Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 15, No. 3, pp. 415 – 439, 2019

# LAGRANGIAN-CONIC RELAXATIONS, PART II: APPLICATIONS TO POLYNOMIAL OPTIMIZATION PROBLEMS*

Naohiko Arima†, Sunyoung Kim‡, Masakazu Kojima§ and Kim-Chuan Toh¶

**Abstract:** We present a moment cone (MC) relaxation and a hierarchy of Lagrangian-SDP relaxations of polynomial optimization problems (POPs) using the unified framework established in Part I. The MC relaxation is derived for a POP of minimizing a polynomial subject to a nonconvex cone constraint and polynomial equality constraints. It is an extension of the completely positive programming relaxation for QOPs. Under a copositivity condition, we characterize the equivalence of the optimal values between the POP and its MC relaxation. A hierarchy of Lagrangian-SDP relaxations, which is parameterized by a positive integer $\omega$, is proposed for an equality constrained POP. It is obtained by combining Lasserre's hierarchy of SDP relaxation of POPs and the basic idea behind the conic and Lagrangian-conic relaxations from the unified framework. We prove under a certain assumption that the optimal value of the Lagrangian-SDP relaxation with the Lagrangian multiplier $\lambda$ and the relaxation order $\omega$ in the hierarchy converges to that of the POP as $\lambda \to \infty$ and $\omega \to \infty$. The hierarchy of Lagrangian-SDP relaxations is designed to be used in combination with the bisection and 1-dimensional Newton methods, which was proposed in Part I, for solving large-scale POPs efficiently and effectively.

**Key words:** *polynomial optimization problem, moment cone relaxation, SOS relaxation, a hierarchy of the Lagrangian-SDP relaxations*

**Mathematics Subject Classification:** *90C20, 90C25, 90C26*

## 1 Introduction

A unified framework for conic and Lagrangian-conic relaxations in Part I [4] was proposed for quadratic optimization problems (QOPs) and polynomial optimization problems (POPs). Let $\mathbb{V}$ be a finite dimensional vector space endowed with an inner product $\langle \cdot, \cdot \rangle$, $\mathbb{K} \subset \mathbb{V}$ a (not necessarily convex) cone, $\boldsymbol{H}^0 \in \mathbb{V}$ and $\boldsymbol{Q}^k \in \mathbb{V}$ $(k = 0, 1, 2, \ldots, m)$. The framework is build on the following primal-dual conic optimization problems (COPs):

$$\zeta^p(\mathbb{K}) \quad := \quad \inf \left\{ \langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle \;\middle|\; \begin{array}{l} \boldsymbol{X} \in \mathbb{K},\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1, \\ \langle \boldsymbol{Q}^k, \boldsymbol{X} \rangle = 0 \ (k = 1, 2, \ldots, m) \end{array} \right\} \tag{1.1}$$

$$\zeta^d(\mathbb{K}) \quad := \quad \sup \left\{ z_0 \;\middle|\; \boldsymbol{Q}^0 + \sum_{k=1}^{m} \boldsymbol{Q}^k z_k - \boldsymbol{H}^0 z_0 \in \mathbb{K}^* \right\}. \tag{1.2}$$

*The research of this author was partially supported by the Japan Science and Technology Agency (JST), the Core Research of Evolutionary Science and Technology (CREST) Research Project.

Here $\mathbb{K}^* = \{\boldsymbol{Y} \in \mathbb{V} : \langle \boldsymbol{X}, \boldsymbol{Y} \rangle \geq 0 \text{ for every } \boldsymbol{X} \in \mathbb{K}\}$ (the dual of $\mathbb{K}$). Let $F(\mathbb{K})$ denote the feasible region $\left\{\boldsymbol{X} \in \mathbb{K},\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1, \langle \boldsymbol{Q}^k, \boldsymbol{X} \rangle = 0 \ (k = 1, 2, \ldots, m)\right\}$ of the primal COP (1.1). As a basic assumption, we impose **Condition (I)**: $F(\mathbb{K}) \neq \emptyset$, $\boldsymbol{H}^0 \in \mathbb{K}^*$ and $\boldsymbol{Q}^k \in \mathbb{K}^*$ $(k = 1, 2, \ldots, m)$.

From a theoretical viewpoint, this framework is intended to unify and generalize the existing results on the completely positive programming (CPP) relaxation of QOPs [1, 9, 10, 17] and its extension to POPs, which was called a moment cone (MC) relaxation in [3]. From a practical viewpoint, the Lagrangian-conic relaxation is proposed to solve large-scale COPs obtained from effective conic relaxations, including doubly nonnegative (DNN) and semidefinite programming (SDP) relaxations of QOPs and POPs. In particular, the CPP relaxation and the sparse DNN relaxation for QOPs in Part I [4] were discussed from these viewpoints.

We consider a general class of POPs of the following form:

$$\zeta^* \quad := \quad \inf\ \left\{ f^0(\boldsymbol{x}) \ \middle|\ \boldsymbol{x} \in \mathbb{J},\ f^k(\boldsymbol{x}) = 0 \ (k = 1, 2, \ldots, m) \right\}, \tag{1.3}$$

where $\mathbb{J}$ denotes a closed (but not necessarily convex) cone in the $n$-dimensional Euclidean space $\mathbb{R}^n$, and $f^k(\boldsymbol{x})$ a real valued polynomial in $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ $(k = 0, 1, 2, \ldots, m)$. In practical applications, $\mathbb{J}$ can be $\mathbb{R}^n$ or $\mathbb{R}^n_+$ (the nonnegative orthant of $\mathbb{R}^n$). We note that even when $\mathbb{J} = \mathbb{R}^n$, any inequality constraint $h(\boldsymbol{x}) \leq 0$ can be incorporated into POP (1.3) if it is converted to an equivalent equality constraint $h(\boldsymbol{x}) + u^2 = 0$ with a slack variable $u \in \mathbb{R}$.

## 1.1 Contribution

As the first application of the framework to POP (1.3), we derive a COP of the form (1.1) with a closed convex cone $\mathbb{K}$, which was introduced by Arima, Kim and Kojima [3], as the MC relaxation of POP. The main emphasis of our discussion here is on a unified treatment of the CPP relaxation of QOPs and their MC relaxation of POPs. More precisely, we construct a nonconvex cone $\boldsymbol{\Gamma}$ in the space $\mathbb{V}$ of symmetric matrices with an appropriate dimension, and symmetric matrices $\boldsymbol{H}^0$, $\boldsymbol{Q}^k \in \mathbb{V}$ $(k = 0, 1, 2, \ldots, m)$ so that COP (1.1) represents POP (1.3), *i.e.*, $\zeta^* = \zeta^p(\boldsymbol{\Gamma})$. Then, we apply the convexification procedure discussed in Part I [4] to COP (1.1) with $\mathbb{K} = \boldsymbol{\Gamma}$ and derive COP (1.1) with $\mathbb{K} = \text{co}\, \boldsymbol{\Gamma}$ for the MC relaxation of POP (1.3). We provide a necessary and sufficient condition for the equivalence the two COPs, *i.e.,* $\zeta^* = \zeta^p(\boldsymbol{\Gamma}) = \zeta^p(\text{co}\, \boldsymbol{\Gamma})$, under Condition (I). This is one of the main theoretical contribution of the paper.

The other contribution of the paper is to propose a hierarchy of Lagrangian-SDP relaxations for POP (1.3) with $\mathbb{J} = \mathbb{R}^n$ by combining Lasserre's hierarchy of SDP relaxations [20] for POPs and Lagrangian-conic relaxation in the framework. We construct a hierarchy of primal-dual pairs of COPs (1.1) and (1.2) with $\mathbb{K} = \mathbb{K}_\omega \subset \mathbb{V}_\omega$, $\boldsymbol{Q}^0 = \boldsymbol{Q}^0_\omega \in \mathbb{V}_\omega$, $\boldsymbol{H}^0 = \boldsymbol{H}^0_\omega \in \mathbb{K}^*_\omega$ and $\boldsymbol{Q}^k = \boldsymbol{Q}^k_\omega \in \mathbb{K}^*_\omega$ $(k = 1, 2, \ldots, m)$. Here $\omega$ denotes a positive integer parameter describing the hierarchy, $\mathbb{V}_\omega$ the space of symmetric matrices with some dimension which monotonically diverges to $\infty$ as $\omega \to \infty$, and $\mathbb{K}_\omega$ the intersection of a positive semidefinite matrix cone in $\mathbb{V}_\omega$ and a linear subspace of $\mathbb{V}_\omega$. Under the so-called Archimedean condition, which requires that the feasible region of POP (1.3) is nonempty and bounded, the optimal value $\zeta^d_\omega(\mathbb{K}_\omega)$ of the dual COP in the hierarchy monotonically converges to the optimal value $\zeta^*$ of POP (1.3) from below as $\omega \to \infty$. (Since the optimal value $\zeta^p_\omega(\mathbb{K}_\omega)$ of the primal COP satisfies $\zeta^d_\omega(\mathbb{K}_\omega) \leq \zeta^p_\omega(\mathbb{K}_\omega) \leq \zeta^*$, it also converges to $\zeta^*$ of POP (1.3) from below as $\omega \to \infty$.)

A fundamental difference between Lasserre's hierarchy of SDP relaxations and ours lies on Condition (I), which is satisfied by ours, but not by Lasserre's. Thus, the entire theory of the unified framework can be applied to ours. In particular, we derive the following primal-dual pair of Lagrangian-SDP relaxation problems:

$$\begin{aligned} \zeta^p_\omega(\lambda, \mathbb{K}_\omega) &:= \inf \left\{ \langle \boldsymbol{Q}^0_\omega + \lambda \boldsymbol{H}^1_\omega, \boldsymbol{X} \rangle \;\middle|\; \boldsymbol{X} \in \mathbb{K}_\omega, \ \langle \boldsymbol{H}^0_\omega, \boldsymbol{X} \rangle = 1 \right\}, \\ \zeta^d_\omega(\lambda, \mathbb{K}_\omega) &:= \sup \left\{ y_0 \;\middle|\; y_0 \in \mathbb{R}, \ \boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \lambda \boldsymbol{H}^1_\omega \in \mathbb{K}^*_\omega \right\}, \end{aligned}$$

where $\boldsymbol{H}^1_\omega = \sum_{k=1}^m \boldsymbol{Q}^k_\omega$, and $\lambda \in \mathbb{R}$ denotes a Lagrangian multiplier prescribed for the problems. The simplicity of these primal-dual COPs enables us to design efficient first-order algorithms for solving the problems. In fact, the dual problem involves only one variable, which makes it possible to effectively utilize the bisection and the 1-dimensional Newton methods proposed in Part I [4]; see also [17]. Moreover, the common optimal value $\zeta^p_\omega(\lambda, \mathbb{K}_\omega) = \zeta^d_\omega(\lambda, \mathbb{K}_\omega)$ bounds $\zeta^d_\omega(\mathbb{K}_\omega)$ from below, and it monotonically converges to $\zeta^d_\omega(\mathbb{K}_\omega)$ as $\lambda \to \infty$. Therefore, under the Archimedean condition, the lower bound $\zeta^p_\omega(\lambda, \mathbb{K}_\omega) = \zeta^d_\omega(\lambda, \mathbb{K}_\omega)$ for the optimal value $\zeta^*$ of POP (1.3) satisfies $\zeta^* - \epsilon < \zeta^d_\omega(\lambda, \mathbb{K}_\omega) \le \zeta^*$ for any $\epsilon > 0$ if sufficiently large $\omega$ and $\lambda$ are taken.

## 1.2 Related results

In addition to the nonemptiness of the feasible region of the primal COP (1.1), Condition (I) requires the copositivity of $\boldsymbol{H}^0 \in \mathbb{V}$ and $\boldsymbol{Q}^k \in \mathbb{V}$ $(k = 1, 2, \ldots, m)$, *i.e.*, $\boldsymbol{H}^0 \in \mathbb{K}^*$ and $\boldsymbol{Q}^k \in \mathbb{K}^*$ $(k = 1, 2, \ldots, m)$. The copositivity condition is originated from a property shared by nonconvex QOPs, including QOPs over the unit simplex [8] and QOPs with linear, binary and complementarity constraints [1, 9], which can be converted to the equivalent CPP problems.

The MC relaxation (1.3) is related to a canonical convexification procedure for POPs by Peña, Vera and Zuluaga [22]. The essential difference lies on the way of convexification (see Section 3 and Theorem 3.2). Furthermore two conditions assumed for the equivalence of a given POP to its convexification are different. The first condition, a hierarchy of copositivity, is weaker than the simple copositivity condition, Condition (I). We need the stronger condition, Condition (I), for consistently deriving the Lagrangian-conic relaxation from (1.1). However, the second condition, zeros at infinity, is stronger than Condition (IV) assumed for the equivalence of POP (3) to its MC relaxation in this paper. When the optimal value $\zeta^*$ of POP (1.3) is finite and Condition (I) is satisfied, Condition (IV) provides a necessary and sufficient condition for the equivalence, while the zeros at infinity condition is merely a sufficient condition. See Section 6 of [3] and Section 3 of Part I [4] for more details.

The MC relaxation was first proposed in [3] by Arima, Kim and Kojima for a homogeneous POP of the form (1.3) where all $f_k(\boldsymbol{x})$ $(k = 0, 1, \ldots, m)$ are homogeneous polynomials with a common degree. A general inhomogeneous POP of the form (1.3) can always be transformed to a homogeneous POP by introducing an auxiliary variable $x_0$ fixed to 1, and their MC relaxation can be applied to the homogenized POP. The MC relaxation obtained this way is equivalent to our MC relaxation applied directly to the original POP using the unified framework. We note, however, that the unified framework not only simplifies the derivation of the MC relaxation of a general POP (1.3) but also relaxes the zeros at infinity condition, which was assumed in [3, 22] for the equivalence of a POP to its MC relaxation (or convexification), into Condition (IV).

Lasserre's hierarchy of SDP relaxation method [20] for a POP is very powerful in theory. The optimal value of each SDP in the hierarchy parameterized by a positive integer $\omega$ provides a lower bound for the optimal value of the POP, and the lower bound monotonically converges to the optimal value under the Archimedean condition. Software packages implementing Lasserre's method [15, 28] apply the primal-dual interior-point method [13, 24, 25] to the generated SDP. Numerical efficiency suffers in this implementation: the size of the SDP to be solved grows very rapidly, as the degree of polynomials involved in the POP becomes larger and/or a value for $\omega$ is increased to obtain a tighter lower bound for the optimal value of the POP.

As a result, exploiting sparsity is essential to solve large scale POPs. In fact, a sparse version of Lasserre's hierarchy of SDP relaxation method was proposed in [26], and it considerably improves the performance in terms of the speed and the size of POPs to be solved. In practical applications, however, POPs that Lasserre's method and its sparse version can solve are still limited to small-medium scale, except for very sparse POPs [26]. As mentioned in Section 1.1, the hierarchy of Lagrangian-SDP relaxations proposed for POPs in this paper inherits the nice theoretical properties from Lasserre's method. But it is designed to solve large scale POPs effectively and efficiently by using first order methods instead of the primal-dual interior-point method.

For simplicity of notation and enumerate, we will describe the hierarchy of Lagrangian-SDP relaxation method for POP (1.3) without taking account of any possible sparsity in polynomials involved there. Exploiting sparsity [14, 16, 19, 21, 26] in the method is necessary to tackle larger scale POPs. We refer to the original version [5] of this paper for a sparse version of the method.

### 1.3 Paper outline

In Section 2, we review the results shown in Part I [4] and describe the notation and symbols used in this paper. We also present how to represent polynomials with symmetric matrices of monomials, and introduce sum of squares (SOS) of polynomials. Section 3 includes the discussion on the MC relaxation of POP (1.3), and Section 4 presents the hierarchy of Lagrangian-SDP relaxations of POP (1.3) with $\mathbb{J} = \mathbb{R}^n$.

## 2 Preliminaries

### 2.1 Conic and Lagrangian-conic optimization problems

The results in Part I [4] are summarized in this subsection. We first list some notation used in Part I. Let $\mathbb{V}$ be a finite dimensional vector space endowed with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$, and $\mathbb{K}$ a nonempty (but not necessarily convex nor closed) cone in $\mathbb{V}$. We denote the dual of $\mathbb{K}$ by $\mathbb{K}^*$, *i.e.*, $\mathbb{K}^* = \{\boldsymbol{Y} \in \mathbb{V} : \langle \boldsymbol{X}, \boldsymbol{Y} \rangle \geq 0 \text{ for every } \boldsymbol{X} \in \mathbb{K}\}$, and the convex hull of $\mathbb{K}$ by co $\mathbb{K}$.

For $\boldsymbol{H}^0,\ \boldsymbol{Q}^k \in \mathbb{V}\ (k = 0, , 2, \ldots, m),\ \boldsymbol{H}^1 = \sum_{k=1}^m \boldsymbol{Q}^k$, let

$$\begin{aligned} F(\mathbb{K}) &= \left\{ \boldsymbol{X} \in \mathbb{V} \;\middle|\; \boldsymbol{X} \in \mathbb{K},\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1,\ \langle \boldsymbol{Q}^k, \boldsymbol{X} \rangle = 0\ (k = 1, 2, \ldots, m) \right\} \\ &\quad \text{(the feasible region of (1.1))}, \\ F_0(\mathbb{K}) &= \left\{ \boldsymbol{X} \in \mathbb{V} \;\middle|\; \boldsymbol{X} \in \mathbb{K},\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 0,\ \langle \boldsymbol{Q}^k, \boldsymbol{X} \rangle = 0\ (k = 1, 2, \ldots, m) \right\}. \end{aligned}$$

In Part I, we introduced the following conditions:

**Condition (I)** $F(\mathbb{K}) \neq \emptyset$, $\boldsymbol{O} \neq \boldsymbol{H}^0 \in \mathbb{K}^*$ and $\boldsymbol{Q}^k \in \mathbb{K}^*$ $(k = 1, 2, \ldots, m)$.

**Condition (II)** $\mathbb{K}$ is closed and convex.

**Condition (III)** $\left\{\boldsymbol{X} \in \boldsymbol{F}(\mathbb{K}) : \langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle \leq \tilde{\zeta}\right\}$ is nonempty and bounded for some $\tilde{\zeta} \in \mathbb{R}$.

**Condition (IV)** $\langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle \geq 0$ for every $\boldsymbol{X} \in F_0(\mathbb{K})$.

For the primal-dual pair (1.1) - (1.2) and the following pair

$$\eta^p(\lambda, \mathbb{K}) \ := \ \inf \left\{ \langle (\boldsymbol{Q}^0 + \lambda \boldsymbol{H}^1), \boldsymbol{X} \rangle \ \middle| \ \boldsymbol{X} \in \mathbb{K}, \ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1 \ \right\}, \tag{2.1}$$

$$\eta^d(\lambda, \mathbb{K}) \ := \ \sup \left\{ y_0 \ \middle| \ \boldsymbol{Q}^0 + \lambda \boldsymbol{H}^1 - \boldsymbol{H}^0 y_0 \in \mathbb{K}^* \ \right\}, \tag{2.2}$$

the following results are shown.

**Theorem 2.1.**

(i) *$\eta^d(\lambda, \mathbb{K})\uparrow\lambda \ = \zeta^d(\mathbb{K}) \leq \zeta^p(\mathbb{K})$ and $\left(\eta^d(\lambda, \mathbb{K}) \leq \eta^p(\lambda, \mathbb{K})\right)\uparrow\lambda \ \leq \zeta^p(\mathbb{K})$ under Condition (I). Here $(\ )\uparrow\lambda$ means a monotonic increase as $\lambda \to \infty$, satisfying the equality or inequality inside the parenthesis if it exists.* (Lemmas 2.1 and 2.2 of [4]).

(ii) *Suppose that Conditions (I) and (II) hold. Then $\left(\eta^d(\lambda, \mathbb{K}) = \eta^p(\lambda, \mathbb{K})\right)\uparrow\lambda \ = \zeta^d(\mathbb{K}) \leq \zeta^p(\mathbb{K})$. If in addition $\eta^p(\lambda, \mathbb{K})$ is finite,* (2.2) *has an optimal solution with the objective value $\eta^d(\lambda, \mathbb{K}) = \eta^p(\lambda, \mathbb{K})$.* (Lemma 2.3 of [4]).

(iii) *$\left(\eta^d(\lambda, \mathbb{K}) = \eta^p(\lambda, \mathbb{K})\right)\uparrow\lambda \ = \zeta^d(\mathbb{K}) = \zeta^p(\mathbb{K})$ under Conditions (I), (II) and (III).* (Lemma 2.5 of [4]).

1. *(iv) Assume that Condition (I) holds. Then*

$$\zeta^p(\mathbb{K}) \ = \ \begin{cases} \zeta^p(co\ \mathbb{K}) & \text{if Condition (IV) holds,} \\ -\infty & \text{otherwise.} \end{cases}$$

(Lemma 3.1 of [4]).

## 2.2 Notation and symbols

For the application of the unified framework to POPs, we use the following notation: Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{R}_+$ the set of nonnegative real numbers, and $\mathbb{Z}_+$ the set of nonnegative integers. Let $|\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i$ for each $\boldsymbol{\alpha} \in \mathbb{Z}^n_+$, where $\alpha_i$ denotes the $i$-th element of $\boldsymbol{\alpha} \in \mathbb{Z}^n_+$. $\mathbb{R}[\boldsymbol{x}]$ is the set of real-valued multivariate polynomials in $x_i \in \mathbb{R}$ $(i = 1, \ldots, n)$, where $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Each polynomial $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ is represented as $f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{F}} f_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$, where $\mathcal{F} \subset \mathbb{Z}^n_+$ is a nonempty finite set, $f_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in \mathcal{F})$ real coefficients, $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n_+$. We assume that $x_i^0 = 1$ even if $x_i = 0$, in particular, $\boldsymbol{x}^{\boldsymbol{0}} = 1$ for any $\boldsymbol{x} \in \mathbb{R}^n$. The support of $f(\boldsymbol{x})$ is defined by $\text{supp}(f(\boldsymbol{x})) = \{\boldsymbol{\alpha} \in \mathcal{F} : f_{\boldsymbol{\alpha}} \neq 0\} \subset \mathbb{Z}^n_+$, and the degree of $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ is defined by $\deg(f(\boldsymbol{x})) = \max\{|\boldsymbol{\alpha}| : \boldsymbol{\alpha} \in \text{supp}(f(\boldsymbol{x}))\}$. For each nonempty subsets $\mathcal{F}$ and $\mathcal{G}$ of $\mathbb{Z}^n_+$, let $\mathcal{F} + \mathcal{G}$ denote their Minkowski sum $\{\boldsymbol{\alpha} + \boldsymbol{\beta} : \boldsymbol{\alpha} \in \mathcal{F},\ \boldsymbol{\beta} \in \mathcal{G}\}$, and let $\mathbb{R}[\boldsymbol{x}, \mathcal{F}] = \{f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}] : \text{supp}(f(\boldsymbol{x})) \subset \mathcal{F}\}$.

Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$. $|\mathcal{F}|$ stands for the number of elements of $\mathcal{F}$. Let $\mathbb{R}^{\mathcal{F}}$ denote the $|\mathcal{F}|$-dimensional Euclidean space whose coordinate are indexed by $\boldsymbol{\alpha} \in \mathcal{F}$. Each vector of $\mathbb{R}^{\mathcal{F}}$ with elements $w_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in \mathcal{F})$ is denoted by $(w_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{F})$ or simply $(w_{\boldsymbol{\alpha}} : \mathcal{F})$. We assume that $(w_{\boldsymbol{\alpha}} : \mathcal{F})$ is a column vector when it is multiplied by a matrix. For every $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}^{\mathcal{F}} = (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})$ denotes the $|\mathcal{F}|$-dimensional (column) vector of monomials $\boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\boldsymbol{x}]$ ($\boldsymbol{\alpha} \in \mathcal{F}$). Hence each polynomial $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F}]$ is represented as $f(\boldsymbol{x}) = \langle (f_{\boldsymbol{\alpha}} : \mathcal{F}), \boldsymbol{x}^{\mathcal{F}} \rangle$. $\mathbb{S}^{\mathcal{F}}$ denotes the linear space of $|\mathcal{F}| \times |\mathcal{F}|$ symmetric matrices with elements $\xi_{\boldsymbol{\alpha\beta}}$ ($\boldsymbol{\alpha} \in \mathcal{F}$, $\boldsymbol{\beta} \in \mathcal{F}$). We use the notation $\square\mathcal{F}$ for the set $\mathcal{F} \times \mathcal{F} = \{(\boldsymbol{\alpha}, \boldsymbol{\beta}) : \boldsymbol{\alpha},\ \boldsymbol{\beta} \in \mathcal{F}\}$.

Each matrix of $\mathbb{S}^{\mathcal{F}}$ is written as $(\xi_{\boldsymbol{\alpha\beta}} : (\alpha, \beta) \in \square\mathcal{F})$ or simply $(\xi_{\boldsymbol{\alpha\beta}} : \square\mathcal{F})$. For $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}^{\mathcal{F}}(\boldsymbol{x}^{\mathcal{F}})^T = (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T = (\boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} : (\alpha, \beta) \in \square\mathcal{F})$ is a rank-1 symmetric matrix of monomials $\boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} \in \mathbb{R}[\boldsymbol{x}]$ ($(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \square\mathcal{F}$), which is denoted by $\boldsymbol{x}^{\square\mathcal{F}}$. $(\boldsymbol{x}^{\mathcal{F}})^T$ means the row vector obtained by taking the transpose of the column vector $\boldsymbol{x}^{\mathcal{F}}$.

For every pair of $\boldsymbol{Q} = (Q_{\boldsymbol{\alpha\beta}} : \square\mathcal{F})$ and $\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}}$, $\langle \boldsymbol{Q}, \boldsymbol{X} \rangle$ denotes the matrix inner product, *i.e.*, $\langle \boldsymbol{Q}, \boldsymbol{X} \rangle = \text{trace}(\boldsymbol{Q}^T \boldsymbol{X}) = \sum_{(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \square\mathcal{F}} Q_{\boldsymbol{\alpha\beta}} \xi_{\boldsymbol{\alpha\beta}}$. With this notation, we often write the quadratic form $(\boldsymbol{x}^{\mathcal{F}})^T \boldsymbol{Q} \boldsymbol{x}^{\mathcal{F}}$ as $\langle \boldsymbol{Q}, \boldsymbol{x}^{\square\mathcal{F}} \rangle$ to indicate that $\boldsymbol{x}^{\square\mathcal{F}} = \boldsymbol{x}^{\mathcal{F}}(\boldsymbol{x}^{\mathcal{F}})^T$ will be replaced by $\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}}$.

Let

$$
\begin{aligned}
\mathbb{S}^{\mathcal{F}}_+ &= \text{the cone of positive semidefinite matrices in } \mathbb{S}^{\mathcal{F}} \\
&= \left\{ (\xi_{\boldsymbol{\alpha\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}} : \begin{array}{l} (w_{\boldsymbol{\alpha}} : \mathcal{F})^T (\xi_{\boldsymbol{\alpha\beta}} : \square\mathcal{F})(w_{\boldsymbol{\alpha}} : \mathcal{F}) \geq 0 \\ \text{for every } (w_{\boldsymbol{\alpha}} : \mathcal{F}) \in \mathbb{R}^{\mathcal{F}} \end{array} \right\}, \\
\mathbb{L}^{\mathcal{F}} &= \left\{ (\xi_{\boldsymbol{\alpha\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}} : \xi_{\boldsymbol{\alpha\beta}} = \xi_{\boldsymbol{\gamma\delta}} \text{ if } \boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\gamma} + \boldsymbol{\delta} \right\}.
\end{aligned}
$$

Then $\mathbb{S}^{\mathcal{F}}_+$ forms a closed convex cone in $\mathbb{S}^{\mathcal{F}}$, and $\mathbb{L}^{\mathcal{F}}$ a linear subspace of $\mathbb{S}^{\mathcal{F}}$. We also see that $\boldsymbol{x}^{\square\mathcal{F}} \in \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{L}^{\mathcal{F}}$ for every $\boldsymbol{x} \in \mathbb{R}^n$. This relation is used repeatedly in the subsequent discussions.

### 2.3 Representing polynomials with symmetric matrices of monomials and sums of squares of polynomials

For a nonempty finite subset $\mathcal{G}$ of $\mathbb{Z}^n_+$, a given polynomial $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{G}]$ is usually represented as the inner product of its coefficient vector $(f_{\boldsymbol{\alpha}} : \mathcal{G})$ and the vector $\boldsymbol{x}^{\mathcal{G}} = (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{G})$ of monomials in the polynomial, *i.e.*, $g(\boldsymbol{x}) = \langle (f_{\boldsymbol{\alpha}} : \mathcal{G}), \boldsymbol{x}^{\mathcal{G}} \rangle$. However, representing a polynomial with a symmetric matrix of monomials is more convenient in the subsequent discussions, in particular, when discussing sum of squares (SOS) of polynomials and conic and Lagrangian-conic relaxations. For the representation of a polynomial $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{G}]$ using a symmetric matrix of monomials, we need to choose a finite subset $\mathcal{F}$ of $\mathbb{Z}^n_+$ satisfying the property $\mathcal{G} \subset \mathcal{F} + \mathcal{F}$. In fact, a smaller-sized $\mathcal{F}$ satisfying this property is preferable for numerical efficiency. See [19] for details of choosing such an $\mathcal{F}$. See also [3].

Let $\mathcal{F}$ be a nonempty subset of $\mathbb{Z}^n_+$ and $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F} + \mathcal{F}]$. Then we can represent the polynomial $f(\boldsymbol{x})$ using the rank-1 symmetric matrix $\boldsymbol{x}^{\square\mathcal{F}} = \boldsymbol{x}^{\mathcal{F}}(\boldsymbol{x}^{\mathcal{F}})^T$ of monomials $\boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}$ ($(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \square\mathcal{F}$) and some $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}$ such that $f(\boldsymbol{x}) = \langle \boldsymbol{Q}, \boldsymbol{x}^{\square\mathcal{F}} \rangle$. Note that $\boldsymbol{x}^{\square\mathcal{F}}$ contains all monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}$ ($\boldsymbol{\alpha} \in \mathcal{F} + \mathcal{F}$), and that the choice of such a $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}$ is not unique as shown in the following example.

**Example 2.2.** Consider the polynomial $f^1(\boldsymbol{x}) = f^1(x_1, x_2)$ in two real variables such that

$$
f^1(\boldsymbol{x}) = 1 - 2x_1 - 2x_2 + x_1^2 + x_2^2 + 2x_1^2 x_2 + 2x_1 x_2^2 + x_1^2 x_2^2.
$$

Let

$$
\mathcal{G} = \{(0,0), (1,0), (0,1), (2,0), (0,2), (2,1), (1,2), (2,2)\},
$$

$$\begin{aligned}(f^1_{\boldsymbol{\alpha}} : \mathcal{G}) &= (1,-2,-2,1,1,2,2,1),\\ \boldsymbol{x}^{\mathcal{G}} &= (x^{\boldsymbol{\alpha}} : \mathcal{G}) = (1, x_1, x_2, x_1^2, x_2^2, x_1^2x_2, x_2^2.x_1^2x_2^2).\end{aligned}$$

Then $\mathbb{R}[\boldsymbol{x},\mathcal{G}] \ni f^1(\boldsymbol{x}) = \langle (f^1_{\boldsymbol{\alpha}} : \mathcal{G}),\, \boldsymbol{x}^{\mathcal{G}}\rangle$. To represent $f^1(\boldsymbol{x})$ using a symmetric matrix of monomials, we can take $\mathcal{F} = \{(0,0),(1,0),(0,1),(1,1)\}$ so that

$$\mathcal{G} \subset \mathcal{F}+\mathcal{F} = \{(0,0),(1,0),(0,1),(2,0),(0,2),(1,1),(2,1),(1,2),(2,2)\}.$$

Let

$$\begin{aligned}\boldsymbol{Q} &= \begin{pmatrix} 1 & -1 & -1 & 0\\ -1 & 1 & 0 & 1\\ -1 & 0 & 1 & 1\\ 0 & 1 & 1 & 1\end{pmatrix} \in \mathbb{S}^{\mathcal{F}},\ \boldsymbol{P} = \begin{pmatrix} 0 & 0 & 0 & -1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ -1 & 0 & 0 & 0\end{pmatrix} \in \mathbb{S}^{\mathcal{F}},\\ \boldsymbol{x}^{\square\mathcal{F}} &= \begin{pmatrix} 1 & x_1 & x_2 & x_1x_2\\ x_1 & x_1^2 & x_1x_2 & x_1^2x_2\\ x_2 & x_1x_2 & x_2^2 & x_1x_2^2\\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^2x_2^2\end{pmatrix}.\end{aligned}$$

Then, $\mathbb{R}[\boldsymbol{x},\mathcal{F}+\mathcal{F}] \ni f^1(\boldsymbol{x}) = \langle \boldsymbol{Q}+\mu\boldsymbol{P},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle$ for every $\mu \in \mathbb{R}$.

**Lemma 2.3.** *Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$, $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}$ and $\boldsymbol{P} \in \mathbb{S}^{\mathcal{F}}$. Then,*

$$\mathbb{R}[\boldsymbol{x},\mathcal{F}+\mathcal{F}] \ni \langle \boldsymbol{Q},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle = \langle \boldsymbol{Q}+\boldsymbol{P},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle \textit{ if and only if } \boldsymbol{P} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}. \tag{2.3}$$

*Proof.* We first show that the linear subspace of $\mathbb{S}^{\mathcal{F}}$ generated by $\{\boldsymbol{x}^{\square\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^n\}$, *i.e.*, $\overline{\mathbb{L}} = \{\lambda\boldsymbol{x}^{\square\mathcal{F}} + \mu\boldsymbol{y}^{\square\mathcal{F}} : \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^n,\ \lambda,\ \mu \in \mathbb{R}\}$, coincides with $\mathbb{L}^{\mathcal{F}}$. By the definition of $\mathbb{L}^{\mathcal{F}}$, we know that $\{\boldsymbol{x}^{\square\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^n\} \subset \mathbb{L}^{\mathcal{F}}$, hence $\overline{\mathbb{L}} \subset \mathbb{L}^{\mathcal{F}}$. It suffices to show that $\dim(\overline{\mathbb{L}}) = \dim(\mathbb{L}^{\mathcal{F}})$. Let $\ell = \dim(\mathbb{L}^{\mathcal{F}})$, which is equivalent to the number of distinct elements in $\mathcal{F}+\mathcal{F}$. Let $\mathbb{M}$ be the linear subspace of $\mathbb{R}^{\ell}$ generated by the set $\{\boldsymbol{x}^{\mathcal{F}+\mathcal{F}} = (\boldsymbol{x}^{\boldsymbol{\gamma}} : \boldsymbol{\gamma} \in \mathcal{F}+\mathcal{F})\} \subset \mathbb{R}^{\ell}$. Then we can identify the linear space $\overline{\mathbb{L}}$ as the linear space $\mathbb{M}$ since each matrix $\boldsymbol{X} = \lambda\boldsymbol{x}^{\square\mathcal{F}} + \mu\boldsymbol{y}^{\square\mathcal{F}} \in \overline{\mathbb{L}}$ corresponds to a vector $\lambda\boldsymbol{x}^{\mathcal{F}+\mathcal{F}} + \mu\boldsymbol{y}^{\mathcal{F}+\mathcal{F}} \in \mathbb{M}$ and vice versa. As a result, $\dim(\overline{\mathbb{L}}) = \dim(\mathbb{M})$. On the other hand, we see that $\dim(\mathbb{M}) = \ell$ since there is no nonzero $(g_{\boldsymbol{\alpha}} : \mathcal{F}+\mathcal{F})$ such that $\langle (g_{\boldsymbol{\alpha}} : \mathcal{F}+\mathcal{F}),\, \boldsymbol{x}^{\mathcal{F}+\mathcal{F}}\rangle = 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Therefore, we obtain that $\dim(\overline{\mathbb{L}}) = \dim(\mathbb{L}^{\mathcal{F}}) = \ell$. Thus we have shown that $\mathbb{L}^{\mathcal{F}} = \overline{\mathbb{L}}$. Now assume that $\mathbb{R}[\boldsymbol{x},\mathcal{F}+\mathcal{F}] \ni \langle \boldsymbol{Q},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle = \langle \boldsymbol{Q}+\boldsymbol{P},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle$. Then $\langle \boldsymbol{P},\, \boldsymbol{X}\rangle = 0$ for all $\boldsymbol{X} \in \overline{\mathbb{L}} = \mathbb{L}^{\mathcal{F}}$, which implies that $\boldsymbol{P} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$. Conversely, if $\boldsymbol{P} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$, then $\mathbb{R}[\boldsymbol{x},\mathcal{F}+\mathcal{F}] \ni \langle \boldsymbol{Q},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle = \langle \boldsymbol{Q}+\boldsymbol{P},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle$ holds from $\boldsymbol{x}^{\square\mathcal{F}} \in \mathbb{L}^{\mathcal{F}}$ for every $\boldsymbol{x} \in \mathbb{R}^n$. □

Lemma 2.3 implies that, for each $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}$, $\{\boldsymbol{Q}+\boldsymbol{P} : \boldsymbol{P} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}\}$ forms an equivalent class in $\mathbb{S}^{\mathcal{F}}$ represented by the common polynomial $f(\boldsymbol{x}) = \langle \boldsymbol{Q},\, \boldsymbol{x}^{\square\mathcal{F}}\rangle$.

We introduce some additional notation for SOS of polynomials. Let

$$\begin{aligned}\mathbb{SOS}[\boldsymbol{x},\mathcal{F}] &= \left\{\sum_{i=1}^{r}(\varphi^i(\boldsymbol{x}))^2 : \varphi^i(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x},\mathcal{F}]\ (i=1,\ldots,r)\ \exists r \in \mathbb{Z}_+\right\}\\ &\quad \text{for every } \mathcal{F} \subset \mathbb{Z}^n_+,\\ \mathbb{SOS}[\boldsymbol{x}] &= \mathbb{SOS}[\boldsymbol{x},\mathbb{Z}^n_+].\end{aligned}$$

We call $\mathbb{SOS}[\boldsymbol{x}]$ *the cone of SOS of polynomials*, and each $f(\boldsymbol{x}) \in \mathbb{SOS}[\boldsymbol{x}]$ *an SOS polynomial.* The following lemma provides a characterization of SOS polynomials.

**Lemma 2.4** ([11]). *Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$. Then*

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{F}] = \left\{ \langle \boldsymbol{Q}, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle : \boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}_+ \right\}.$$

In Example 2.2, the matrix $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}$ itself is not positive semidefinite. But if we choose $\mu = 1$, then the matrix $\boldsymbol{Q}^1 = \boldsymbol{Q} + \mu \boldsymbol{P} \in \mathbb{S}^{\mathcal{F}}$ is positive semidefinite. Hence $f^1(\boldsymbol{x}) = \langle \boldsymbol{Q}^1, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{F}]$ by the lemma above. In fact, we see that

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{F}] \ni f^1(\boldsymbol{x}) = \langle \boldsymbol{Q}^1, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle = (x_1 + x_2 + x_1 x_2 - 1)^2, \tag{2.4}$$

where

$$\mathcal{F} \;\; = \;\; \{(0,0), (1,0), (0,1), (1,1)\}, \; \boldsymbol{Q}^1 = \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{S}^{\mathcal{F}}_+. \tag{2.5}$$

The following lemmas will be used in Sections 3 and 4.

**Lemma 2.5.** *Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$ and $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}$. Then $\langle \boldsymbol{Q}, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{F}]$ if $\boldsymbol{Q} \in \mathbb{S}^{\mathcal{F}}_+ + \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$.*

*Proof.* By Lemma 2.4, $\langle \boldsymbol{Q}, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{F}]$ if and only if the identity $\langle \boldsymbol{Q}, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle = \langle \boldsymbol{V}, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle$ holds for some $\boldsymbol{V} \in \mathbb{S}^{\mathcal{F}}_+$. Hence, by Lemma 2.3, $\langle \boldsymbol{Q}, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{F}]$ if and only if $\boldsymbol{Q} = \boldsymbol{V} + \boldsymbol{P}$ for some $\boldsymbol{V} \in \mathbb{S}^{\mathcal{F}}_+$ and $\boldsymbol{P} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$. Therefore, the desired result follows. □

**Lemma 2.6.** *Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$ and $\mathbb{M}$ a linear subspace of $\mathbb{S}^{\mathcal{F}}$.*

(i) $\mathbb{S}^{\mathcal{F}}_+ \cap \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} = \{\boldsymbol{O}\}$.

(ii) *The set $\{\boldsymbol{X} \in \mathbb{S}^{\mathcal{F}} : \boldsymbol{X} + \boldsymbol{Y} \in B, \; \boldsymbol{X} \in \mathbb{S}^{\mathcal{F}}_+, \; \boldsymbol{Y} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}\}$ is bounded for every bounded subset $B$ of $\mathbb{S}^{\mathcal{F}}$.*

(iii) $(\mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{M}) + (\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M})$ *is closed.*

*Proof.* (i) Let $q = |\mathcal{F}|$. Since the set of monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}$ ($\boldsymbol{\alpha} \in \mathcal{F}$) is independent, *i.e.*, there is no nonzero $(g_{\boldsymbol{\alpha}} : \mathcal{F})$ such that $(g_{\boldsymbol{\alpha}} : \mathcal{F})^T (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})$ is identically zero for all $\boldsymbol{x} \in \mathbb{R}^n$, there exist $\boldsymbol{x}_j \in \mathbb{R}^n$ $(j = 1, \ldots, q)$ such that the $q \times q$ matrix $\boldsymbol{A} = \left((\boldsymbol{x}_1^{\boldsymbol{\alpha}} : \mathcal{F}), (\boldsymbol{x}_2^{\boldsymbol{\alpha}} : \mathcal{F}), \ldots, (\boldsymbol{x}_q^{\boldsymbol{\alpha}} : \mathcal{F})\right)$ is nonsingular. Let $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_+ \cap \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$. Then $\langle \boldsymbol{A}\boldsymbol{A}^T, \, \overline{\boldsymbol{X}} \rangle = \langle \sum_{j=1}^q (\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})^T, \, \overline{\boldsymbol{X}} \rangle = \sum_{j=1}^q \langle (\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})^T, \, \overline{\boldsymbol{X}} \rangle = 0$. Since $\boldsymbol{A}\boldsymbol{A}^T$ is a $q \times q$ positive definite matrix and $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_+$, the identity $\langle \boldsymbol{A}\boldsymbol{A}^T, \, \overline{\boldsymbol{X}} \rangle = 0$ implies that $\overline{\boldsymbol{X}} = \boldsymbol{O}$.

(ii) For some bounded subset $B$ of $\mathbb{S}^{\mathcal{F}}$, assume on the contrary that there exists a sequence $\{\boldsymbol{Z}^p = \boldsymbol{X}^p + \boldsymbol{Y}^p : p = 1, 2, \ldots\}$ satisfying $\boldsymbol{Z}^p = \boldsymbol{X}^p + \boldsymbol{Y}^p \in B, \; \boldsymbol{X}^p \in \mathbb{S}^{\mathcal{F}}_+, \; \boldsymbol{Y}^p \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$ $(p = 1, 2, \ldots)$ and $\|\boldsymbol{X}^p\| \to \infty$ as $p \to \infty$. We may assume without loss of generality that $\mathbb{S}^{\mathcal{F}}_+ \ni \boldsymbol{X}^p / \|\boldsymbol{X}^p\|$ converges to some nonzero $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_+$ as $p \to \infty$. Hence $\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \ni$

$\boldsymbol{Y}^p/\|\boldsymbol{X}^p\| = \boldsymbol{Z}^p/\|\boldsymbol{X}^p\| - \boldsymbol{X}^p/\|\boldsymbol{X}^p\|$ converges to $-\overline{\boldsymbol{X}}$ as $p \to \infty$. Since $\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$ is a linear subspace of $\mathbb{S}^{\mathcal{F}}$, we obtain that $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_+ \cap \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$. By (i), we know that $\overline{\boldsymbol{X}} = \boldsymbol{O}$. This is a contradiction.

(iii) Suppose that $\boldsymbol{Z}^p = \boldsymbol{X}^p + \boldsymbol{Y}^p$, $\boldsymbol{X}^p \in \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{M}$, $\boldsymbol{Y}^p \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M}$ $(p = 1, 2, \dots)$ and $\boldsymbol{Z}^p \to \overline{\boldsymbol{Z}}$ for some $\overline{\boldsymbol{Z}} \in \mathbb{S}^{\mathcal{F}}$ as $p \to \infty$. Since the sequence $\{\boldsymbol{Z}^p = \boldsymbol{X}^p + \boldsymbol{Y}^p : p = 1, 2, \dots, \}$ is bounded and $\boldsymbol{X}^p \in \mathbb{S}^{\mathcal{F}}_+$, $\boldsymbol{Y}^p \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$ $(p = 1, 2, \dots)$, the sequence $\{\boldsymbol{X}^p \in \mathbb{S}^{\mathcal{F}}_+ : p = 1, 2, \dots\}$ is bounded. Hence we may assume that it converges to some $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_+$. It follows that $\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M} \ni \boldsymbol{Y}^p = \boldsymbol{Z}^p - \boldsymbol{X}^p \to \overline{\boldsymbol{Z}} - \overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}$ as $p \to \infty$. Since both $\mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{M}$ and $\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M}$ are closed subsets of $\mathbb{S}^{\mathcal{F}}$, we know that $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{M}$ and $\overline{\boldsymbol{Z}} - \overline{\boldsymbol{X}} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M}$. Therefore, we have shown that $\overline{\boldsymbol{Z}} = \overline{\boldsymbol{X}} + (\overline{\boldsymbol{Z}} - \overline{\boldsymbol{X}}) \in (\mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{M}) + (\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M})$. □

## 3 A Class of POPs and their Covexification

Consider a class of POPs of the form (1.3). Assume that $\mathbb{J}$ is a nonempty closed (but not necessarily convex) cone in $\mathbb{R}^n$, and $f^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F} + \mathcal{F}]$ $(k = 0, 1, \dots, m)$ for a nonempty finite subset $\mathcal{F}$ of $\mathbb{Z}^n_+$ including $\boldsymbol{0} \in \mathbb{Z}^n_+$. For practical applications, we focus on $\mathbb{R}^n$, $\mathbb{R}^n_+$ and $\mathbb{R}^{\ell} \times \mathbb{R}^{n-\ell}_+$ with $1 \leq \ell \leq n-1$ for the cone $\mathbb{J}$, but the theoretical results in this section are valid for any closed cone in $\mathbb{R}^n$. We also assume throughout this section that the feasible region of (1.3) is nonempty.

We transform POP (1.3) into the COP of the form (1.1) to present the moment cone (MC) relaxation of POPs. Let us take $\mathbb{S}^{\mathcal{F}}$ for the underlying linear space $\mathbb{V}$, and represent each polynomial $f^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F} + \mathcal{F}]$ as $f^k(\boldsymbol{x}) = \langle \boldsymbol{Q}^k, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle$ $(k = 0, 1, \dots, m)$, where $\boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{F}}$. Let $\boldsymbol{\Delta}^{\mathcal{F}}_1 = \left\{\boldsymbol{x}^{\square \mathcal{F}} \in \mathbb{S}^{\mathcal{F}} \ : \boldsymbol{x} \in \mathbb{J}\right\}$. Then, POP (1.3) can be rewritten as

$$\zeta^* \quad := \quad \inf \left\{ \langle \boldsymbol{Q}^0, \, \boldsymbol{X} \rangle \;\middle|\; \boldsymbol{X} \in \boldsymbol{\Delta}^{\mathcal{F}}_1, \; \langle \boldsymbol{Q}^k, \, \boldsymbol{X} \rangle = 0 \; (k = 1, 2, \dots, m) \right\}. \tag{3.1}$$

We consider the following illustrative example:

**Example 3.1.** As in Example 2.2, we take $n = 2$ and $\mathcal{F} = \{(0,0), (1,0), (0,1), (1,1)\}$. Let $m = 1$, $\mathbb{R}[\boldsymbol{x}, \mathcal{F}+\mathcal{F}] \ni f^0(\boldsymbol{x}) = \langle \boldsymbol{Q}^0, \, \boldsymbol{x}^{\square \mathcal{F}} \rangle$ for some $\boldsymbol{Q}^0 \in \mathbb{S}^{\mathcal{F}}$, and let $f^1(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F}+\mathcal{F}]$ be a sum of squares of polynomial given in (2.4). Let $\mathbb{J} = \mathbb{R}^2_+$. In this case, it is obvious that the feasible region of POP (1.3) is bounded and contains two points $\boldsymbol{x} = (1, 0)$ and $\boldsymbol{x} = (0, 1)$, thus, its finite minimum value $\zeta^*$ is attained at some feasible solution. By definition, we see that $\boldsymbol{\Delta}^{\mathcal{F}}_1 = \{\boldsymbol{x}^{\square \mathcal{F}} \in \mathbb{S}^{\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^2_+\}$.

The problem (3.1) is in a form similar to COP (1.1), but we still need to embed $\boldsymbol{\Delta}^{\mathcal{F}}_1$ in a cone $\mathbb{K} \subset \mathbb{S}^{\mathcal{F}}$ and introduce an inhomogeneous equality constraint $\langle \boldsymbol{H}^0, \, \boldsymbol{X} \rangle = 1$ such that $\boldsymbol{\Delta}^{\mathcal{F}}_1 = \{\boldsymbol{X} \in \mathbb{K} : \langle \boldsymbol{H}^0, \, \boldsymbol{X} \rangle = 1\}$. This can be achieved by two methods. The first method is to take the conic hull of $\boldsymbol{\Delta}^{\mathcal{F}}_1$ such that

$$\boldsymbol{\Delta}^{\mathcal{F}} \quad = \quad \left\{\mu \boldsymbol{X} \in \mathbb{S}^{\mathcal{F}} : \mu \geq 0, \; \boldsymbol{X} \in \boldsymbol{\Delta}^{\mathcal{F}}_1\right\} = \left\{\mu \boldsymbol{x}^{\square \mathcal{F}} \in \mathbb{S}^{\mathcal{F}} : \mu \geq 0, \; \boldsymbol{x} \in \mathbb{J}\right\}.$$

The second method is to homogenize $\boldsymbol{\Delta}^{\mathcal{F}}_1$ such that

$$\boldsymbol{\Gamma}^{\mathcal{F}} \quad = \quad \left\{(x_0^{\tau - |\boldsymbol{\alpha}|} \boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau - |\boldsymbol{\alpha}|} \boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T \in \mathbb{S}^{\mathcal{F}} : (x_0, \boldsymbol{x}) \in \mathbb{R}_+ \times \mathbb{J}\right\},$$

where $\tau = \max\{|\boldsymbol{\alpha}| : \boldsymbol{\alpha} \in \mathcal{F}\}$. We note that, for $x_0 = 0$ and $\boldsymbol{x} \in \mathbb{R}^n$,

$$x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} = \begin{cases} 0 & \text{if } \tau > |\boldsymbol{\alpha}| \\ \boldsymbol{x}^{\boldsymbol{\alpha}} & \text{otherwise, } i.e., \text{ if } \tau = |\boldsymbol{\alpha}|. \end{cases} \tag{3.2}$$

Both $\boldsymbol{\Delta}^{\mathcal{F}}$ and $\boldsymbol{\Gamma}^{\mathcal{F}}$ are cones in $\mathbb{S}^{\mathcal{F}}$. The first construction of the cone $\boldsymbol{\Delta}^{\mathcal{F}}$ was (implicitly) employed in [22], while the construction of the second cone $\boldsymbol{\Gamma}^{\mathcal{F}}$ is an extension of the one discussed in Section 5 of Part I [4] for a class of linearly constrained QOPs with complementarity constraints. We are mainly interested in the second one.

Let $\boldsymbol{H}^0$ be a matrix in $\mathbb{S}^{\mathcal{F}}$ with 1 in the $(\boldsymbol{0},\boldsymbol{0})$-th element and 0 elsewhere. Then, we see that $\langle \boldsymbol{H}^0,\, \boldsymbol{x}^{\square\mathcal{F}}\rangle = \langle \boldsymbol{H}^0,\, (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T\rangle = \boldsymbol{x}^{\boldsymbol{0}}\boldsymbol{x}^{\boldsymbol{0}} = 1$ for every $\boldsymbol{x} \in \mathbb{R}^n$, and that $\langle \boldsymbol{H}^0,\, \boldsymbol{X}\rangle = X_{\boldsymbol{00}} = x_0^{2\tau}$ for every $\boldsymbol{X} = (x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T \in \boldsymbol{\Gamma}^{\mathcal{F}}$. It follows that

$$\begin{aligned}
&\left\{\boldsymbol{X} \in \boldsymbol{\Delta}^{\mathcal{F}} : \langle \boldsymbol{H}^0,\, \boldsymbol{X}\rangle = 1\right\} \\
&= \left\{\mu\boldsymbol{x}^{\square\mathcal{F}} \in \mathbb{S}^{\mathcal{F}} : \boldsymbol{x} \in \mathbb{J},\ \mu \geq 0,\ \langle \boldsymbol{H}^0,\, \mu\boldsymbol{x}^{\square\mathcal{F}}\rangle = 1\right\} \;=\; \boldsymbol{\Delta}_1^{\mathcal{F}}, \\
&\left\{\boldsymbol{X} \in \boldsymbol{\Gamma}^{\mathcal{F}} : \langle \boldsymbol{H}^0,\, \boldsymbol{X}\rangle = 1\right\} \\
&= \left\{(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T : (x_0, \boldsymbol{x}) \in \mathbb{R}_+ \times \mathbb{J},\ x_0 = 1\right\} \;=\; \boldsymbol{\Delta}_1^{\mathcal{F}}.
\end{aligned} \tag{3.3}$$

Therefore, both COP (1.1) with $\mathbb{K} = \boldsymbol{\Delta}^{\mathcal{F}}$ and COP (1.1) with $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$ are equivalent to POP (3.1), and $\zeta^p(\boldsymbol{\Delta}^{\mathcal{F}}) = \zeta^p(\boldsymbol{\Gamma}^{\mathcal{F}}) = \zeta^*$. In both cases, $F(\mathbb{K}) \neq \emptyset$ since we have assumed that the feasible region of POP (1.3) is nonempty.

For Example 3.1, we see that

$$\begin{aligned}
\boldsymbol{\Delta}^{\mathcal{F}} &= \left\{\begin{pmatrix} \mu & \mu x_1 & \mu x_2 & \mu x_1x_2 \\ \mu x_1 & \mu x_1^2 & \mu x_1x_2 & \mu x_1^2x_2 \\ \mu x_2 & \mu x_1x_2 & \mu x_2^2 & \mu x_1x_2^2 \\ \mu x_1x_2 & \mu x_1^2x_2 & \mu x_1x_2^2 & \mu x_1^2x_2^2 \end{pmatrix} \in \mathbb{S}_+^{\mathcal{F}} : \begin{pmatrix} \mu \\ x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}_+^3 \right\}, \\
\boldsymbol{\Gamma}^{\mathcal{F}} &= \left\{\begin{pmatrix} x_0^4 & x_0^3x_1 & x_0^3x_2 & x_0^2x_1x_2 \\ x_0^3x_1 & x_0^2x_1^2 & x_0^2x_1x_2 & x_0x_1^2x_2 \\ x_0^3x_2 & x_0^2x_1x_2 & x_0^2x_2^2 & x_0x_1x_2^2 \\ x_0^2x_1x_2 & x_0x_1^2x_2 & x_0x_1x_2^2 & x_1^2x_2^2 \end{pmatrix} \in \mathbb{S}_+^{\mathcal{F}} : \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}_+^3 \right\}.
\end{aligned} \tag{3.4}$$

It can be easily verified that $\boldsymbol{\Delta}^{\mathcal{F}} \cap \boldsymbol{\Gamma}^{\mathcal{F}}$ coincides with the union of

$$\left\{\begin{pmatrix} \mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{S}_+^{\mathcal{F}} : \mu \in \mathbb{R}_+ \right\}$$

and

$$\left\{\begin{pmatrix} 1 & x_1 & x_2 & x_1x_2 \\ x_1 & x_1^2 & x_1x_2 & x_1^2x_2 \\ x_2 & x_1x_2 & x_2^2 & x_1x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^2x_2^2 \end{pmatrix} \in \mathbb{S}_+^{\mathcal{F}} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}_+^2 \right\}$$

Hence they are different.

Now, we are ready to present the main theorem of this section.

**Theorem 3.2.** *Suppose that Condition (I) holds for* $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$ *and* $\mathbb{K} = \mathbf{\Delta}^{\mathcal{F}}$. *Then,*

(i) $\zeta^p(co\ \mathbf{\Delta}^{\mathcal{F}}) = \zeta^*$.

(ii) *Assume that* $\zeta^*$ *is finite, we have that*

$$\zeta^p(co\ \mathbf{\Gamma}^{\mathcal{F}}) = \begin{cases} \zeta^* & \textit{if Condition (IV) holds for } \mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}, \\ -\infty & \textit{otherwise.} \end{cases}$$

*Proof.* We observe that

$$F_0(\mathbf{\Delta}^{\mathcal{F}}) = \left\{ \boldsymbol{X} \in \mathbb{S}^{\mathcal{F}} \ \middle| \ \begin{array}{c} \boldsymbol{X} \in \mathbf{\Delta}^{\mathcal{F}},\ \langle \boldsymbol{H}^0,\ \boldsymbol{X} \rangle = 0, \\ \langle \boldsymbol{Q}^k,\ \boldsymbol{X} \rangle = 0\ (k = 1, 2, \ldots, m) \end{array} \right\} = \{\boldsymbol{O}\}.$$

Therefore Condition (IV) holds for $\mathbb{K} = \mathbf{\Delta}^{\mathcal{F}}$, and both assertions (i) and (ii) follow from (iv) of Theorem 2.1. ☐

We note that COP (1.1) with $\mathbb{K} = \text{co}\ \mathbf{\Delta}^{\mathcal{F}}$ is similar (essentially equivalent) to the one obtained by applying the canonical convexification procedure [22] to POP (1.3). Next, we investigate Conditions (I) in detail for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$ and $\mathbb{K} = \mathbf{\Delta}^{\mathcal{F}}$ in cases where $\mathbb{J} = \mathbb{R}^n$ and $\mathbb{J} = \mathbb{R}^n_+$. First, suppose that $\mathbb{J} = \mathbb{R}^n$. Then both $\mathbf{\Delta}^{\mathcal{F}}$ and $\mathbf{\Gamma}^{\mathcal{F}}$ are included in $\mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{L}^{\mathcal{F}}$. Hence their duals $\left(\mathbf{\Delta}^{\mathcal{F}}\right)^*$ and $\left(\mathbf{\Gamma}^{\mathcal{F}}\right)^*$ include $\text{cl}\left(\mathbb{S}^{\mathcal{F}}_+ + \left(\mathbb{L}^{\mathcal{F}}\right)^\perp\right) = \mathbb{S}^{\mathcal{F}}_+ + \left(\mathbb{L}^{\mathcal{F}}\right)^\perp$ (see Lemma 2.6). Therefore, if $\boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{F}}_+ + \left(\mathbb{L}^{\mathcal{F}}\right)^\perp$, or, equivalently, if $f^k(\boldsymbol{x}) = \langle \boldsymbol{Q}^k,\ \boldsymbol{x}^{\square\mathcal{F}} \rangle$ is an SOS polynomial ($k = 1, 2, \ldots, m$) (see Lemma 2.3 and 2.4), then Condition (I) is satisfied for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$ and $\mathbb{K} = \mathbf{\Delta}^{\mathcal{F}}$. If a polynomial equation $g(\boldsymbol{x}) = 0$ is given, it is equivalent to have $(g(\boldsymbol{x}))^2 = 0$. Thus, Condition (I) for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$ and $\mathbb{K} = \mathbf{\Delta}^{\mathcal{F}}$ is not a strong assumption.

Suppose that $\mathbb{J} = \mathbb{R}^n_+$. Then,

$$\begin{aligned} \mathbf{\Delta}^{\mathcal{F}} &= \left\{ \mu \boldsymbol{x}^{\square\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^n_+,\ \mu \geq 0 \right\} \subset (\mathbb{C}^{\mathcal{F}})^* \cap \mathbb{L}^{\mathcal{F}} \subset \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{N}^{\mathcal{F}} \cap \mathbb{L}^{\mathcal{F}}, \\ \left(\mathbf{\Delta}^{\mathcal{F}}\right)^* &\supset \text{cl}\ \left(\mathbb{C}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^\perp\right) \supset \text{cl}\ \left(\mathbb{S}^{\mathcal{F}}_+ + \mathbb{N}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^\perp\right), \end{aligned}$$

where

$$\begin{aligned} \mathbb{C}^{\mathcal{F}} &= \left\{ \boldsymbol{Y} \in \mathbb{S}^{\mathcal{F}} : \begin{array}{l} (\xi_{\boldsymbol{\alpha}} : \mathcal{F})^T \boldsymbol{Y} (\xi_{\boldsymbol{\alpha}} : \mathcal{F}) \geq 0 \\ \text{for every } (\xi_{\boldsymbol{\alpha}} : \mathcal{F}) \geq \boldsymbol{0} \end{array} \right\} \text{ (the copositive cone)}, \\ (\mathbb{C}^{\mathcal{F}})^* &= \left\{ \boldsymbol{X} \in \mathbb{S}^{\mathcal{F}} : \langle \boldsymbol{X},\ \boldsymbol{Y} \rangle \geq 0 \text{ for every } \boldsymbol{Y} \in \mathbb{C}^{\mathcal{F}} \right\} \\ &= \text{co}\ \left\{ (\xi_{\boldsymbol{\alpha}} : \mathcal{F})(\xi_{\boldsymbol{\alpha}} : \mathcal{F})^T \in \mathbb{S}^{\mathcal{F}} : (\xi_{\boldsymbol{\alpha}} : \mathcal{F}) \geq \boldsymbol{0} \right\} \\ &\quad \text{(the completely positive cone)}, \\ \mathbb{N}^{\mathcal{F}} &= \left\{ \boldsymbol{X} \in \mathbb{S}^{\mathcal{F}} : X_{\boldsymbol{\alpha\beta}} \geq 0\ ((\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \square\mathcal{F}) \right\} \\ &\quad \text{(the cone of nonnegative matrices)}. \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbf{\Gamma}^{\mathcal{F}} &= \left\{ \mu \boldsymbol{x}^{\square\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^n_+,\ \mu \geq 0 \right\} \subset (\mathbb{C}^{\mathcal{F}})^* \cap \mathbb{L}^{\mathcal{F}} \subset \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{N}^{\mathcal{F}} \cap \mathbb{L}^{\mathcal{F}}, \\ \left(\mathbf{\Gamma}^{\mathcal{F}}\right)^* &\supset \text{cl}\ \left(\mathbb{C}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^\perp\right) \supset \text{cl}\ \left(\mathbb{S}^{\mathcal{F}}_+ + \mathbb{N}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^\perp\right). \end{aligned}$$

Therefore, if $\boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{F}}_+ + \mathbb{N}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^{\perp}$ (or, less restrictively, if $\boldsymbol{Q}^k \in \mathbb{C}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^{\perp}$) $(k = 1, 2, \ldots, m)$, then Condition (I) is satisfied for $\mathbb{K} = \boldsymbol{\Delta}^{\mathcal{F}}$ and $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$.

Now we focus on Condition (IV) with $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$. Recall that $\tau = \max\{|\boldsymbol{\alpha}| : \boldsymbol{\alpha} \in \mathcal{F}\}$. Let $\overline{\mathcal{F}} = \{\boldsymbol{\alpha} \in \mathcal{F} : |\boldsymbol{\alpha}| = \tau\}$, and define $\overline{\boldsymbol{Q}}^k = (\overline{Q}^k_{\boldsymbol{\alpha\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}}$ and $\bar{f}^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \overline{\mathcal{F}} + \overline{\mathcal{F}}]$ such that

$$\overline{Q}^k_{\boldsymbol{\alpha\beta}} = \begin{cases} Q^k_{\boldsymbol{\alpha\beta}} & \text{if } \boldsymbol{\alpha},\ \boldsymbol{\beta} \in \overline{\mathcal{F}} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \bar{f}^k(\boldsymbol{x}) = \langle \overline{\boldsymbol{Q}}^k,\, \boldsymbol{x}^{\square\mathcal{F}} \rangle$$

$(k = 0, 1, \ldots, m)$.

**Lemma 3.3.** *Condition (IV) with* $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$ *holds if and only if*

$$\bar{f}^0(\boldsymbol{x}) \geq 0 \text{ if } \boldsymbol{x} \in \mathbb{J} \ \text{ and } \bar{f}^k(\boldsymbol{x}) = 0 \ (k = 1, 2, \ldots, m). \tag{3.5}$$

*Proof.* (i) "If part": Suppose that (3.5) holds. Let $\boldsymbol{X} \in F_0(\boldsymbol{\Gamma}^{\mathcal{F}})$. By the definition of $F_0(\boldsymbol{\Gamma}^{\mathcal{F}})$, there is an $(x_0, \boldsymbol{x}) \in \mathbb{R}_+ \times \mathbb{J}$ such that

$$\begin{aligned} \boldsymbol{X} &= (x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T, \\ 0 &= \langle \boldsymbol{H}^0,\, \boldsymbol{X} \rangle = X_{\boldsymbol{00}} = x_0^2, \ \textit{i.e.}, x_0 = 0, \\ 0 &= \langle \boldsymbol{Q}^k,\, \boldsymbol{X} \rangle = \left\langle \boldsymbol{Q}^k,\, (x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T \right\rangle \ (k = 1, 2, \ldots, m). \end{aligned}$$

If follows from (3.2), $x_0 = 0$ and the first identity above that

$$X_{\boldsymbol{\alpha\beta}} = \begin{cases} \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} & \text{if } \boldsymbol{\alpha},\ \boldsymbol{\beta} \in \overline{\mathcal{F}}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

Hence we obtain that

$$\langle \boldsymbol{Q}^k,\, \boldsymbol{X} \rangle = \left\langle \overline{\boldsymbol{Q}}^k,\, \boldsymbol{x}^{\square\mathcal{F}} \right\rangle = \bar{f}^k(\boldsymbol{x}) \ (k = 0, 1, 2, \ldots, m). \tag{3.7}$$

Therefore (3.5) implies Condition (IV) with $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$.

(ii) "Only if part": Suppose that Condition (IV) with $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$ holds. For an arbitrary chosen $\boldsymbol{x} \in \mathbb{J}$ such that $\bar{f}^k(\boldsymbol{x}) = 0$ $(k = 1, 2, \ldots, m)$, let $x_0 = 0$ and $\boldsymbol{X} = (x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T$. Then (3.6) and (3.7) follows. Therefore Condition (IV) with $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$ implies (3.5). □

The condition (3.5) holds in the following cases:

(a) $\overline{\boldsymbol{Q}}^0 = \boldsymbol{O} \in \mathbb{S}^{\mathcal{F}}$ or $\bar{f}^0(\boldsymbol{x})$ is an identically zero polynomial, i.e., $\deg(f^0(\boldsymbol{x})) < 2\tau$.

(b) $\{\boldsymbol{x} \in \mathbb{J} : \bar{f}^k(\boldsymbol{x}) = 0 \ (k = 1, 2, \ldots, m)\} = \{\boldsymbol{0}\}$.

We note that (a) can be always satisfied by choosing a nonempty finite subset $\mathcal{F}$ of $\mathbb{Z}^n_+$ such that $f^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F} + \mathcal{F}]$ $(k = 0, 1, 2, \ldots, m)$ and $\deg(f^0(\boldsymbol{x})) < 2\tau = 2\max\{|\boldsymbol{\alpha}| : \boldsymbol{\alpha} \in \mathcal{F}\}$, and that (b) implies that the feasible region of POP (1.3) is bounded.

For Example 3.1, we observe that $F_0(\boldsymbol{\Gamma}^{\mathcal{F}}) =$

$$\left\{ \boldsymbol{X} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1^2 x_2^2 \end{pmatrix} \in \mathbb{S}^{\mathcal{F}}_+ : \begin{array}{l} (x_1, x_2) \in \mathbb{R}^2_+, \\ \langle \boldsymbol{Q}^1,\, \boldsymbol{X} \rangle = Q^1_{(1,1)(1,1)} X_{(1,1)(1,1)} = 0 \end{array} \right\} = \{\boldsymbol{O}\},$$

where $\boldsymbol{Q}^1$ is given as in (2.5). Thus, Condition (IV) is satisfied for $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$.

If the cone co $\boldsymbol{\Gamma}$ is closed, *i.e.*, Condition (II) is satisfied for $\mathbb{K} = \text{co}\, \boldsymbol{\Gamma}$ in addition to Condition (I), then we can introduce the primal-dual pair of Lagrangian-conic relaxation problems (2.1) and (2.2) for $\mathbb{K} = \text{co}\, \boldsymbol{\Gamma}$, and the relation $\left(\eta^d(\lambda, \mathbb{K}) = \eta^p(\lambda, \mathbb{K})\right)\!\uparrow\!\lambda\ = \zeta^d(\mathbb{K}) \leq \zeta^p(\mathbb{K})$ follows by (ii) of Theorem 2.1. Here $(\ )\!\uparrow\!\lambda$ means a monotonic increase as $\lambda \to \infty$, satisfying the equality in the parenthesis. The cone $\boldsymbol{\Gamma}^{\mathcal{F}}$ as well as its convex hull co $\boldsymbol{\Gamma}^{\mathcal{F}}$, however, are not necessarily closed. In fact, $\boldsymbol{\Gamma}^{\mathcal{F}}$ given in (3.4) is not closed. To see this, let

$$\overline{\boldsymbol{X}} \ = \ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Then $\overline{\boldsymbol{X}} \not\in \boldsymbol{\Gamma}^{\mathcal{F}}$, but the matrix $(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{F})^T \in \boldsymbol{\Gamma}^{\mathcal{F}}$ with $(x_0, x_1, x_2) = (\epsilon, \epsilon, 1/\epsilon)$ converges to $\overline{\boldsymbol{X}}$ as $\epsilon \to 0^+$.

**Lemma 3.4.** *Assume that* $\boldsymbol{0} \in \mathcal{F}$ *and* $\tau\boldsymbol{e}_i \in \mathcal{F}$ $(i = 1, 2, \dots, n)$*, where* $\boldsymbol{e}_i$ *denotes the i-th unit coordinate vector in* $\mathbb{R}^n$*. Then,* $\boldsymbol{\Gamma}^{\mathcal{F}}$ *and co* $\boldsymbol{\Gamma}^{\mathcal{F}}$ *are closed.*

*Proof.* Let $\{\boldsymbol{X}^p : p = 1, 2, \dots\}$ be a sequence in $\boldsymbol{\Gamma}^{\mathcal{F}}$ converging to $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}$. By $\boldsymbol{X}^p \in \boldsymbol{\Gamma}^{\mathcal{F}}$, there exists $(x_0^p, \boldsymbol{x}^p) \in \mathbb{R}_+ \times \mathbb{J}$ such that $X^p_{\boldsymbol{\alpha\beta}} = (x_0^p)^{\tau-|\boldsymbol{\alpha}|}(\boldsymbol{x}^p)^{\boldsymbol{\alpha}}(x_0^p)^{\tau-|\boldsymbol{\beta}|}(\boldsymbol{x}^p)^{\boldsymbol{\beta}}$, which converges to $\overline{X}_{\boldsymbol{\alpha\beta}}$ as $p \to \infty$ for $((\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \square\mathcal{F})$. Specifically, $(x_0^p)^{\tau-|\boldsymbol{\alpha}|}(\boldsymbol{x}^p)^{\boldsymbol{\alpha}}(x_0^p)^{\tau-|\boldsymbol{\alpha}|}(\boldsymbol{x}^p)^{\boldsymbol{\alpha}}$ converges to $\overline{X}_{\boldsymbol{\alpha\alpha}}$ for $\boldsymbol{\alpha} = \boldsymbol{0} \in \mathcal{F}$ and $\boldsymbol{\alpha} = \tau\boldsymbol{e}_i \in \mathcal{F}$ $(i = 1, 2, \dots, n)$. Thus, $(x_0^p)^{2\tau}$ and $(x_i^p)^{2\tau}$ $(i = 1, 2, \dots, n)$ converge to $\overline{X}_{\boldsymbol{00}}$ and $\overline{X}_{(\tau\boldsymbol{e}_i)(\tau\boldsymbol{e}_i)}$ $(i = 1, 2, \dots, n)$, respectively. This implies that the sequence $\{(x_0^p, \boldsymbol{x}^p) : p = 1, 2, \dots\}$ is bounded, and we can take a subsequence which converges to $(\bar{x}_0, \bar{\boldsymbol{x}}) \in \mathbb{R}_+ \times \mathbb{J}$. Therefore,

$$\overline{\boldsymbol{X}} = ((\bar{x}_0)^{\tau-|\boldsymbol{\alpha}|}(\bar{\boldsymbol{x}})^{\boldsymbol{\alpha}} : \mathcal{F})((\bar{x}_0)^{\tau-|\boldsymbol{\alpha}|}(\bar{\boldsymbol{x}})^{\boldsymbol{\alpha}} : \mathcal{F})^T \in \boldsymbol{\Gamma}^{\mathcal{F}},$$

and we have shown that $\boldsymbol{\Gamma}^{\mathcal{F}}$ is closed. The closedness of co $\boldsymbol{\Gamma}^{\mathcal{F}}$ follows from Lemma 3.1 of [3]. ☐

Without loss of generality, the assumption of the previous lemma can be satisfied by adding $\boldsymbol{0} \in \mathbb{Z}^n_+$ and $\tau\boldsymbol{e}_i \in \mathbb{Z}^n_+$ $(i = 1, 2, \dots, n)$ to $\mathcal{F}$ if necessary. For Example 3.1, one can add $(2, 0)$ and $(0, 2)$ to $\mathcal{F}$ to satisfy the assumption.

The MC relaxation problem proposed in [3] as an extension of the CPP relaxation for QOPs to POPs is essentially equivalent to COP (1.1) with $\mathbb{K} = \text{co}\, \boldsymbol{\Gamma}^{\mathcal{F}}$. A hierarchy of copositivity conditions assumed there is weaker than Condition (I) and can be regarded as a generalization of Condition (I). We have assumed a stronger condition here, Condition (I), to consistently derive the Lagrangian-conic relaxation (2.1) in the unified framework. On the other hand, the additional condition on zeros at infinity assumed in Lemma 3.1 of [4], which was also assumed in [22] for a canonical convexification procedure for a class of POPs, is stronger than Condition (IV).

If $\mathcal{F} = \left\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : |\boldsymbol{\alpha}| \leq 1\right\}$, then POP (1.3) becomes a QOP. In this case, $\mathbb{L}^{\mathcal{F}} = \mathbb{S}^{\mathcal{F}}$ and $(\mathbb{L}^{\mathcal{F}})^{\perp} = \{\boldsymbol{0}\}$, and the previous discussions correspond to Section 4.2 of Part I [4], where the convexification of a linearly constrained QOP with complementarity condition was discussed.

## 4 A Hierarchy of Lagrangian-SDP Relaxations for POPs

In this section, we propose a hierarchy of Lagrangian-SDP relaxations for POPs by combining the approach in [2, 4, 17] for deriving the Lagrangian-CPP and Lagrangian-DNN relaxations for a class of QOPs with the hierarchy of SDP relaxations proposed by [20] for POPs. The motivation for combining these two approaches, which have been studied almost independently, is to develop efficient and effective numerical methods for POPs.

Fixing $\mathbb{J} = \mathbb{R}^n$ in POP (1.3), we consider a POP of the form

$$\zeta^* \quad := \quad \inf \left\{ f^0(\boldsymbol{x}) \;\middle|\; \boldsymbol{x} \in \mathbb{R}^n,\ f^k(\boldsymbol{x}) = 0 \ (k = 1, 2, \ldots, m) \right\} \tag{4.1}$$

throughout this section. Then an SOS relaxation of POP (4.1) is given by

$$\bar{\zeta}^d_\infty \quad := \quad \sup \left\{ z_0 \in \mathbb{R} \;\middle|\; f^0(\boldsymbol{x}) - z_0 \in \mathbb{SOS}[\boldsymbol{x}] + \sum_{k=1}^m \mathbb{R}[\boldsymbol{x}] f^k(\boldsymbol{x}) \right\}. \tag{4.2}$$

For main theoretical results which we will establish, we introduce

**Condition (VI)** (Archimedean condition) The feasible region $\left\{\boldsymbol{x} \in \mathbb{R}^n : f^k(\boldsymbol{x}) = 0 \right.$ $\left.(k = 1, 2, \ldots, m)\right\}$ is nonempty and bounded, and the set $\{\boldsymbol{x} : p(\boldsymbol{x}) \geq 0\}$ is bounded for some $p(\boldsymbol{x}) \in \mathbb{SOS}[\boldsymbol{x}] + \sum_{k=1}^m \mathbb{R}[\boldsymbol{x}] f^k(\boldsymbol{x})$.

**Lemma 4.1.**

(i) $\bar{\zeta}^d_\infty \leq \zeta^*$.

(ii) *If Condition (VI) holds, then* $\bar{\zeta}^d_\infty = \zeta^*$.

*Proof.* (i) If $\boldsymbol{x}$ is a feasible solution of POP (4.1) and $z_0$ is a feasible solution of (4.2), then $f^0(\boldsymbol{x}) - z_0 \in \mathbb{SOS}[\boldsymbol{x}]$, which implies $z_0 \leq f^0(\boldsymbol{x})$. Thus $\bar{\zeta}^d_\infty \leq \zeta^*$ follows.

(ii) We know that the feasible region of POP (4.1) can be written as

$$\left\{ \boldsymbol{x} \in \mathbb{R}^n : f^k(\boldsymbol{x}) \geq 0,\ -f^k(\boldsymbol{x}) \geq 0 \ (k = 1, 2, \ldots, m) \right\}$$

and that $\mathbb{SOS}[\boldsymbol{x}] - \mathbb{SOS}[\boldsymbol{x}] = \mathbb{R}^n[\boldsymbol{x}]$. Then the lemma follows directly from Lemma 4.1 of [23]. See also Section 4 of [20]. □

## 4.1 Reducing SOS problem (4.2) to a simpler SOS problem

In this subsection, we establish the equivalence between SOS problem (4.2) and the following simpler SOS problem

$$\zeta^d_\infty \quad := \quad \sup \left\{ y_0 \in \mathbb{R} \;\middle|\; \begin{array}{l} f^0(\boldsymbol{x}) - y_0 \in \mathbb{SOS}[\boldsymbol{x}] - \displaystyle\sum_{k=1}^m y_k \boldsymbol{\Theta}[\boldsymbol{x}] (f^k(\boldsymbol{x}))^2, \\ (y_0, y_1, \ldots, y_m) \in \mathbb{R}^{1+m} \end{array} \right\}, \tag{4.3}$$

where

$$\begin{array}{l} \mathcal{A}_\tau = \left\{ \boldsymbol{\alpha} \in \mathbb{Z}^n_+ : |\boldsymbol{\alpha}| \leq \tau \right\} \ (\tau \in \mathbb{Z}_+),\ \theta_\tau(\boldsymbol{x}) = \displaystyle\sum_{\boldsymbol{\alpha} \in \mathcal{A}_\tau} \boldsymbol{x}^{2\boldsymbol{\alpha}} \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\tau], \\ \boldsymbol{\Theta}[\boldsymbol{x}] = \{\theta_\tau(\boldsymbol{x}) : \tau \in \mathbb{Z}_+\} \subset \mathbb{SOS}[\boldsymbol{x}] \subset \mathbb{R}[\boldsymbol{x}]. \end{array} \tag{4.4}$$

The last relation in (4.4) implies that $\left\{ y_k \boldsymbol{\Theta}[\boldsymbol{x}] f^k(\boldsymbol{x}) : y_k \in \mathbb{R} \right\} \subset \mathbb{R}[\boldsymbol{x}]$ $(k = 1, 2, \ldots, m)$; hence (4.3) can be regarded as a subproblem of (4.2).The following lemma shows their equivalence.

**Theorem 4.2.** $\bar{\zeta}^d_\infty = \zeta^d_\infty$.

*Proof.* (i) Proof of $\zeta^d_\infty \leq \bar{\zeta}^d_\infty$. We know that $\mathbb{R}[\boldsymbol{x}] \supset -y_k \boldsymbol{\Theta}[\boldsymbol{x}] f^k(\boldsymbol{x})$ for every $y_k \in \mathbb{R}$ ($k = 1,\dots,m$). This implies that if $(y_0, y_1, \dots, y_m)$ is a feasible solution of (4.3), then $z_0 = y_0$ is a feasible solution of (4.2). Therefore, the inequality $\zeta^d_\infty \leq \bar{\zeta}^d_\infty$ follows.

(ii) Proof of $\zeta^d_\infty \geq \bar{\zeta}^d_\infty$. Let $z_0$ be a feasible solution of (4.2) and $\epsilon$ an arbitrary positive number. We show that there is a feasible solution $(y_0, y_1, \dots, y_m)$ of (4.3) with objective value $y_0 = z_0 - \epsilon$. Since $z_0$ is feasible solution of (4.2), we see that

$$\mathbb{SOS}[\boldsymbol{x}] \quad \ni \quad f^0(\boldsymbol{x}) - z_0 - \sum_{k=1}^m \psi^k(\boldsymbol{x}) f^k(\boldsymbol{x}) \tag{4.5}$$

for some $\psi^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ ($k = 1,\dots,m$). Let $\tau^k = \deg(\psi^k(\boldsymbol{x}))$, so that $\psi^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{A}_{\tau^k}]$ ($k = 1,\dots,m$). Then, each polynomial $\psi^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{A}_{\tau^k}]$ can be represented as

$$\psi^k(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \psi^k_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \ (k = 1,\dots,m).$$

Substituting these identities into the relation (4.5), we get

$$\mathbb{SOS}[\boldsymbol{x}] \quad \ni \quad f^0(\boldsymbol{x}) - z_0 - \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \psi^k_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x}). \tag{4.6}$$

Choose a $\rho > 0$ such that $\epsilon - \left(\sum_{k=1}^m |\mathcal{A}_{\tau^k}|\right)(1/(2\rho))^2 > 0$, and let $y_k = \max\{(\rho\psi^k_{\boldsymbol{\alpha}})^2 : \boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}\}$ ($k = 1,\dots,m$). Then,

$$\begin{aligned}
\mathbb{SOS}[\boldsymbol{x}] \quad \ni \quad & \left(\epsilon - \left(\sum_{k=1}^m |\mathcal{A}_{\tau^k}|\right)(1/(2\rho))^2\right) \\
& + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \left(\rho\psi^k_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x}) + 1/(2\rho)\right)^2 \\
& + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \left(y_k - (\rho\psi^k_{\boldsymbol{\alpha}})^2\right) (\boldsymbol{x}^{\boldsymbol{\alpha}} f^i(\boldsymbol{x}))^2 \\
= \quad & \epsilon + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \psi^i_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x}) \ + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} y_k \left(\boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x})\right)^2.
\end{aligned}$$

It follows from (4.6) that

$$\begin{aligned}
\mathbb{SOS}[\boldsymbol{x}] \quad \ni \quad & f^0(\boldsymbol{x}) - z_0 - \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \psi^k_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x}) \\
& + \epsilon + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} \psi^k_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x}) \ + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} y_k \left(\boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x})\right)^2 \\
= \quad & f^0(\boldsymbol{x}) - (z_0 - \epsilon) + \sum_{k=1}^m \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k}} y_k \left(\boldsymbol{x}^{\boldsymbol{\alpha}} f^k(\boldsymbol{x})\right)^2 \\
= \quad & f^0(\boldsymbol{x}) - (z_0 - \epsilon) + \sum_{k=1}^m y_k \theta_{\tau^k}(\boldsymbol{x}) \left(f^k(\boldsymbol{x})\right)^2.
\end{aligned}$$

Therefore, we have shown that

$$\begin{aligned} f^0(\boldsymbol{x}) - (z_0 - \epsilon) &\in \mathbb{SOS}[\boldsymbol{x}] - \sum_{k=1}^m y_k \theta_{\tau^k}(\boldsymbol{x}) \left(f^k(\boldsymbol{x})\right)^2 \\ &\subset \mathbb{SOS}[\boldsymbol{x}] - \sum_{k=1}^m y_k \boldsymbol{\Theta}[\boldsymbol{x}] \left(f^k(\boldsymbol{x})\right)^2, \end{aligned}$$

and that $(y_0, y_1) = (z_0 - \epsilon, z_1)$ is a feasible solution of SOS problem (4.3). Thus we have $\zeta^d_\infty \geq y_0 = z_0 - \epsilon$ for all $\epsilon > 0$. This implies that $\zeta^d_\infty \geq z_0$ for any feasible $z_0$ of (4.2). Hence $\zeta^d_\infty \geq \bar{\zeta}^d_\infty$. □

## 4.2 A hierarchy of finite SOS subproblems of (4.3) for numerical computation

The SOS problem (4.3) that attains the exact optimal value $\zeta^*$ of POP (4.1) under the assumption of Lemma 4.1 cannot be solved numerically because the degree of sum of squares of polynomials involved is not bounded. For numerical computation of lower bounds of $\zeta^*$ which converges to $\zeta^*$, we introduce a hierarchy of SOS subproblems of (4.3) by bounding the degree of the SOS polynomials to be used with an increasing sequence of finite integers as in a similar way to the hierarchy of Lasserre's SDP relaxation [20] for POPs.

Let $\omega_{\min} = \max\{\lceil \deg(f^0(\boldsymbol{x}))/2 \rceil,\ \deg(f^k(\boldsymbol{x}))\ (k = 1, 2, \ldots, m)\}$. For every $\omega \in \mathbb{Z}_+$ not less than $\omega_{\min}$, we consider the following SOS problem:

$$\zeta^d_\omega \quad := \quad \sup \left\{ y_0 \in \mathbb{R} \;\middle|\; \begin{array}{l} f^0(\boldsymbol{x}) - y_0 + \displaystyle\sum_{k=1}^m y_k \theta_{\tau^k}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega], \\ (y_0, y_1, \ldots, y_m) \in \mathbb{R}^{1+m} \end{array} \right\}, \quad (4.7)$$

where

$$\tau^k(\omega) \quad = \quad \omega - \deg(f^k(\boldsymbol{x}))\ (k = 1, \ldots, m). \quad (4.8)$$

We note that

$$\begin{aligned} &f^0(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{A}_{\omega_{\min}} + \mathcal{A}_{\omega_{\min}}] \subset \mathbb{R}[\boldsymbol{x}, \mathcal{A}_\omega + \mathcal{A}_\omega], \\ &\theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega] \subset \mathbb{R}[\boldsymbol{x}, \mathcal{A}_\omega + \mathcal{A}_\omega]\ (k = 1, 2, \ldots, m), \end{aligned} \quad (4.9)$$

Therefore, the degree of polynomials in the SOS problem (4.7) is bounded by $2\omega$. This SOS problem can be solved as an SDP (4.16), as shown in the next subsection.

**Lemma 4.3.** *Suppose that $(y_0, y_1, \ldots, y_m)$ is a feasible solution of (4.7). Then $(y_0, y'_1, \ldots, y'_m)$ is a feasible solution of (4.7) with the same objective value if $y'_k \geq y_k$ $(k = 1, 2, \ldots, m)$.*

*Proof.* The assertion follows from $\theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega]$ $(k = 1, 2, \ldots, m)$. □

**Lemma 4.4.** *Let $\omega_{\min} \leq \omega \in \mathbb{Z}_+$. Then $\zeta^d_\omega \leq \zeta^d_\infty$ and $\zeta^d_\omega$ converges to $\zeta^d_\infty$ monotonically from below as $\omega \to \infty$.*

The inequality $\zeta^d_\omega \leq \zeta^d_\infty$ follows from the definitions of $\theta_\tau(\boldsymbol{x})$ and $\boldsymbol{\Theta}[\boldsymbol{x}]$ in (4.4). Letting $\omega_{\min} \leq \omega_1 < \omega_2$, we show that $\zeta^d_{\omega_1} \leq \zeta^d_{\omega_2}$. Suppose that $(y_0, y_1, \ldots, y_m)$ is a feasible solution of (4.7) with $\omega = \omega_1$. By Lemma 4.3, we may assume that $y_k \geq 0$ $(k = 1, 2, \ldots, m)$. Then,

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_{\omega_1}] \ni f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\tau^k(\omega_1)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2, \quad (4.10)$$

$$\tau^k(\omega_2) > \tau^k(\omega_1) \ (k = 1, \ldots, m),$$

$$\theta_{\tau^k(\omega_2)}(\boldsymbol{x}) \left(f^k(\boldsymbol{x})\right)^2 - \theta_{\tau^k(\omega_1)}(\boldsymbol{x}) \left(f^k(\boldsymbol{x})\right)^2 = \left(\sum_{\boldsymbol{\alpha} \in \mathcal{F}^k} \boldsymbol{x}^{2\boldsymbol{\alpha}}\right) \left(f^k(\boldsymbol{x})\right)^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_{\omega_2}] \quad (k = 1, \ldots, m), \quad (4.11)$$

where $\mathcal{F}^k = \mathcal{A}_{\tau^k(\omega_2)} \backslash \mathcal{A}_{\tau^k(\omega_1)} = \left\{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k(\omega_2)} : \boldsymbol{\alpha} \notin \mathcal{A}_{\tau^k(\omega_1)}\right\} = \left\{\boldsymbol{\alpha} \in \mathcal{A}_{\tau^k(\omega_2)} : |\boldsymbol{\alpha}| > \tau_1^k\right\}$ $(k = 1, \ldots, m)$. It follows from $y_k \geq 0$ $(k = 1, 2, \ldots, m)$, (4.10) and (4.11) that

$$\begin{aligned}
\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_{\omega_2}] &\supset \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_{\omega_1}] + \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_{\omega_2}] \\
&\ni f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\tau^k(\omega_1)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 \\
&\quad + \sum_{k=1}^m y_k \left(\theta_{\tau^k(\omega_2)}(\boldsymbol{x}) \left(f^k(\boldsymbol{x})\right)^2 - \theta_{\tau^k(\omega_1)}(\boldsymbol{x}) \left(f^k(\boldsymbol{x})\right)^2\right) \\
&= f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\tau^k(\omega_2)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2. \qquad (4.12)
\end{aligned}$$

Hence, $(y_0, y_1, \ldots, y_m)$ remains a feasible solution of SOS problem (4.7) with $\omega = \omega_2$. We have shown that $\zeta^d_{\omega_1} \leq \zeta^d_{\omega_2}$.

Finally, we show that $\zeta^d_\omega$ converges to $\zeta^d_\infty$ as $\omega \to \infty$. Let $\epsilon > 0$. Then there exists a feasible solution $(y_0, y_1, \ldots, y_m)$ of (4.3) such that $y_0 \geq \zeta^d_\infty - \epsilon$. Thus,

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega] \ni f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\sigma^k}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2$$

for some $\sigma^k \in \mathbb{Z}_+$ $(k = 1, 2, \ldots, m)$ and some $\omega \geq \max\left\{\sigma^k + \deg\left(f^k(\boldsymbol{x})\right) : k = 1, 2, \ldots, m\right\}$. Now, define $\tau^k(\omega)$ $(k = 1, 2, \ldots, m)$ by (4.8). Then $\tau^k(\omega) \geq \sigma^k$ $(k = 1, 2, \ldots, m)$. Assuming $y_k \geq 0$ $(k = 1, 2, \ldots, m)$ without loss of generality by Lemma 4.3, we can prove that

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega] \ni f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2$$

by the same way as (4.12) has been derived from (4.10). As a result, $(y_0, y_1, \ldots, y_k)$ is a feasible solution of (4.7) with the objective value $y_0 \geq \zeta^d_\infty - \epsilon$. This implies that $\zeta^d_\infty - \epsilon \leq \zeta^d_\omega$. We already know that $\zeta^d_\omega \leq \zeta^d_{\omega_2} \leq \zeta^d_\infty$ if $\omega < \omega_2$. Since $\epsilon > 0$ arbitrary, we have shown that $\zeta^d_\omega$ converges to $\zeta^d_\infty$ as $\omega \to \infty$.

## 4.3 Reducing SOS problem (4.7) to a COP

To derive a COP of the form (1.2) equivalent to SOS problem (4.7), we need to convert the SOS condition

$$f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega] \qquad (4.13)$$

to a linear matrix inequality. By (4.9), we can represent the left hand side of (4.13) as

$$f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 = \left\langle \boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \sum_{k=1}^m \boldsymbol{Q}^k_\omega y_k,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \right\rangle. \qquad (4.14)$$

Here $\boldsymbol{H}^0_\omega$ and $\boldsymbol{Q}^k_\omega$ $(k = 0, 1, \ldots, m)$ are matrices in $\mathbb{S}^{\mathcal{A}_\omega}$ chosen such that

$$\begin{array}{rcl} f^0(\boldsymbol{x}) & = & \left\langle \boldsymbol{Q}^0_\omega,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \right\rangle \in \mathbb{R}[\boldsymbol{x}, \mathcal{A}_\omega + \mathcal{A}_\omega], \\ 1 & = & \left\langle \boldsymbol{H}^0_\omega,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \right\rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega], \\ \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 & = & \left\langle \boldsymbol{Q}^k_\omega,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \right\rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega] \ (k = 1, 2, \ldots, m). \end{array} \tag{4.15}$$

Specifically, $\boldsymbol{H}^0_\omega$ is a matrix in $\mathbb{S}^{\mathcal{A}_\omega}$ whose elements are all zeros except $H^0_{\boldsymbol{00}} = 1$.

By (4.14), we can rewrite the SOS condition (4.13) as

$$\left\langle \boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \sum_{k=1}^m \boldsymbol{Q}^k_\omega y_k,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \right\rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega].$$

By Lemma 2.5, we know that the inclusion relation above is equivalent to $\boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \sum_{k=1}^m \boldsymbol{Q}^k_\omega y_k \in \mathbb{S}^{\mathcal{A}_\omega}_+ + \left(\mathbb{L}^{\mathcal{A}_\omega}\right)^\perp$. Thus, letting $\mathbb{K}_\omega = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{L}^{\mathcal{A}_\omega}$, we obtain the following primal-dual pair of COPs:

$$\zeta^p_\omega \ := \ \inf \left\{ \langle \boldsymbol{Q}^0_\omega,\, \boldsymbol{X} \rangle \ \middle| \ \begin{array}{l} \boldsymbol{X} \in \mathbb{K}_\omega,\ \langle \boldsymbol{H}^0_\omega,\, \boldsymbol{X} \rangle = 1, \\ \langle \boldsymbol{Q}^k_\omega,\, \boldsymbol{X} \rangle = 0 \ (k = 1, 2, \ldots, m) \end{array} \right\}. \tag{4.16}$$

$$\zeta^d_\omega \ := \ \sup \left\{ y_0 \in \mathbb{R} \ \middle| \ \begin{array}{l} \boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \sum_{k=1}^m \boldsymbol{Q}^k_\omega y_k \in \mathbb{K}^*_\omega \\ (y_0, y_1, \ldots, y_m) \in \mathbb{R}^{1+m} \end{array} \right\}. \tag{4.17}$$

In particular, the problem (4.17) is equivalent to the SOS problem (4.7). We note that $\mathbb{K}^*_\omega = \mathbb{S}^{\mathcal{A}_\omega}_+ + \mathbb{L}^{\mathcal{A}_\omega}$, and that both $\mathbb{K}_\omega$ and $\mathbb{K}^*_\omega$ closed. See Lemma 2.6.

## 4.4 Lagrangian-SDP relaxations of COPs (4.16) and (4.17)

If we take $\mathbb{K} = \mathbb{K}_\omega = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{L}^{\mathcal{A}_\omega}$, $\boldsymbol{H}^0 = \boldsymbol{H}^0_\omega$ and $\boldsymbol{Q}^k_\omega$ $(k = 0, 1, 2, \ldots, k)$, then the problems (4.16) and (4.17) coincide with the primal-dual pair of COPs (1.1) and (1.2), respectively. We are now ready to apply the general discussions on COPs (1.1) and (1.2) given in Sections 2, Part I [4]. Let $\boldsymbol{H}^1_\omega = \sum_{k=1}^m \boldsymbol{Q}^k_\omega$ for their Lagrangian-conic relaxation problems (2.1) and (2.2). Then we obtain:

$$\eta^p_\omega(\lambda) \ := \ \inf \left\{ \langle \boldsymbol{Q}^0_\omega + \lambda \boldsymbol{H}^1_\omega,\, \boldsymbol{X} \rangle \ \middle| \ \boldsymbol{X} \in \mathbb{K}_\omega,\ \langle \boldsymbol{H}^0_\omega,\, \boldsymbol{X} \rangle = 1 \right\}. \tag{4.18}$$

$$\eta^d_\omega(\lambda) \ := \ \sup \left\{ y_0 \in \mathbb{R} \ \middle| \ \boldsymbol{Q}^0_\omega + \lambda \boldsymbol{H}^1_\omega - \boldsymbol{H}^0_\omega y_0 \in \mathbb{K}^*_\omega,\ y_0 \in \mathbb{R} \right\}. \tag{4.19}$$

**Theorem 4.5.** *Assume that the feasible region of POP* (4.1) *is nonempty. Let* $\omega_{\min} \leq \omega \in \mathbb{Z}_+$. *The following results hold.*

(i) $\eta^d_\omega(\lambda) = \eta^p_\omega(\lambda) \leq \zeta^d_\omega$ *for every* $\lambda \in \mathbb{R}$. *The problem* (4.19) *attains the optimal value* $\eta^d_\omega(\lambda)$ *at a feasible solution if* $\eta^d_\omega(\lambda)$ *is finite.*

(ii) $\eta^d_\omega(\lambda)$ *converges to* $\zeta^d_\omega$ *monotonically from below as* $\lambda \to \infty$.

(iii) *For any* $\epsilon > 0$, *there exist* $\hat{\omega} \in \mathbb{Z}_+$ *and* $\hat{\lambda} \in \mathbb{R}$ *such that* $\zeta^d_\infty - \epsilon \leq \eta^d_\omega(\lambda) \leq \zeta^d_\infty$ *holds for every* $\omega \geq \hat{\omega}$ *and every* $\lambda \geq \hat{\lambda}$; *roughly speaking* $\eta^d_\omega(\lambda)$ *converges to* $\zeta^d_\infty$ *from below as* $\lambda > 0$ *and* $\omega \in \mathbb{Z}_+$ *both tend to* $\infty$.

(iv) *In addition, if Condition (VI) holds, we can replace* $\zeta^d_\infty$ *by the optimal value* $\zeta^*$ *of POP* (4.1).

*Proof.* For (i) and (ii), it suffices to show that Conditions (I) and (II) hold for $\mathbb{K} = \mathbb{K}_\omega = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{L}^{\mathcal{A}_\omega}$. Then, both assertions follow from (ii) of Theorem 2.1. It follows from the assumption that the feasible region of (4.16) is nonempty. By construction, we know that $\boldsymbol{H}^0_\omega,\ \boldsymbol{Q}^k_\omega \in \mathbb{S}^{\mathcal{A}_\omega}_+$ $(k = 1, 2, \ldots, m)$. On the other hand, $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{L}^{\mathcal{A}_\omega} \subset \mathbb{S}^{\mathcal{A}_\omega}_+$. Therefore $\mathbb{K}^* \supset \mathbb{S}^{\mathcal{A}_\omega}_+$, and Condition (I) follows. On the other hand we know that $\mathbb{S}^{\mathcal{A}_\omega}_+$ and $\mathbb{L}^{\mathcal{A}_\omega}$ are both closed convex cones. And, so is their intersection. Therefore, Condition (II) holds for $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{L}^{\mathcal{A}_\omega}$.

(iii) We have derived the Lagrangian-SDP relaxation problem (4.19) of (4.1) from SDP (4.16). By the same argument as above, we can prove directly that (4.19) is equivalent to SOS problem:

$$\eta^d_\omega(\lambda) \ := \ \sup \left\{ y_0 \in \mathbb{R} \ \middle| \ f^0(\boldsymbol{x}) - y_0 + \lambda \sum_{k=1}^{m} \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\omega],\ y_0 \in \mathbb{R} \right\}.$$

Thus it can be shown that if $\omega_{\min} \leq \omega_1 < \omega_2$, then the inequality $\eta^d_{\omega_1}(\lambda) \leq \eta^d_{\omega_2}(\lambda)$ holds for every $\lambda \geq 0$. Let $\epsilon > 0$. Then, by Lemma 4.4, we can find an $\hat{\omega} \in \mathbb{Z}_+$ such that $\zeta^d_\infty - \epsilon/2 \leq \zeta^d_\omega \leq \zeta^d_\infty$ for every $\omega \geq \hat{\omega}$. Furthermore, by (ii), we can find a $\hat{\lambda} \in \mathbb{R}$ such that $\zeta^d_{\hat{\omega}} - \epsilon/2 \leq \eta^d_{\hat{\omega}}(\lambda) \leq \zeta^d_{\hat{\omega}}$ for every $\lambda \geq \hat{\lambda}$. Hence we obtain $\zeta^d_\infty - \epsilon \leq \zeta^d_{\hat{\omega}} - \epsilon/2 \leq \eta^d_{\hat{\omega}}(\lambda)$. Now if $\omega \geq \hat{\omega}$ and $\lambda \geq \hat{\omega}$, then $\zeta^d_\infty - \epsilon \leq \eta^d_{\hat{\omega}}(\lambda) \leq \eta^d_\omega(\lambda) \leq \zeta^d_\infty$, where the last inequality follows from the discussion above. □

In order to solve the primal-dual pair of COPs (4.18) and (4.19), it is possible to apply the bisection and 1-dimensional Newton methods proposed in Part I [4]. See also the numerical method in [17], consisting of a bisection method (Algorithm A of [17]), a proximal alternating direction multiplier method [12] (Algorithm B) and an accelerated proximal gradient method [7] (Algorithm C).

Now we present a simpler way of deriving the Lagrangian-SDP relaxation, COP (4.18) directly from POP (4.1). Choose $\lambda > 0$ and $\omega \in \mathbb{Z}_+$ such that $\omega \geq \omega_{\min}$ where $\omega_{\min} = \max\{\lceil \deg(f_0(\boldsymbol{x}))/2 \rceil, \deg(f_k(\boldsymbol{x}))\ (k = 1, 2, \ldots, m)\}$. Consider the unconstrained POP

$$\zeta^p(\lambda) \ := \ \inf \left\{ g^0(\lambda, \omega, \boldsymbol{x}) \ : \ \boldsymbol{x} \in \mathbb{R}^n \right\}, \tag{4.20}$$

where $g^0(\lambda, \omega, \boldsymbol{x}) = f^0(\boldsymbol{x}) + \lambda \sum_{k=1}^m \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2$. (See (4.4) and (4.8) for the definition of $\theta_{\tau^k(\omega)}(\boldsymbol{x})$). Since $\theta_{\tau^k(\omega)}(\boldsymbol{x}) \geq 1$ for every $\boldsymbol{x} \in \mathbb{R}^n$, the term $\lambda \sum_{k=1}^m \theta_{\tau^k(\omega)}(\boldsymbol{x})(f^k(\boldsymbol{x}))^2$ added to the objective polynomial $f^0(\boldsymbol{x})$ of POP (4.1) serves as a penalty for violating the equality constraints $f^k(\boldsymbol{x}) = 0$ $(k = 1, 2, \ldots, m)$ of POP (4.1). By the construction of the matrices $\boldsymbol{H}^0_\omega$, $\boldsymbol{Q}^k_\omega$ $(k = 0, 1, \ldots, m)$ (see (4.15)) and $\boldsymbol{H}^1_\omega = \sum_{k=1}^m \boldsymbol{Q}^k_\omega$, the unconstrained POP (4.20) can be rewritten as

$$\zeta^p_\omega(\lambda) \ := \ \inf \left\{ \langle \boldsymbol{Q}^0_\omega + \lambda \boldsymbol{H}^1_\omega,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \rangle \ : \ \langle \boldsymbol{H}^0_\omega,\, \boldsymbol{x}^{\square \mathcal{A}_\omega} \rangle = 1,\ \boldsymbol{x} \in \mathbb{R}^n \right\}.$$

Replacing $\boldsymbol{x}^{\square \mathcal{A}_\omega}$ by a variable matrix $\boldsymbol{X} \in \mathbb{K}_\omega = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{L}^\omega$, we obtain the Lagrangian-SDP relaxation COP (4.18) and its dual (4.19). It follows from $\boldsymbol{x}^{\square \mathcal{A}_\omega} \in \mathbb{K}_\omega$ that $\eta^p_\omega(\lambda) \leq \zeta^p_\omega(\lambda) \leq \zeta^*$. We also see by Theorem 4.5 that if the feasible reagion of POP (4.1) is nonempty and

Condition (VI) holds then $\zeta^p_\omega(\lambda)$ converges to the optimal value $\zeta^*$ of POP (4.1) as $\lambda > 0$ and $\omega \in \mathbb{Z}_+$ both tend to $\infty$. The derivation of the Lagrangian-SDP relaxation COP (4.18) and its dual (4.19) are equivalent to the application of Lasserre's hierarchy of SDP relaxations [20] to (4.20).

### 4.5 A numerical example

We illustrate how $\eta^d_\omega(\lambda)$ converges to the optimal value $\zeta^*$ of POP (4.1) as $\omega$ and $\lambda$ both tend to $\infty$ with a simple numerical example. We also roughly evaluate the performance of our Lagrangian-SDP relaxation formulation (4.18) and (4.19) combined with the PB (projection-bisection) method, which was originally proposed in [17] (see also Section 4 of Part I [4] and Section 2 of [6]) for large scale COPs of the form (1.1).

Let $f^0(\boldsymbol{x})$ be a randomly generated polynomial with degree 4, $m = 1$ and $f^1(\boldsymbol{x}) = \sum_{i=1}^n (2x_i - 1)^2 - 1$, for a POP of the form (4.1). Since the feasible region is nonempty and bounded, the problem has a finite optimal value $\zeta^*$. The numerical experiments were conducted in MATLAB on Mac Pro with Intel Xeon E5 8 core CPU (3.0 GHz). For the PB method, we modified the MATLAB code of the PB method (Algorithm 2.2 in [6]) developed for a class of QOPs. For the primal-dual interior-point method, SeDuMi [24] and SDPT3 [25] were used with default parameters for accuracy. The relative accuracy $\delta = 1.0e$-4 was used for the PB method. We note that the PB method solves the dual problem (4.19) while SeDuMi and SDPT3 solve the primal-dual pair of (4.18) and (4.19) simaltaneously.

In Table 1, the changes of $\eta^d_\omega(\lambda)$ as $\lambda$ and $\omega$ increase are shown for $n$ fixed to 3. The optimal value $\zeta^*$ of POP (4.1) was computed by SparsePOP [26, 27] with SeDuMi. We note that SparsePOP is a MATLAB implementation of Lasserre's hierarchy of SDP relaxations of POPs [20]. When the PB method was applied to (4.19) with $\omega = 2$, $\eta^d_\omega(\lambda)$ increased with $\lambda$ from 100 to 51200, but it then decreased due to numerical errors. On the other hand, when SeDuMi was applied to the primal-dual pair of (4.18) and (4.19), $\eta^d_\omega(\lambda)$ increased with $\lambda$ for each fixed $\omega = 2,\ 3,\ 4$, and it also increased with $\omega$ for each fixed $\lambda$. It eventually attained a better lower bound $-5.02473$ of the optimal value $\zeta^* = -5.02431$ of POP (4.1) at $\omega = 4$ and $\lambda = 102400$. We also observe in Table 1 that SeDuMi provided a better lower bound in less cpu time.

Table 1: $n = 3$. The optimal value is -5.02431. The default parameters for accuracy were used when applying SeDuMi and SDPT3 to (4.18) and (4.19). For the PB method applied to (4.19), the relative accuracy $\delta = 1.0e$-4 was used.

| | Lower Bounds (seconds) | | | |
|---|---|---|---|---|
| $\omega$ | 2 | | 3 | 4 |
| $\lambda$ | PB | SeDuMi | SeDuMi | SeDuMi |
| 100 | -5.48483(1.90e0) | -5.48480(5.97e-2) | -5.41977(7.82e-2) | -5.39395(2.42e-1) |
| 200 | -5.26430(1.52e0) | -5.26423(4.65e-2) | -5.23750(8.58e-2) | -5.22897(2.21e-1) |
| 400 | -5.14748(1.63e0) | -5.14733(6.07e-2) | -5.13594(7.63e-2) | -5.13342e(2.20e-1) |
| 800 | -5.08717e(1.34e0) | -5.08670(3.98e-2) | -5.08160(8.43e-2) | -5.08093(2.44e-1) |
| 1600 | -5.05603(1.28e0) | -5.05574(6.02e-2) | -5.05336(7.36e-2) | -5.05320(2.28e-1) |
| 3200 | -5.04078(1.52e0) | -5.04008(7.08e-2) | -5.03894(1.01e-1) | -5.03890(2.38e-1) |
| 6400 | -5.03540(1.30e0) | -5.03221(4.54e-2) | -5.03165(9.23e-2) | -5.03164(2.01e-1) |
| 12800 | -5.03148(1.33e0) | -5.02826(8.10e-2) | -5.02798(9.34e-2) | -5.02796(1.85e-1) |
| 25600 | -5.02933(1.71e0) | -5.02628(4.89e-2) | -5.02615(1.08e-1) | -5.02614(2.44e-1) |
| 51200 | -5.02689(1.84e0) | -5.02529(4.86e-2) | -5.02515(8.54e-2) | -5.02518(2.09e-1) |
| 102400 | -5.06741(1.48e0) | -5.02479(8.26e-2) | -5.02476(9.95e-2) | -5.02473(2.21e-1) |

We now consider larger problems. The PB method is compared with the direct application of SparsePOP, which invokes SeDuMi or SDPT3 as an SDP solver, to POP (4.1) and the applications of SeDuMi and SDPT3 to (4.18) and (4.19). Table 2 shows the case where $n = 20$ and $\omega = 2$. We observe that

(a) SparsePOP with SDPT3 and SparsePOP with SeDuMi applied directly to POP (4.1) attained the best lower bound, and also performed faster than SDPT3 and SeDuMi applied to (4.18) and (4.19), respectively.

(b) SDPT3 and SeDuMi applied to (4.18) and (4.19) worked effectively in computing better bounds but required longer time than the PB method,

(c) the PB method with smaller $\lambda$ up to 1600 attained a reasonable lower bound in less cpu time but the lower bounds generated got worse for larger $\lambda$ due to numerical errors.

Table 2: $n = 20$ and $\omega = 2$. The default parameters for accuracy were used when applying SeDuMi and SDPT3 to (4.18) and (4.19). For the PB method applied to (4.19), the relative accuracy $\delta = 1.0e$-4 was used.

| | Lower BoundS (seconds) | | |
|---|---|---|---|
| | SparsePOP with SDPT3 for (4.1): -1.92956e+1(5.45e+2) | | |
| | SparsePOP with SeDuMi for (4.1): -1.92956e+1(5.14e+3) | | |
| $\lambda$ | PB | SDPT3 | SeDuMi |
| 100 | -2.45517e1(6.34e1) | -2.45504e1(1.73e3) | -2.45504e1(1.70e4) |
| 200 | -2.23581e1(7.56e1) | -2.23569e1(1.76e3) | -2.23569e1(1.70e4) |
| 400 | -2.10353e1(8.54e1) | -2.10322e1(1.82e3) | -2.10322e1(1.66e4) |
| 800 | -2.02549e1(1.53e2) | -2.02489e1(1.64e3) | -2.02489e1(1.65e4) |
| 1600 | -1.98115e1(1.18e2) | -1.98030e1(1.77e3) | -1.98028e1(1.71e4) |
| 3200 | -2.14671e1(1.47e2) | -1.95590e1(1.73e3) | -1.95589e1(1.73e4) |
| 6400 | -2.82489e1(4.10e2) | -1.94301e1(1.86e3) | -1.94300e1(1.73e4) |
| 12800 | -4.09698e1(3.00e2) | -1.93637e1(1.82e3) | -1.93636e1(1.73e4) |

Table 3 shows the case for $n = 25,\ 30$ and $\omega = 2$. We note that the observation (a) made in the previous case is no longer true in these cases. In general, the dual SDP problem derived as Lasserre's SDP relaxation [20] of a POP has no interior-feasible solution. This is a difficulty that many SDP solvers such as SDPT3 based on the primal-dual interior-point method cannot properly handle. For SeDuMi, we observe that it can effectively deal with such degeneracy, however, it is too slow to process these problems. As far as the speed is concerned, SDPT3 is much faster than SeDuMi for large SDPs, but its speed is still not fast enough and its memory consumption is too large for these problems. The observation (c) on the PB method remains valid and it successfully provided a lower bound for the unknown optimal value of POP (4.1) with $n = 30$ which neither SeDuMi nor SDPT3 could process.

The numerical results reported in Tables 1 to 3 display high potential of our approach based on the Lagrangian-SDP relaxation formulation (4.18) and (4.19) combined with the PB method. As seen in the Tables, the current implementation of the PB method lacks stability for large $\lambda$. A more stable and efficient implementation of the PB method is necessary before fully evaluating the performance of our approach. See also the numerical results reported in [3, 6, 17].

Table 3: $n = 25,\ 30$ and $\omega = 2$. The default parameters for accuracy were used when applying SeDuMi and SDPT3 to (4.18) and (4.19). For the PB method applied to (4.19), the relative accuracy $\delta = 1.0e\text{-}4$ was used. † means that SDPT3 stops with error (termcode = -5, gap = 3.6180e6, pinfeas = 1.1035e-4 and dinfeas = 7.5934e-2). ‡ means that SDPT3 terminated with error (termcode = -5, gap = 1.7981e-3, pinfeas = 7.1405e-11 and dinfeas = 8.0319e-12).

| | Lower Bounds (seconds) | | | |
|---|---|---|---|---|
| $n$ | 25 | | 30 | |
| | SparsePOP+SDPT3: -2.9375e6†(2.5476e3) | | SparsePOP+SDPT3: Out of memory | |
| $\lambda$ | PB | SDPT3 | PB for (4.19) | SDPT3 |
| 100 | -2.88618e1(1.61e2) | -2.88610e1 (1.45e4) | -3.42261e1(3.84e2) | Out of memory |
| 200 | -2.62152e1(1.99e2) | -2.62075e1 (1.49e4) | -3.11497e1(7.75e2) | |
| 400 | -2.45866e1(2.15e2) | -2.45837e1 (1.55e4) | -2.90901e1(4.67e2) | |
| 800 | -2.36180e1(2.38e2) | -2.36062e1 (1.57e4) | -2.78837e1(6.17e2) | |
| 1600 | -2.30519e1(2.90e2) | -2.30385e1 (1.55e4) | -2.72198e1(5.32e2) | |
| 3200 | -3.80742e1(3.90e2) | -2.27227e1 (1.57e4) | -3.09553e1(6.17e2) | |
| 6400 | -2.83800e1(3.37e2) | -3.10021e1‡ (1.24e4) | -3.24649e1(1.08e3) | |
| 12800 | -3.91523e1(4.79e2) | -2.24664e1 (1.49e4) | -5.17664e1(9.43e2) | |

## 5 Concluding Remarks

For POP (1.3) with $\mathbb{J} = \mathbb{R}^n_+$, two different approaches can be used. The first one is the hierarchy of Lagrangian-DNN relaxations, obtained by replacing SDP cones with DNN cones in the construction of the hierarchy of Lagrangian-SDP relaxations in Section 4. The second one is the hierarchy of Lagrangian-SDP relaxation for the reformulated equality constrained POP over $\mathbb{R}^n$, obtained from adding $x_i - x_{i+n}^2 = 0$ $(i = 1, 2, \ldots, n)$ and replacing the cone $\mathbb{R}^n_+$ by $\mathbb{R}^{2n}$. The lower bounds generated by the first hierarchy of Lagrangian-DNN relaxations may not be theoretically guaranteed to converge to the optimal value of the original POP. However, it may work effectively and efficiently for practical problems with a low relaxation order.

Preliminary numerical results have been reported in Section 4. As mentioned in Section 4.5, there remain some issues to be investigated for a stable and efficient implementation of the PB method for the hierarchy of Lagrangian-SDP relaxations of general POPs. The 1-dimensional Newton method proposed in Part I [4] may increase the numerical efficiency.

In addition, handling sparsity in an efficient manner is an important issue. Although a hierarchy of sparse Lagrangian-SDP relaxations was presented in the original version [5] of this paper, it is excluded here for the simplicity of the discussions in this paper. We hope to report extensive numerical results in the near future.

## References

[1] N. Arima, S. Kim and M. Kojima, A quadratically constrained quadratic optimization model for completely positive cone programming, *SIAM J. Optim.* 23 (2013) 2320–2340.

[2] N. Arima, S. Kim and M. Kojima, Simplified copositive and Lagrangian relaxations for linearly constrained quadratic optimization problems in continuous and binary variables, *Pacific J. of Optim.* 10 (2014) 437–451.

[3] N. Arima, S. Kim and M. Kojima, Extension of completely positive cone relaxation to polynomial optimization, *J. Optimiz. Theory App.* 168 (2016) 884-900 .

[4] N. Arima, S. Kim, M. Kojima and K. C. Toh, Lagrangian-conic relaxations, part I: a unified framework and its applications to quadratic optimization problems, to appear in *Pacific J. Optim.* 14 (2017) 161–192.

[5] N. Arima, S. Kim, M. Kojima, and K. C. Toh. Lagrangian-conic relaxations, part II: applications to polynomial optimization problems, Research Report http://www.opt.c.titech.ac.jp/kojima/articles/B-476.pdf B-476, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, January (2014).

[6] N. Arima, S. Kim, M. Kojima, and K.C. Toh, A robust Lagrangian-DNN method for a class of quadratic optimization problems, *Comput. Optim. and Appl.* 66 (2017) 453–479.

[7] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sci.* 2 (2009)183–202 .

[8] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. Quist and T. Terlaky, On copositive programming and standard quadratic optimization problems, *J. Global Optim.* 18 (2000) 301–320.

[9] S. Burer, On the copositive representation of binary and continuous non- convex quadratic programs, *Math. Program.* 120 (2009) 479–495.

[10] S. Burer, Copositive programming, in: Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications, International Series in Operational Research and Management Science, M. Anjos and J. Lasserre (eds), Springer, 2011, pp. 201–218

[11] M.D. Choi, T.Y. Lam and B. Reznick, Sums of squares of real polynomials, *Proc. Symposia Pure Math.*, Choi, M.D., Lam, T.Y., Reznick, B.: Sums of squares of real polynomials. Proc. Symposia Pure Math. 58 (2), 103-126, 1995.

[12] M. Fazel, T. K. Pong, D. Sun and P. Tseng, Hankel matrix rank minimization with applications to system identification and realization, *SIAM J. Matrix Anal. Appl.* 34 (2013) 946–977.

[13] K. Fujisawa, M. Fukuda, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata and M. Yamashita, SDPA (SemiDefinite Programming Algorithm) user's manual — Version 7.05, Research Report B-448, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152-8552, Japan (2008).

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

[15] D. Henrion and J. B. Lasserre, 'GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi, *ACM Trans. Math. Softw.* 29 (2003) 165–194.

[16] S. Kim, M. Kojima, M. Mevissen and M. Yamashita, Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion, *Math. Program.* 129 (2011) 33–68.

[17] S. Kim, M. Kojima and K. C. Toh, A Lagrangian-DNN relaxation: a fast method for computing tight lower bounds for a class of quadratic optimization problems, *Math. Program.* 156 (2016) 161–187.

[18] S. Kim, M. Kojima and K. C. Toh, Doubly nonnegative relaxations for quadratic and polynomial optimization problems with binary and box constraints, Research Report B-483, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152-8552, Japan, July (2016).

[19] M. Kojima, S. Kim and H. Waki, Sparsity in sums of squares of polynomials, *Math. Program.* 103 (2005) 45–62 .

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

[21] J. B. Lasserre, Convergent SDP-relaxation in polynomial optimization with sparsity, *SIAM J. Optim.* 17 (1993) 822–843.

[22] J. Peña, J. C. Vera and L. F. Zuluaga, Positive polynomial on unbounded equality-constrained domains, Manuscript. Available at http://www.optimization-online.org/DB_FILE/2011/05/3047.pdf, June (2011).

[23] M. Putinar, Positive polynomials on compact semi-algebraic sets, *Indiana University Math. J.* 42 (1993) 969–984.

[24] J.F. Strum, SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, *Optim. Methods Softw.* 11 & 12 (1999) 625–653.

[25] R.H. Tütüntü, K.C. Toh and M.J. Todd, Solving semidefinite-quadratic-linear programs using SDPT3, *Math. Program.* 95 (2003) 189–217.

[26] H. Waki, S. Kim, M. Kojima and M. Muramatsu, Sums of squares and semidefinite program relaxation for polynomial optimization problems with structured sparsity, *SIAM J. Optim.* 17 (2006) 218–242.

[27] H. Waki, S. Kim, M. Kojima, M. Muramatsu and H. Sugimoto, Algorithm 883: Sparsepop: A sparse semidefinite programming relaxation of polynomial pptimization problems, *ACM Trans. Math. Softw.* 35 (2008) 15.

[28] H. Waki, S. Kim, M. Kojima, M. Muramtasu and H. Sugimoto, Algorithm 883: SparsePOP: A sparse semidefinite programming relaxation of polynomial optimization problems, *ACM Trans. Math. Softw.* **35** (2008) 15.

---

*Manuscript received 22 February 2016*
*revised 21 December 2016*
*accepted for publication 26 April 2017*

Naohiko Arima
Research and Development Initiative & JST CREST, Chuo University
1-13-27, Kasuga, Bunkyo-ku, Tokyo 112-8551 Japan
E-mail address: nao_arima@me.com

Sunyoung Kim
Department of Mathematics, Ewha W. University
52 Ewhayeodae-gil, Sudaemoon-gu, Seoul 120-750, Korea
E-mail address: skim@ewha.ac.kr

Masakazu Kojima
Department of Industrial and Systems Engineering, Chuo University
Tokyo 192-0393, Japan
E-mail address: kojima@is.titech.ac.jp

Kim-Chuan Toh
Department of Mathematics, National University of Singapore
10 Lower Kent Ridge Road, Singapore 119076
E-mail address: mattohkc@nus.edu.sg