

[9] J. B. Lasserre: "A sum of squares approximation of nonnegative polynomials". *SIAM Journal on Optimization,* 16 (2006), 751–765.

[10] J. B. Lasserre: "Convergent SDP-relaxations in polynomial optimization with sparsity". *SIAM Journal on Optimization,* 17 (2006), 822–843.

[11] J. B. Lasserre and T. Netzer: "SOS approximations of nonnegative polynomials via simple high degree perturbations". *Mathematische Zeitschrift,* 256 (2007), 99–112.

[12] M. Laurent: "Sums of squares, moment matrices and optimization over polynomials", In M. Putinar and S. Sullivant (eds.): *IMA Volume Emerging Applications of Algebraic Geometry,* 149 (Springer, Berlin, 2009), 157–270.

[13] P. A. Parrilo: "Semidefinite programming relaxations for semialgebraic problems". *Mathematical Programming*, 96 (2003), 293–320.

[14] A. Prestel and C. N. Delzell: *Positive Polynomials* (Springer-Verlag, Berlin, 2001).

[15] H. Waki, S. Kim, M. Kojima and M. Muramatsu: "Sums of Squares and Semidefinite Programming Relaxations with Structured Sparsity". *SIAM Journal on Optimization*, 17 (2006), 218–242.

[16] H. Waki, S. Kim, M. Kojima, M. Muramatsu and H. Sugimoto, "SparsePOP : a Sparse Semidefinite Programming Relaxation of Polynomial Optimization Problems". *ACM Transactions on Mathematical Software*, 15 (2008) 15:1–15:13.

[17] H. Waki, M. Nakata and M. Muramatsu: "Strange Behaviors of Interior-point Methods for Solving Semidefinite Programming Problems in Polynomial Optimization". *To appear in Computational Optimization and Applications*, (DOI: 10.1007/s10589-011-9437-8).

[18] H. Waki: "How to generate weakly infeasible semidefinite programs via Lasserre's relaxations for polynomial optimization". *To appear in Optimization Letters*, (DOI: 10.1007/s11590-011-0384-1).