ISSN 1342-2804

Research Reports on Mathematical and Computing Sciences

Lagrangian-Conic Relaxations, Part II: Applications to Polynomial Optimization Problems

Naohiko Arima, Sunyoung Kim,
Masakazu Kojima, and Kim-Chuan Toh

January 2014, B–476

Department of
Mathematical and
Computing Sciences
Tokyo Institute of Technology

SERIES B: Operations Research

B-476 Lagrangian-Conic Relaxations, Part II: Applications to Polynomial Optimization Problems

Naohiko Arima*, Sunyoung Kim†, Masakazu Kojima‡, and Kim-Chuan Toh♯

January 2014

Abstract. We present the moment cone (MC) relaxation and a hierarchy of sparse 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 sparse Lagrangian-SDP relaxations, which is parameterized by a positive integer ω called the relaxation order, is proposed for an equality constrained POP. It is obtained by combining a sparse variant of 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 λ and the relaxation order ω in the hierarchy converges to that of the POP as $\lambda \to \infty$ and $\omega \to \infty$. The hierarchy of sparse 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.

Keywords. Polynomial optimization problem, moment cone relaxation, SOS relaxation, a hierarchy of the Lagrangian-SDP relaxations, exploiting sparsity.

AMS Classification. 90C22, 90C25, 90C26.

* Research and Development Initiative & JST CREST, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo 112-8551 Japan. 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. (nao_arima@me.com).

† Department of Mathematics, Ewha W. University, 52 Ewhayeodae-gil, Sudaemoon-gu, Seoul 120-750 Korea. The research of this author was supported by NRF 2012-R1A1A2-038982. (skim@ewha.ac.kr).

‡ Research and Development Initiative & JST CREST, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo 112-8551 Japan. 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. (kojima@is.titech.ac.jp).

♯ Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, Singapore 119076. Research supported in part by Ministry of Education Academic Research Fund R-146-000-168-112 (mattohkc@nus.edu.sg).

1 Introduction

A unified framework for conic and Lagrangian-conic relaxations in Part I [5] was proposed for quadratic optimization problems (QOPs) and polynomial optimization problems (POPs). From a theoretical viewpoint, this framework is intended to unify and generalize the existing results on the completely positive (CPP) relaxation of QOPs [2, 8, 9, 16] and its extension to POPs, which was called the moment cone (MC) relaxation in [4]. From a practical viewpoint, the Lagrangian-conic relaxation is proposed to solve large-scale conic optimization problems (COPs) obtained from effective conic relaxations, including doubly nonnegative (DNN) and semidefinite programming (SDP) relaxations of QOPs and POPs, by first-order algorithms. In particular, the CPP relaxation and the sparse DNN relaxation for QOPs in Part I were discussed from these viewpoints.

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

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

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 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ $(k = 0, 1, 2, \ldots, m)$.

The first purpose of this paper is to derive the moment cone (MC) relaxation for POP (1). The MC relaxation introduced by Arima, Kim and Kojima [4] for this form of POP under a hierarchy of copositivity conditions is an extension of the CPP relaxation. 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 first convert POP (1) into an equivalent COP which serves as the starting optimization model in the framework. Then, we can apply the general results established on the COP in Part I [5] to derive not only the MC relaxation but also the Lagrange-MC relaxation of (1).

Let $\mathbb{V}$ be a finite dimensional vector space endowed with an inner product $\langle \cdot, \cdot \rangle$, and $\mathbb{K} \subset \mathbb{V}$ a (not necessarily convex) cone. Let $\boldsymbol{H}^0 \in \mathbb{V}$ and $\boldsymbol{Q}^k \in \mathbb{V}$ $(k = 0, 1, 2, \ldots, m)$. The primal COP is of the form:

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

As a basic assumption, the following *copsitivity condition* (Condition (I)) is assumed: $\boldsymbol{O} \neq \boldsymbol{H}^0 \in \mathbb{K}^*$ and $\boldsymbol{Q}^k \in \mathbb{K}^*$ $(k = 1, 2, \ldots, m)$, where $\mathbb{K}^* = \{\boldsymbol{Y} \in \mathbb{V} : \langle \boldsymbol{X}, \boldsymbol{Y} \rangle \geq 0 \text{ for every } \boldsymbol{X} \in \mathbb{K}\}$ (the dual of $\mathbb{K}$). The copositivity condition assumed here is stronger than the hierarchy of copositivity conditions of [4]. The stronger condition is necessary for deriving the Lagrangian-conic relaxation consistently.

For the application of the framework to POP (1), 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($\boldsymbol{\Gamma}$) in (2) represents POP (1), *i.e.*, $\zeta^* = \zeta^p(\boldsymbol{\Gamma})$. Then, the MC relaxation of POP (1) is derived as its convexification, *i.e.*, "the strongest convex relaxation" of COP($\boldsymbol{\Gamma}$) in (2) by replacing $\boldsymbol{\Gamma}$ with its convex hull, co $\boldsymbol{\Gamma}$. For the equivalence between COP($\boldsymbol{\Gamma}$) in (2) and its convexification COP(co $\boldsymbol{\Gamma}$), or for the identity $\zeta^p(\boldsymbol{\Gamma}) = \zeta^p(\text{co}\,\boldsymbol{\Gamma})$, we assume the copositivitiy condition for $\mathbb{K} = \boldsymbol{\Gamma}$. This condition is satisfied for any cone $\mathbb{J} \in \mathbb{R}^n$ if each polynomial $f^k(\boldsymbol{x})$ is a sum of squares polynomials, thus, if $f^k(\boldsymbol{x})$ is replaced by $f^k(\boldsymbol{x})^2$ $(k = 1, 2, \ldots, m)$ in POP (1). Under the copositivity condition, we provide a condition (Condition (IV)) that characterizes the equivalence between POP (1) and its MC relaxation based on Theorem 3.1 of Part I [5].

The second purpose of applying the unified framework to POPs is to establish a theoretical foundation for *effective*, *efficient* and *stable* numerical methods, which can be implemented with first-order algorithms, for solving large-scale POPs. The MC relaxation, *i.e.*, COP(co $\boldsymbol{\Gamma}$) in (2), can be regarded as the most effective relaxation in terms of the quality of the lower bound provided for the optimal value of POP (1). It is, however, numerically intractable even when $\mathbb{J}$ is a simple closed convex cone such as $\mathbb{R}^n$ or $\mathbb{R}^n_+$ (the nonnegative orthant of $\mathbb{R}^n$) in (1). As a numerically tractable relaxation of COP(co $\boldsymbol{\Gamma}$) in (2) for $\mathbb{J} = \mathbb{R}^n_+$, the doubly nonnegative (DNN) relaxation obtained by choosing a DNN cone for $\mathbb{K}$ in COP (2) and the Lagrangian-DNN relaxation of POP (1) can be considered. In the recent paper [16], the Lagrangian-DNN relaxation for a class of QOPs with complementarity constraints was implemented with first-order algorithms. It was shown through numerical results on some well-known test problems for nonconvex QOPs that the Lagrangian-DNN relaxation combined with first-order algorithms provided tight lower bounds for their optimal values efficiently and stably. See also Section 5 of Part I [5]. Those results indicate that applying the methods in [16] to POP (1) with $\mathbb{J} = \mathbb{R}^n_+$ is an important subject that needs further investigation.

For an alternative to the MC relaxation, a sparse variant [17, 20, 22, 28] of the hierarchy of SDP relaxations proposed by Lasserre [21] can be used for an equality and inequality constrained POP. The hierarchy is parameterized by a positive integer ω called the relaxation order. This relaxation method is numerically tractable and as effective as the MC relaxation, in the sense that under a certain assumption, the optimal value of the SDP relaxation with the order ω converges to the optimal value of the given POP as $\omega \to \infty$. This is a theoretically strong result, but solving the SDP relaxations of increasing size with ω is known to be numerically hard. In fact, SDP solvers [12, 26, 27] based on primal-dual interior-point methods have difficulties in solving SDP relaxation problems with the dimension larger than several thousands.

We propose a hierarchy of sparse Lagrangian-SDP relaxations for an equality constrained POP of the form (1) with $\mathbb{J} = \mathbb{R}^n$, by combining the hierarchy of sparse SDP relaxations and the Lagrangian-conic relaxation in the unified framework. Notice that inequality constraints can also be dealt with in the form (1) since any polynomial inequality $g(\boldsymbol{x}) \geq 0$ is equivalent to the equality $g(\boldsymbol{x}) - v^2 = 0$ with a slack variable $v \in \mathbb{R}$. Assuming that the polynomials $f^k(\boldsymbol{x})$ $(k = 0, 1, \ldots, m)$ are sparse, we first embed the structured sparsity characterized by a chordal graph as in [17, 20, 22, 28], and derive a sparse sum of squares (SOS) problem equivalent to POP (1) with $\mathbb{J} = \mathbb{R}^n$ by applying a sparse variant (Corollary 3.3 of [22], Theorem 1 of [20]) of Putinar's lemma (Lemma 4.1 of [25]) for representing positive polynomials by SOS polynomials. We then transform the SOS problem into a simpler SOS problem with a structure that leads to the copositivity condition in the unified framework. The transformed SOS problem is still not numerically solvable because SOS polynomials with any degree can be used in the problem. By restricting the degrees of SOS polynomials by the relaxation order ω, we construct a hierarchy of numerically solvable SOS problems from the transformed SOS problem. Finally, we convert the hierarchy of SOS problems into the hierarchy of SDPs in the linear space $\mathbb{V}_\omega$ of symmetric matrices with the dimension determined by the maximum degree of the polynomials $f^k(\boldsymbol{x})$ $(k = 0, 1, 2, \ldots, m)$ and ω.

Each SDP with the relaxation order ω in the hierarchy is of the form:

$$\zeta^d_\omega(\mathbb{K}_\omega) = \sup \left\{ y_0 \;\middle|\; \boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \sum_{k=1}^{m} \boldsymbol{Q}^k_\omega y_k \in \mathbb{K}^*_\omega \right\} \tag{3}$$

for some cone $\mathbb{K}_\omega \subset \mathbb{V}_\omega$, symmetric matrices $\boldsymbol{H}^0_\omega$, $\boldsymbol{Q}^k_\omega \in \mathbb{V}_\omega$ ($k = 0, 1, 2, \ldots, m$). This problem corresponds to the dual of (2). It is constructed so that it satisfies the copositivity condition for $\mathbb{K} = \mathbb{K}_\omega$. Thus, the entire theory of the unified framework can be applied. In particular, we can derive the following primal-dual pair of Lagrangian-SDP relaxation problems:

$$\zeta^p_\omega(\lambda, \mathbb{K}_\omega) = \inf \left\{ \langle \boldsymbol{Q}^0_\omega + \lambda \boldsymbol{H}^1_\omega, \boldsymbol{X} \rangle \;\middle|\; \boldsymbol{X} \in \mathbb{K}_\omega, \; \langle \boldsymbol{H}^0_\omega, \boldsymbol{X} \rangle = 1 \right\}, \tag{4}$$

$$\zeta^d_\omega(\lambda, \mathbb{K}_\omega) = \sup \left\{ y_0 \;\middle|\; y_0 \in \mathbb{R}, \; \boldsymbol{Q}^0_\omega - \boldsymbol{H}^0_\omega y_0 + \lambda \boldsymbol{H}^1_\omega \in \mathbb{K}^*_\omega \right\}, \tag{5}$$

where $\boldsymbol{H}^1_\omega = \sum_{k=1}^m \boldsymbol{Q}^k_\omega$, and $\lambda \in \mathbb{R}$ denotes a Lagrangian multiplier (or parameter) prescribed for the problems. These problems are very simple so that efficient first-order algorithms can be designed to solve the problems. In fact, the dual problem (5) involves only one variable, which makes it possible to effectively utilize the bisection and 1-dimensional Newton methods proposed in Part I [5]; see Also [16]. They also inherit the structured sparsity characterized by a chordal graph from the one embedded in POP (1) with $\mathbb{J} = \mathbb{R}^n$, for example, $\mathbb{K}^*$ is described as the Minkovski sum of positive semidefinite cones of small dimensions and linear subspaces of $\mathbb{V}_\omega$. Moreover, the following theoretical results will be established: for every relaxation order ω and Lagrange multiplier λ,

- the optimal value $\zeta^d_\omega(\mathbb{K}_\omega)$ of (3) bounds the optimal value ζ^* of POP (1) with $\mathbb{J} = \mathbb{R}^n$ from below, and it monotonically converges to ζ^* as $\omega \to \infty$,
- the optimal value $\zeta^d_\omega(\lambda, \mathbb{K}_\omega)$ of (5) bounds the optimal value $\zeta^d_\omega(\mathbb{K}_\omega)$ of (3) from below, and it monotonically converges to $\zeta^d_\omega(\mathbb{K}_\omega)$ as $\lambda \to \infty$,
- the primal problem (4) is strictly feasible (*i.e.*, there exists a primal feasible solution that lies in the relative interior of the cone $\mathbb{K}_\omega$) and it has an optimal solution with the optimal value $\zeta^d_\omega(\lambda, \mathbb{K}_\omega) = \zeta^p_\omega(\lambda, \mathbb{K}_\omega)$ (the strong duality).

The first two results imply that the lower bound $\zeta^d_\omega(\lambda, \mathbb{K}_\omega)$ for the optimal value ζ^* of POP (1) satisfies $\zeta^* - \epsilon < \zeta^d_\omega(\lambda, \mathbb{K}_\omega)$ for any $\epsilon > 0$ if sufficiently large ω and λ are taken. The last result contributes to the numerical stability of first-order algorithms.

In section 2, we review the results shown in Part I [5] and describe the notation and symbols used in this paper. We also present how to represent polynomials with symmetric matrices of monomials, and introduce SOS of polynomials. Section 3 includes the discussion on the MC relaxation of POP (1), and section 4 presents the hierarchy of sparse Lagrangian-SDP relaxations of POP (1) with $\mathbb{J} = \mathbb{R}^n$.

2 Preliminaries

2.1 Conic and Lagrangian-conic optimization problems

The results in Part I [5] 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 $\boldsymbol{H}^0,\ \boldsymbol{Q}^k \in \mathbb{V}\ (k = 0, , 2, \dots, m),\ \boldsymbol{H}^1 = \sum_{k=1}^m \boldsymbol{Q}^k$, let

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

In Part I, we introduced the following conditions:

Condition (I) $\boldsymbol{O} \neq \boldsymbol{H}^0 \in \mathbb{K}^*$, $\boldsymbol{Q}^k \in \mathbb{K}^*$ $(k = 1, 2, \dots, m)$ (copositivity condition).

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 pairs of problems,

$$
\begin{aligned}
\zeta^p(\mathbb{K}) &:= \inf \left\{ \langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle \mid \boldsymbol{X} \in F(\mathbb{K}) \right\} \\
&= \inf \left\{ \langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle \;\middle|\; \begin{array}{l} \boldsymbol{X} \in \mathbb{K},\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1, \\ \langle \boldsymbol{Q}^k, \boldsymbol{X} \rangle = 0\ (k = 1, 2, \dots, m) \end{array} \right\}. \qquad (6) \\
\zeta^d(\mathbb{K}) &:= \sup \left\{ z_0 \;\middle|\; \boldsymbol{Q}^0 + \sum_{k=1}^m \boldsymbol{Q}^k z_k - \boldsymbol{H}^0 z_0 \in \mathbb{K}^* \right\}, \qquad (7) \\
\eta^p(\lambda, \mathbb{K}) &:= \inf \left\{ \langle (\boldsymbol{Q}^0 + \lambda \boldsymbol{H}^1), \boldsymbol{X} \rangle \;\middle|\; \boldsymbol{X} \in \mathbb{K},\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1 \right\}, \qquad (8) \\
\eta^d(\lambda, \mathbb{K}) &:= \sup \left\{ y_0 \;\middle|\; \boldsymbol{Q}^0 + \lambda \boldsymbol{H}^1 - \boldsymbol{H}^0 y_0 \in \mathbb{K}^* \right\}, \qquad (9)
\end{aligned}
$$

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).*

(ii) $\left(\eta^d(\lambda, \mathbb{K}) = \eta^p(\lambda, \mathbb{K})\right) \uparrow \lambda = \zeta^d(\mathbb{K}) \leq \zeta^p(\mathbb{K})$ *under Conditions (I) and (II).*

(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).*

(iv) $\zeta^p(\mathbb{K}) = \zeta^p(\text{co } \mathbb{K})$ *under Conditions (I) and (IV).*

Here $\uparrow \lambda$ *means "increases monotonically as* $\lambda \to \infty$*".*

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 α_i denotes the ith 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 $\mathrm{supp}(f(\boldsymbol{x})) = \{\boldsymbol{\alpha} \in \mathcal{F} : f_{\boldsymbol{\alpha}} \neq 0\} \subset \mathbb{Z}^n_+$, and the degree of $f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ is defined by $\deg(f(\boldsymbol{x})) = \max\{|\boldsymbol{\alpha}| : \boldsymbol{\alpha} \in \mathrm{supp}(f(\boldsymbol{x}))\}$. 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 $\mathbb{R}[\boldsymbol{x}, \mathcal{F}] = \{f(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}] : \mathrm{supp}(f(\boldsymbol{x})) \subset \mathcal{F}\}$.

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. If $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}^{\mathcal{F}} = (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})$ 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 $\xi_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ $(\boldsymbol{\alpha} \in \mathcal{F},\ \boldsymbol{\beta} \in \mathcal{F})$. 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 $(\xi_{\boldsymbol{\alpha}\boldsymbol{\beta}} : (\alpha, \beta) \in \square\mathcal{F})$ or simply $(\xi_{\boldsymbol{\alpha}\boldsymbol{\beta}} : \square\mathcal{F})$. If $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}^{\mathcal{F}}(\boldsymbol{x}^{\mathcal{F}})^T = (\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T = (\boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} : (\alpha, \beta) \in \square\mathcal{F})$ 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 $\boldsymbol{x}^{\square\mathcal{F}}$. $(\boldsymbol{x}^{\mathcal{F}})^T$ denotes the row vector obtained by taking the transpose of the column vector $\boldsymbol{x}^{\mathcal{F}}$.

For every pair of $\boldsymbol{Q} = (Q_{\boldsymbol{\alpha}\boldsymbol{\beta}} : \square\mathcal{F})$ and $\boldsymbol{X} = (X_{\boldsymbol{\alpha}\boldsymbol{\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}}$, $\langle \boldsymbol{Q},\ \boldsymbol{X} \rangle$ denotes the matrix inner product, *i.e.*, $\langle \boldsymbol{Q},\ \boldsymbol{X} \rangle = \mathrm{trace}(\boldsymbol{Q}^T \boldsymbol{X}) = \sum_{(\boldsymbol{\alpha},\boldsymbol{\beta})\in\square\mathcal{F}} Q_{\boldsymbol{\alpha}\boldsymbol{\beta}} \xi_{\boldsymbol{\alpha}\boldsymbol{\beta}}$. 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 $\boldsymbol{X} = (X_{\boldsymbol{\alpha}\boldsymbol{\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}}$.

Let

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

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 nonnegative polynomials on $\mathbb{R}^n$ and $\mathbb{R}^n_+$, and exploiting sparsity in 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 [4].

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

$$f^1(\boldsymbol{x}) \;=\; 1 - 2x_1 - 2x_2 + x_1^2 + x_2^2 + 2x_1^2x_2 + 2x_1x_2^2 + x_1^2x_2^2.$$

Let

$$\begin{aligned} \mathcal{G} &= \{(0,0),(1,0),(0,1),(2,0),(0,2),(2,1),(1,2),(2,2)\}, \\ (f^1_{\boldsymbol{\alpha}} : \mathcal{G}) &= (1,-2,-2,1,1,2,2,1), \\ \boldsymbol{x}^{\mathcal{G}} &= (x^{\boldsymbol{\alpha}} : \mathcal{G}) = (1, x_1, x_2, x_1^2, x_2^2, x_1^2x_2, x_2^2.x_1^2x_2^2). \end{aligned}$$

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 2.1. *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,*

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

Proof. We first show that the linear subspace of $\mathbb{S}^{\mathcal{F}}$ generated by $\left\{\boldsymbol{x}^{\square\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^n\right\}$, *i.e.*,

$$\overline{\mathbb{L}} = \left\{\lambda\boldsymbol{x}^{\square\mathcal{F}} + \mu\boldsymbol{y}^{\square\mathcal{F}} : \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^n,\ \lambda,\ \mu \in \mathbb{R}\right\},$$

coincides with $\mathbb{L}^{\mathcal{F}}$. By the definition of $\mathbb{L}^{\mathcal{F}}$, we know that $\left\{\boldsymbol{x}^{\square\mathcal{F}} : \boldsymbol{x} \in \mathbb{R}^n\right\} \subset \mathbb{L}^{\mathcal{F}}$, 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 different 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 2.1 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 an SOS polynomials.

Lemma 2.2. *[10] 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 2.1, 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

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{F}] \ni f^1(\boldsymbol{x}) = \langle \boldsymbol{Q}^1,\, \boldsymbol{x}^{\square\mathcal{F}} \rangle = (x_1 + x_2 + x_1 x_2 - 1)^2, \tag{10}$$

where

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

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

Lemma 2.3. *Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$ and $\mathbb{M}$ a linear subspace of $\mathbb{S}^{\mathcal{F}}$. Then, $(\mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{M}) + (\left(\mathbb{L}^{\mathcal{F}}\right)^{\perp} \cap \mathbb{M})$ is closed.*

Proof. We first show that 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 $\mathbb{S}^{\mathcal{F}}$. 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. For some bounded 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 $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{F}}_{+} \cap \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$. Recall that $\boldsymbol{x}_j^{\mathcal{F}} = (\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})^T \in \mathbb{L}^{\mathcal{F}}$. Hence $\langle \boldsymbol{A}\boldsymbol{A}^T, \overline{\boldsymbol{X}}\rangle = \langle \sum_{j=1}^{q}(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})^T, \overline{\boldsymbol{X}}\rangle = \sum_{j=1}^{q}\langle(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})(\boldsymbol{x}_j^{\boldsymbol{\alpha}} : \mathcal{F})^T, \overline{\boldsymbol{X}}\rangle = 0$. 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 $\overline{\boldsymbol{X}} = \boldsymbol{O}$. This is a contradiction. Thus we have shown that $\{\boldsymbol{X} \in \mathbb{S}^{\mathcal{F}} : \boldsymbol{X}+\boldsymbol{Y} \in B,\ \boldsymbol{X} \in \mathbb{S}^{\mathcal{F}}_{+},\ \boldsymbol{Y} \in \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}\}$ is bounded for every bounded subset B of $\mathbb{S}^{\mathcal{F}}$.

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

3 A class of polynomial optimization problems and their covexification

Consider a class of POPs of the form (1). Assume that $\mathbb{J}$ is a nonempty closed (but not necessarily convex) cone in $\mathbb{R}^n$, and $f^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{F}+\mathcal{F}]$ ($k = 0,1,\ldots,m$) 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 transform POP (1) into the COP of the form (6) to present the moment cone (MC) relaxation of POPs in the unified framework of Section 2, Part I [5]. The MC relaxation was proposed by [4] as an extension of the completely positive (CPP) relaxation [2, 8] for a class of linearly constrained QOPs in continuous and binary variables. We also recall that the CPP relaxation for the class of QOPs was discussed in Section 4 of [16].

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 $f^k(\boldsymbol{x}) = \langle \boldsymbol{Q}^k, \boldsymbol{x}^{\square\mathcal{F}}\rangle$ ($k = 0,1,\ldots,m$), 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) can be rewritten as

$$\zeta^* \ := \ \inf \left\{ \langle \boldsymbol{Q}^0, \boldsymbol{X}\rangle \ \middle|\ \boldsymbol{X} \in \boldsymbol{\Delta}^{\mathcal{F}}_1,\ \langle \boldsymbol{Q}^k, \boldsymbol{X}\rangle = 0 \ (k = 1,2,\ldots,m) \ \right\}. \tag{11}$$

We consider the following illustrative example:

Example 3.1. As in Example 2.1, 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 (10). Let $\mathbb{J} = \mathbb{R}^2_+$. In this case, it is obvious that the feasible region of POP (14) is bounded and contains two points $\boldsymbol{x} = (1,0)$ and $\boldsymbol{x} = (0,1)$, thus, its finite minimum value ζ^* 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 (11) is in a form similar to COP (6), 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

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

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

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

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

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

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 [24], while the second $\boldsymbol{\Gamma}^{\mathcal{F}}$ in [4].

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

$$\begin{aligned} &\left\{\boldsymbol{X} \in \boldsymbol{\Delta}^{\mathcal{F}} : \langle \boldsymbol{H}^0,\, \boldsymbol{X}\rangle = 1\right\} \\ &= \left\{\mu\boldsymbol{x}^{\square\mathcal{F}} \in \mathbb{S}^{\mathcal{F}} : \boldsymbol{x} \in \mathbb{J},\ \mu \geq 0,\ \langle \boldsymbol{H}^0,\, \mu\boldsymbol{x}^{\square\mathcal{F}}\rangle = 1\right\} = \boldsymbol{\Delta}^{\mathcal{F}}_1, \\ &\left\{\boldsymbol{X} \in \boldsymbol{\Gamma}^{\mathcal{F}} : \langle \boldsymbol{H}^0,\, \boldsymbol{X}\rangle = 1\right\} \\ &= \left\{(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T : (x_0,\boldsymbol{x}) \in \mathbb{R}_+ \times \mathbb{J},\ x_0 = 1\right\} = \boldsymbol{\Delta}^{\mathcal{F}}_1. \end{aligned} \tag{12}$$

Therefore, both COP (6) with $\mathbb{K} = \boldsymbol{\Delta}^{\mathcal{F}}$ and COP (6) with $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$ are equivalent to POP (11), and $\zeta^p(\boldsymbol{\Delta}^{\mathcal{F}}) = \zeta^p(\boldsymbol{\Gamma}^{\mathcal{F}}) = \zeta^*$.

For Example 3.1, we see that

$$\begin{aligned} \boldsymbol{\Delta}^{\mathcal{F}} &= \left\{\begin{pmatrix} \mu & \mu x_1 & \mu x_2 & \mu x_1x_2 \\ \mu x_1 & \mu x_1^2 & \mu x_1x_2 & \mu x_1^2x_2 \\ \mu x_2 & \mu x_1x_2 & \mu x_2^2 & \mu x_1x_2^2 \\ \mu x_1x_2 & \mu x_1^2x_2 & \mu x_1x_2^2 & \mu x_1^2x_2^2 \end{pmatrix} \in \mathbb{S}^{\mathcal{F}}_+ :\ (\mu,x_1,x_2) \in \mathbb{R}^3_+ \right\}, \\ \boldsymbol{\Gamma}^{\mathcal{F}} &= \left\{\begin{pmatrix} x_0^4 & x_0^3x_1 & x_0^3x_2 & x_0^2x_1x_2 \\ x_0^3x_1 & x_0^2x_1^2 & x_0^2x_1x_2 & x_0x_1^2x_2 \\ x_0^3x_2 & x_0^2x_1x_2 & x_0^2x_2^2 & x_0x_1x_2^2 \\ x_0^2x_1x_2 & x_0x_1^2x_2 & x_0x_1x_2^2 & x_1^2x_2^2 \end{pmatrix} \in \mathbb{S}^{\mathcal{F}}_+ :\ (x_0,x_1,x_2) \in \mathbb{R}^3_+ \right\}. \end{aligned} \tag{13}$$

Obviously, for every $x_1 > 0$ and every $x_2 > 0$, the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1^2 x_2^2 \end{pmatrix}$$

is contained in $\mathbf{\Gamma}^{\mathcal{F}}$ but not in $\mathbf{\Delta}^{\mathcal{F}}$.

Now, we are ready to apply all the discussions in Section 3 of Part I [5]. In particular, the following theorem is obtained directly from Theorem 3.1 of [5].

Theorem 3.1. *Suppose that Condition (I) holds for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$. Then,*

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

(ii) Assume that ζ^* is finite, we have that

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

Next we investigate Conditions (I) and (IV) in detail for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$, and in particular, $\mathbb{J} = \mathbb{R}^n$ and $\mathbb{J} = \mathbb{R}^n_+$. First, suppose that $\mathbb{J} = \mathbb{R}^n$. Then it follows that $\mathbf{\Gamma}^{\mathcal{F}} \subset \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{L}^{\mathcal{F}}$, and $\left(\mathbf{\Gamma}^{\mathcal{F}}\right)^* \supset \text{cl}\left(\mathbb{S}^{\mathcal{F}}_+ + \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}\right) = \mathbb{S}^{\mathcal{F}}_+ + \left(\mathbb{L}^{\mathcal{F}}\right)^{\perp}$. (See Lemma 2.3 for the last equality). 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 a sum of squares of polynomials ($k = 1, 2, \ldots, m$) (see Lemma 2.1 and 2.2), then Condition (I) is satisfied for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$. (Recall that $\boldsymbol{O} \neq \boldsymbol{H}^0 \in \mathbb{S}^{\mathcal{F}}_+$ by definition). If a polynomial equation $g(\boldsymbol{x}) = 0$ is given, it is equivalent to have $(g(\boldsymbol{x}))^2 = 0$. Thus, Condition (I) for $\mathbb{K} = \mathbf{\Gamma}^{\mathcal{F}}$ is not a strong assumption.

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

$$\mathbf{\Gamma}^{\mathcal{F}} \subset (\mathbb{C}^{\mathcal{F}})^* \cap \mathbb{L}^{\mathcal{F}} \subset \mathbb{S}^{\mathcal{F}}_+ \cap \mathbb{N}^{\mathcal{F}} \cap \mathbb{L}^{\mathcal{F}},$$

and

$$\left(\mathbf{\Gamma}^{\mathcal{F}}\right)^* \supset \text{cl}\ \left(\mathbb{C}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^{\perp}\right) \supset \text{cl}\ \left(\mathbb{S}^{\mathcal{F}}_+ + \mathbb{N}^{\mathcal{F}} + (\mathbb{L}^{\mathcal{F}})^{\perp}\right),$$

where

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

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 $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$.

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 $\overline{\mathcal{F}} = \{\boldsymbol{\alpha} \in \mathcal{F} : |\boldsymbol{\alpha}| = \tau\}$. Suppose that

$$\boldsymbol{X} \;=\; (x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \mathcal{F})^T \in F_0(\boldsymbol{\Gamma}^{\mathcal{F}})$$

Then,

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

As a result, if we define $\overline{\boldsymbol{Q}}^k = (\overline{Q}^k_{\boldsymbol{\alpha}\boldsymbol{\beta}} : \square\mathcal{F}) \in \mathbb{S}^{\mathcal{F}}$ and $\bar{f}^k(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \overline{\mathcal{F}} + \overline{\mathcal{F}}]$ such that

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

($k = 0, 1, \ldots, m$), we can rewrite Condition (IV) for $\mathbb{K} = \boldsymbol{\Gamma}^{\mathcal{F}}$ as

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

Usual cases where the condition above holds are:

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

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 $\deg(f^0(\boldsymbol{x})) < 2\tau$, and that (b) implies that the feasible region of POP (1) is bounded.

For Example 3.1, we observe that

$$F_0(\boldsymbol{\Gamma}^{\mathcal{F}}) = \left\{ \boldsymbol{X} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1^2x_2^2 \end{pmatrix} \in \mathbb{S}^{\mathcal{F}}_+ : \begin{array}{l} (x_1, x_2) \in \mathbb{R}^2_+, \\ \langle \boldsymbol{Q}^1,\, \boldsymbol{X} \rangle = Q^1_{(1,1)(1,1)} X_{(1,1)(1,1)} = 0 \end{array} \right\} = \{\boldsymbol{O}\},$$

where $\boldsymbol{Q}^1$ is given as in (10). 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 (8) and (9) for $\mathbb{K} = \text{co}\,\boldsymbol{\Gamma}$, and apply the discussions given in Section 3 of Part I [4]. In particular, 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; see Theorem 2.1. 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 (13) 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 $(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{F})(x_0^{\tau-|\boldsymbol{\alpha}|}\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{F})^T \in \boldsymbol{\Gamma}$ with $(x_0, x_1, x_2) = (\epsilon, \epsilon, 1/\epsilon)$ and $\epsilon > 0$ converges to $\overline{\boldsymbol{X}}$.

Lemma 3.1. *Assume that* $\mathbf{0} \in \mathcal{F}$ *and* $\tau \boldsymbol{e}_i \in \mathcal{F}$ $(i = 1, 2, \ldots, n)$*, where* $\boldsymbol{e}_i$ *denotes the i-th unit coordinate vector in* $\mathbb{R}^n$*. Then,* $\mathbf{\Gamma}^{\mathcal{F}}$ *and co* $\mathbf{\Gamma}^{\mathcal{F}}$ *are closed.*

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

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

and we have shown that $\mathbf{\Gamma}^{\mathcal{F}}$ is closed. The closedness of co $\mathbf{\Gamma}^{\mathcal{F}}$ follows from Lemma 3.1 of [4]. □

Without loss of generality, the assumption of the previous lemma can be satisfied by adding $\mathbf{0} \in \mathbb{Z}^n_+$ and $\tau\boldsymbol{e}_i \in \mathbb{Z}^n_+$ $(i = 1, 2, \ldots, n)$ 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 [4] as an extension of the CPP relaxation for QOPs to POPs is essentially equivalent to COP (6) with $\mathbb{K} = \text{co}\,\mathbf{\Gamma}^{\mathcal{F}}$. 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 (8) in the unified framework. On the other hand, the additional condition on zeros at infinity assumed in [4], which was also assumed in [24] for a canonical convexification procedure for a class of POPs, is stronger than Condition (IV). See Condition (IV)' and Lemma 3.1 of [5].

If $\mathcal{F} = \{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : |\boldsymbol{\alpha}| \leq 1\}$, then POP (1) becomes a QOP. In this case, $\mathbb{L}^{\mathcal{F}} = \mathbb{S}^{\mathcal{F}}$ and $(\mathbb{L}^{\mathcal{F}})^{\perp} = \{\mathbf{0}\}$, and the previous discussions correspond to Section 4.2 of Part I [5], where the convexification of a linearly constrained QOP with complementarity condition was discussed.

4 A hierarchy of sparse Lagrangian-SDP relaxations for POPs

The primal COP (6) and its Lagrangian-conic relaxation (8) have been mainly used when applying the unified framework of Part I [5] to QOPs and POPs, and their duals. As a result of this application, the problem of computing tight lower bounds for the optimal values of general POPs is reduced to the dual COP (7).

In this section, we propose a hierarchy of Lagrangian-SDP relaxations for POPs by combining the approach in [3, 5, 16] for deriving the Lagrangian-CPP and Lagrangian-DNN relaxations for a class of QOPs with the hierarchy of SDP relaxations proposed by [21] 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. We simultaneously take account the important issue of exploiting sparsity [17, 22, 28] in the proposed hierarchy of Largangian-SDP relaxations for POPs. As a result, the description may be slightly complicated, but exploiting sparsity is essential for numerical efficiency of solving large-scale POPs.

4.1 A class of equality constrained POPs with a structured sparsity

Let us fix $\mathbb{J} = \mathbb{R}^n$ in POP (1) throughout this section. We deal with a class of equality constrained POP of the form:

$$\zeta^* \;=\; \inf\; \left\{ f^0(\boldsymbol{x}) \;\middle|\; \boldsymbol{x} \in \mathbb{R}^n,\; f^i(\boldsymbol{x}) = 0 \; (i = 1, 2, \ldots, m_0) \right\}. \tag{14}$$

As mentioned in Section 1, we can include inequality constraints since a polynomial inequality $g(\boldsymbol{x}) \geq 0$ can be rewritten as $g(\boldsymbol{x}) - v^2 = 0$ with a slack variable $v \in \mathbb{R}$. We assume that POP (14) is sparse. More precisely, each equality constraint $f^k(\boldsymbol{x}) = 0$ involves only a small subset of the variables $x_1, x_2, \ldots, x_n$ and the objective polynomial $f^0(\boldsymbol{x})$ consists of monomials whose variables include only a small number of pairs of x_i and x_j. Under this assumption, a structured sparsity [17, 22, 28] can be embed in POP (14). Let

$$\begin{aligned}
C^k &\subset \{1, \ldots, n\},\; C_k \neq \emptyset,\; (k = 1, \ldots, m),\; \bigcup_{k=1}^{m} C^k = \{1, \ldots, n\}, \\
\mathcal{A}_\tau &= \left\{ \boldsymbol{\alpha} \in \mathbb{Z}^n_+ : |\boldsymbol{\alpha}| \leq \tau \right\} \; (\tau \in \mathbb{Z}_+), \\
\mathcal{A}^k_\tau &= \left\{ \boldsymbol{\alpha} \in \mathcal{A}_\tau : \alpha_i = 0 \text{ if } i \notin C^k \right\} \; (\tau \in \mathbb{Z}_+, k = 1, \ldots, m), \\
\mathcal{A}^k_\infty &= \bigcup_{\tau \in \mathbb{Z}_+} \mathcal{A}^k_\tau = \left\{ \boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \alpha_i = 0 \text{ if } i \notin C_k \right\}.
\end{aligned}$$

Then, the sparsity structure of POP (14) is described as

$$f^0(\boldsymbol{x}) \in \sum_{k=1}^{m} \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_\infty] \;\text{ and }\; f^i(\boldsymbol{x}) \in \bigcup_{k=1}^{m} \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_\infty] \; (i = 1, 2, \ldots, m_0).$$

From the second inclusion relation, we can partition the index set $\{1, 2, \ldots, m_0\}$ of equality constraints in (14) into m sets I^k $(k = 1, 2, \ldots, m)$ such that

$$f^i(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_\infty] \text{ for every } i \in I^k.$$

Now, we rewrite POP (14) as

$$\zeta^* \;=\; \inf\; \left\{ f^0(\boldsymbol{x}) \;\middle|\; \boldsymbol{x} \in \mathbb{R}^n,\; f^i(\boldsymbol{x}) = 0 \; (i \in I^k,\; k = 1, 2, \ldots, m) \right\}. \tag{15}$$

As an illustrative example, we consider the same problem as Example 3.1.

Example 4.1. Let

$$\begin{aligned}
n &= 4,\; m_0 = 3, \\
\mathcal{F} &= \{\boldsymbol{0}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2, 2\boldsymbol{e}_3, 2\boldsymbol{e}_4\} \subset \mathbb{Z}^4_+, \\
& \quad \text{where } \boldsymbol{e}_i \in \mathbb{Z}^4_+ \text{ denotes the } i\text{th unit coordinate vector}, \\
f^0(\boldsymbol{x}) &\in \mathbb{R}[\boldsymbol{x}, \mathcal{F}^0 + \mathcal{F}^0], \text{ where } \mathcal{F}_0 = \{\boldsymbol{0}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2\}, \\
f^1(\boldsymbol{x}) &= x_1 - x_3^2, \quad f^2(\boldsymbol{x}) = x_2 - x_4^2, \quad f^3(\boldsymbol{x}) = x_1 + x_2 + x_1 x_2 - 1.
\end{aligned}$$

Then, the POP in Example 3.1 can be rewritten as

$$\text{minimize} \quad \left\{ f^0(\boldsymbol{x}) \;\middle|\; \begin{array}{l} f^1(\boldsymbol{x}) = x_1 - x_3^2 = 0, \quad f^2(\boldsymbol{x}) = x_2 - x_4^2 = 0, \\ f^3(\boldsymbol{x}) = x_1 + x_2 + x_1 x_2 - 1 = 0 \end{array} \right\}.$$

Define $I^k = \{k\}$ $(k = 1, 2, 3)$, $C^1 = \{1, 3\}$, $C^2 = \{2, 4\}$ and $C^3 = \{1, 2\}$. Then, the problem above can be expressed as in (15).

We impose the following two conditions on POP (15) in the subsequent discussions.

Condition (A)

(a) Each C^k $(k = 1, 2, \ldots, m)$ is a nonempty maximal subset among the family $\{C^1, C^2, \ldots, C^m\}$, *i.e.*, there is no C^s $(s \neq k)$ that includes C^k.

(b) For every $k \in \{1, \ldots, m-1\}$, there is an $s \geq k+1$ such that

$$C^k \cap \left(C^{k+1} \cup \cdots \cup C^m\right) \subset C^s \text{ and } \neq C^s.$$

Condition (B) For each $k = 1, 2, \ldots, m$, there is an $i \in I_k$ such that $f^i(\boldsymbol{x}) = \rho_k - \sum_{i \in C^k} x_i^2$ for some $\rho_k > 0$.

Let $G(N, E)$ be an undirected graph with the node set $N = \{1, 2, \ldots, n\}$ and the edge set

$$E = \{(i, j) \in N \times N : i \neq j \text{ and } (i, j) \in C^k \text{ for some } k = 1, 2, \ldots, m\},$$

where each edge $(i, j) \in E$ is identified with $(j, i) \in E$. It is known that Condition (A) provides a necessary and sufficient condition for $G(N, E)$ to be a chordal graph and for $C^1, C^2, \ldots, C^m$ to be its maximal cliques. The property (b) of Condition (A) is called the running intersection property. See [7] for the definition of a chordal graph and its properties.

Condition (B) implies that the feasible region of POP (15) is bounded. Conversely, if the feasible region is bounded, we can add the constraints

$$\rho_k - \sum_{i \in C^k} x_i^2 - x_{n+k}^2 = 0 \ (k = 1, 2, \ldots, m)$$

for some $\rho_k > 0$, and replace $\boldsymbol{x}$ by $(\boldsymbol{x}, x_{n+1}, \ldots, x_{n+m})$ and C^k by $C^k \cup \{n + k\}$ $(k = 1, 2, \ldots, m)$, where x_{n+k} denotes a slack variable $(k = 1, 2, \ldots, m)$. See [17, 22, 28] for more details on how the structured sparsity can be constructed from a given POP. Condition (B) can be weakened if Theorem 1 of [20] is used instead of Corollary 3.3 of [22] in the proof of Lemma 4.1. But its description is slightly complicated, so we prefer to use Condition (B).

We can easily verify that C^1, C^2 and C^3 constructed in Example 4.1 satisfy Condition (A). In this case, the corresponding graph $G(N, E)$ does not have any cycle, which is a trivial chordal graph. Although Example 4.1 can be modified to satisfy Condition (B), we use Example 4.1 for ease of discussion.

4.2 A sparse SOS problem equivalent to POP (15)

By applying the sparse SOS relaxation [17, 22, 28] to POP (15), we obtain the following SOS problem:

$$\bar{\zeta}^d_\infty \;:=\; \sup\left\{ z_0 \in \mathbb{R} \;\middle|\; \begin{array}{l} f^0(\boldsymbol{x}) - z_0 \\ \in \displaystyle\sum_{k=1}^m \left(\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty] + \sum_{i \in I_k} \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_\infty] f^i(\boldsymbol{x}) \right) \end{array} \right\}. \tag{16}$$

Lemma 4.1. *Suppose that the feasible region of POP* (15) *is nonempty and that Conditions (A) and (B) hold. Then* $\zeta^* = \bar{\zeta}^d_\infty$.

Proof. By Condition (B), the feasible region of POP (15) is bounded, which together with the nonemptyness of the feasible region ensures that the optimal value ζ^* is finite. Let z_0 be a feasible solution of SOS problem (16). Then there exist $\varphi^k(\boldsymbol{x}) \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty]$ and $\psi^i(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_\infty]$ $(i \in I_k,\ k = 1, 2, \ldots, m)$ such that

$$f^0(\boldsymbol{x}) - z_0 \;=\; \sum_{k=1}^m \left(\varphi^k(\boldsymbol{x}) + \sum_{i \in I_k} \psi^i(\boldsymbol{x}) f^i(\boldsymbol{x}) \right) \text{ for all } \boldsymbol{x} \in \mathbb{R}^n,$$

which implies that $f^0(\boldsymbol{x}) - z_0 \geq 0$ for every feasible solution $\boldsymbol{x}$ of POP (15). Thus we have shown that $\zeta^* \geq \bar{\zeta}^d_\infty$. To prove the converse inequality, let

$$\begin{aligned} S &= \left\{ \boldsymbol{x} \in \mathbb{R}^n : f^i(\boldsymbol{x}) = 0 \ (i \in I_k,\ k = 1, \ldots, m) \right\} \\ &= \left\{ \boldsymbol{x} \in \mathbb{R}^n : \begin{array}{l} f^i(\boldsymbol{x}) \geq 0 \ (i \in I_k,\ k = 1, \ldots, m), \\ -f^i(\boldsymbol{x}) \geq 0 \ (i \in I_k,\ k = 1, \ldots, m) \end{array} \right\}, \end{aligned}$$

and let ϵ be an arbitrary positive number. Then $f_0(\boldsymbol{x}) - (\zeta^* - \epsilon) > 0$ for all $\boldsymbol{x} \in S$. By Corollary 3.3 of [22],

$$\begin{aligned} & f_0(\boldsymbol{x}) - (\zeta^* - \epsilon) \\ \in & \sum_{k=1}^m \left(\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty] + \sum_{i \in I_k} \left(\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty] f^i(\boldsymbol{x}) - \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty] f^i(\boldsymbol{x}) \right) \right) \\ = & \sum_{k=1}^m \left(\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty] + \sum_{i \in I_k} \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_\infty] f^i(\boldsymbol{x}) \right). \end{aligned}$$

We have shown that for any positive ϵ, $z_0 = \zeta^* - \epsilon$ is a feasible solution of (16). Thus $\bar{\zeta}^d_\infty \geq \zeta^* - \epsilon$ for all $\epsilon > 0$. Therefore, $\zeta^* \leq \bar{\zeta}^d_\infty$. □

4.3 Replacing SOS problem (16) by a simpler SOS

In this subsection, we establish the equivalence between SOS problem (16) and the following SOS problem, which is simpler than (16).

$$\zeta^d_\infty \;:=\; \sup\left\{ y_0 \in \mathbb{R} \;\middle|\; \begin{array}{l} f^0(\boldsymbol{x}) - y_0 \\ \in \displaystyle\sum_{k=1}^m \left(\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\infty] - y_k \sum_{i \in I_k} \boldsymbol{\Theta}^k[\boldsymbol{x}] (f^i(\boldsymbol{x}))^2 \right) \\ (y_0, y_1, \ldots, y_m) \in \mathbb{R}^{1+m} \end{array} \right\}, \tag{17}$$

where

$$\begin{aligned}\theta^k_\tau(\boldsymbol{x}) &= \sum_{\boldsymbol{\alpha}\in\mathcal{A}^k_\tau} \boldsymbol{x}^{2\boldsymbol{\alpha}} \in \mathbb{SOS}[\boldsymbol{x},\mathcal{A}^k_\tau], \\ \boldsymbol{\Theta}^k[\boldsymbol{x}] &= \{\theta^k_\tau(\boldsymbol{x}) : \tau \in \mathbb{Z}_+\} \subset \mathbb{SOS}[\boldsymbol{x},\mathcal{A}^k_\infty].\end{aligned} \tag{18}$$

The second relation in (18) indicates that (17) is simpler than (16). The following lemma shows the equivalence and plays a key role in analyzing the SOS problem (16) in the unified framework presented in Part I [5].

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

Proof. (i) Proof of $\zeta^d_\infty \leq \bar{\zeta}^d_\infty$. We know that

$$\mathbb{R}[\boldsymbol{x},\mathcal{A}^k_\infty] \supset -y_k\boldsymbol{\Theta}^k[\boldsymbol{x}]f^i(\boldsymbol{x}) \text{ for every } y_k \in \mathbb{R}\ (i \in I_k,\ k = 1,\ldots,m).$$

This implies that if $(y_0.y_1,\ldots,y_m)$ is a feasible solution of (17), then $z_0 = y_0$ is a feasible solution of (16). 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 (16) and ϵ an arbitrary positive number. We show that there is a feasible solution $(y_0,y_1,\ldots,y_m)$ of (17) with objective value $y_0 = z_0 - \epsilon$. Since z_0 is feasible solution of (16), we see that

$$\sum_{k=1}^m \mathbb{SOS}[\boldsymbol{x},\mathcal{A}^k_\infty] \ \ni\ f^0(\boldsymbol{x}) - z_0 - \sum_{k=1}^m\sum_{i\in I_k}\psi^i(\boldsymbol{x})f^i(\boldsymbol{x}) \tag{19}$$

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

$$\psi^i(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}^k_\tau}\psi^i_{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\alpha}}\ (i \in I_k,\ k = 1,\ldots,m).$$

Substituting these identities into the relation (19), we get

$$\sum_{k=1}^m \mathbb{SOS}[\boldsymbol{x},\mathcal{A}^k_\infty] \ \ni\ f^0(\boldsymbol{x}) - z_0 - \sum_{k=1}^m\sum_{i\in I_k}\sum_{\boldsymbol{\alpha}\in\mathcal{A}^k_\tau}\psi^i_{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\alpha}}f^i(\boldsymbol{x}). \tag{20}$$

Choose a $\rho > 0$ such that $\epsilon - \sum_{k=1}^m\sum_{i\in I_k}\sum_{\boldsymbol{\alpha}\in\mathcal{A}^k_\tau}(1/(2\rho))^2 > 0$, and let

$$z_k = \max\{(\rho\psi^i_{\boldsymbol{\alpha}})^2 : \boldsymbol{\alpha} \in \mathcal{A}^k_\tau,\ i \in I_k\}\ (k = 1,\ldots,m).$$

Then,

$$\begin{aligned}
\sum_{k=1}^{m} \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}_\infty^k] \ \ni \ & \left(\epsilon - \sum_{k=1}^{m} \sum_{i \in I_k} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_\tau^k} (1/(2\rho))^2 \right) \\
& + \sum_{k=1}^{m} \sum_{i \in I_k} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_\tau^k} \left(\rho \psi_{\boldsymbol{\alpha}}^i \boldsymbol{x}^{\boldsymbol{\alpha}} f^i(\boldsymbol{x}) + 1/(2\rho) \right)^2 \\
& + \sum_{k=1}^{m} \sum_{i \in I_k} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_\tau^k} \left(z_k - (\rho \psi_{\boldsymbol{\alpha}}^i)^2 \right) (\boldsymbol{x}^{\boldsymbol{\alpha}} f^i(\boldsymbol{x}))^2 \\
= \ & \epsilon + \sum_{k=1}^{m} \sum_{i \in I_k} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_\tau^k} \psi_{\boldsymbol{\alpha}}^i \boldsymbol{x}^{\boldsymbol{\alpha}} f^i(\boldsymbol{x}) \ + \ \sum_{k=1}^{m} \sum_{i \in I_k} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_\tau^k} z_k \left(\boldsymbol{x}^{\boldsymbol{\alpha}} f^i(\boldsymbol{x}) \right)^2 .
\end{aligned}$$

It follows from (20) that

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

Therefore, we have shown that

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

and that $(y_0, y_1, \dots, y_m) = (z_0 - \epsilon, z_1, \dots, z_m)$ is a feasible solution of SOS problem (17). Thus we have $\zeta_\infty^d \geq y_0 = z_0 - \epsilon$ for all $\epsilon > 0$. This implies that $\zeta_\infty^d \geq z_0$ for any feasible z_0 of (16). Hence $\zeta_\infty^d \geq \bar{\zeta}_\infty^d$. □

4.4 A hierarchy of finite SOS subproblems of (17) for numerical computation

The SOS problem (17) that attains the exact optimal value ζ^* of POP (15) cannot be solved numerically because the degree of sum of squares of polynomials involved is not bounded.

For numerical computation of lower bounds of ζ^* which converges to ζ^*, we introduce a hierarchy of SOS subproblems of (17) by bounding the degree of the SOS polynomials to be used with an increasing sequence of finite integers.

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

$$\zeta^d_\omega \ := \ \sup \left\{ y_0 \in \mathbb{R} \ \middle| \ \begin{array}{l} f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \sum_{i \in I_k} \theta^k_{\tau^i}(\boldsymbol{x}) (f^i(\boldsymbol{x}))^2 \\ \in \sum_{k=1}^m \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega], \\ (y_0, y_1, \dots, y_m) \in \mathbb{R}^{1+m} \end{array} \right\}, \tag{21}$$

where

$$\tau^i \ = \ \omega - \deg(f^i(\boldsymbol{x}))\ (i \in I_k,\ k = 1, \dots, m). \tag{22}$$

We note that

$$\begin{array}{l} f^0(\boldsymbol{x}) \in \sum_{k=1}^m \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_{\omega_{\min}} + \mathcal{A}^k_{\omega_{\min}}] \subset \mathbb{R}[\boldsymbol{x}, \mathcal{A}_\omega + \mathcal{A}_\omega], \\ \sum_{i \in I_k} \theta^k_{\tau^i}(\boldsymbol{x}) (f^i(\boldsymbol{x}))^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega] \subset \mathbb{R}[\boldsymbol{x}, \mathcal{A}_\omega + \mathcal{A}_\omega]\ (k = 1, 2, \dots, m). \end{array} \tag{23}$$

Therefore, the degree of polynomials in the SOS problem (21) is bounded by 2ω. This SOS problem can be solved as an SDP (31), as shown in the next subsection.

Now, we apply the discussion above to Example 4.1. First, observe that

$$\deg(f^i(\boldsymbol{x})) = 2\ (i = 1, 2, 3),\ \omega_{\min} = 2.$$

If we take $\omega = \omega_{\min} = 2$, then

$$\begin{array}{rcl} \tau^i & = & \omega - \deg(f^i(\boldsymbol{x})) = 0, \\ A^k_{\tau^i} & = & \{\boldsymbol{0}\} \subset \mathbb{Z}^4_+,\ \theta^k_{\tau^i} = 1\ (i = I_k = \{i\},\ k = 1, 2, 3). \end{array}$$

If $\omega = \omega_{\min} + 1 = 3$, then

$$\begin{array}{rcl} \tau^i & = & \omega - \deg(f^i(\boldsymbol{x})) = 1\ (i = 1, 2, 3), \\ A^1_{\tau^1} & = & \{\boldsymbol{0}, \boldsymbol{e}_1, \boldsymbol{e}_3\} \subset \mathbb{Z}^4_+,\ A^2_{\tau^2} = \{\boldsymbol{0}, \boldsymbol{e}_2, \boldsymbol{e}_4\} \subset \mathbb{Z}^4_+,\ A^3_{\tau^3} = \{\boldsymbol{0}, \boldsymbol{e}_1, \boldsymbol{e}_2\} \subset \mathbb{Z}^4_+, \\ \theta^1_{\tau^1} & = & 1 + x_1^2 + x_3^2,\ \theta^2_{\tau^2} = 1 + x_2^2 + x_4^2,\ \theta^2_{\tau^2} = 1 + x_1^2 + x_2^2, \end{array} \tag{24}$$

where $\boldsymbol{e}_j \in \mathbb{Z}^4_+$ denote the jth unit coordinate vector $(j = 1, 2, 3, 4)$.

Lemma 4.3. *Suppose that $\mathbb{Z}_+ \ni \omega \geq \omega_{\min}$. Then, $\zeta^d_\omega \uparrow\omega = \zeta^d_\infty$.*

Proof. The inequality $\zeta^d_\omega \leq \zeta^d_\infty$ follows from the definitions of $\theta^k_\tau(\boldsymbol{x})$ and $\boldsymbol{\Theta}^k[\boldsymbol{x}]$ in (18). Letting $\omega_{\min} \leq \omega_1 < \omega_2$, we now show that $\zeta^d_{\omega_1} \leq \zeta^d_{\omega_2}$. Suppose that $(y_0, y_1, \dots, y_m)$ is a feasible solution of (21) with $\omega = \omega_1$. Then,

$$\sum_{k=1}^m \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_{\omega_1}] \ \ni \ f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \sum_{i \in I_k} \theta^k_{\tau^i_1}(\boldsymbol{x}) (f^i(\boldsymbol{x}))^2, \tag{25}$$

where

$$\tau_1^i = \omega_1 - \deg(f^i(\boldsymbol{x})) \ (i \in I_k, \ k = 1,\ldots,m).$$

Let

$$\tau_2^i = \omega_2 - \deg(f^i(\boldsymbol{x})) \ (i \in I_k, \ k = 1,\ldots,m).$$

Then,

$$\begin{aligned}
&\tau_2^i > \tau_1^i \ (i \in I_k, \ k = 1,\ldots,m), \\
&\theta^k_{\tau_2^i}(\boldsymbol{x}) \left(f^i(\boldsymbol{x})\right)^2 - \theta^k_{\tau_1^i}(\boldsymbol{x}) \left(f^i(\boldsymbol{x})\right)^2 = \left(\sum_{\boldsymbol{\alpha}\in\mathcal{F}^i} \boldsymbol{x}^{2\boldsymbol{\alpha}}\right) \left(f^i(\boldsymbol{x})\right)^2 \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_{\omega_2}] \\
&\qquad (i \in I_k, \ k = 1,\ldots,m),
\end{aligned}$$

where

$$\mathcal{F}^i = \mathcal{A}_{\tau_2^i} \backslash \mathcal{A}_{\tau_1^i} = \left\{\boldsymbol{\alpha} \in \mathcal{A}_{\tau_2^i} : \boldsymbol{\alpha} \notin \mathcal{A}_{\tau_1^i}\right\} = \left\{\boldsymbol{\alpha} \in \mathcal{A}_{\tau_2^i} : |\boldsymbol{\alpha}| > \tau_1^i\right\} \quad (i \in I_k, \ k = 1,\ldots,m).$$

It follows from (25) that

$$\begin{aligned}
\sum_{k=1}^m \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_{\omega_2}] \ni\ & f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \sum_{i\in I_k} \theta^k_{\tau_1^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 \\
& + \sum_{k=1}^m y_k \sum_{i\in I_k} \left(\theta^k_{\tau_2^i}(\boldsymbol{x}) \left(f^i(\boldsymbol{x})\right)^2 - \theta^k_{\tau_1^i}(\boldsymbol{x}) \left(f^i(\boldsymbol{x})\right)^2\right) \\
=\ & f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^m y_k \sum_{i\in I_k} \theta^k_{\tau_2^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2. \qquad (26)
\end{aligned}$$

Hence, $(y_0, y_1, \ldots, y_m)$ remains a feasible solution of SOS problem (21) with $\omega = \omega_2$. We have shown that $\zeta^d_{\omega_1} \leq \zeta^d_{\omega_2}$.

Finally, we show that ζ^d_ω converges to ζ^d_∞ as $\omega \to \infty$. Let $\epsilon > 0$. Then there exists a feasible solution $(y_0, y_1, \ldots, y_m)$ of (17) such that $y_0 \geq \zeta^d_\infty - \epsilon$. Thus,

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

for some $\sigma^i \in \mathbb{Z}_+$ $(i \in I_k, \ k = 1, 2, \ldots, m)$ and some

$$\omega \geq \max\left\{\sigma^i + \deg\left(f^i(\boldsymbol{x})\right) : i \in I_k, \ k = 1, 2, \ldots, m\right\}.$$

Now, define τ^i $(i \in I_k, \ k = 1, 2, \ldots, m)$ by (22). Then $\tau^i \geq \sigma^i$ $(i \in I_k, \ k = 1, 2, \ldots, m)$, and we can prove that

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

by the same way as (26) has been derived from (25). As a result, $(y_0, y_1, \ldots, y_m)$ is a feasible solution of (21) 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 ζ^d_ω converges to ζ^d_∞ as $\omega \to \infty$. □

4.5 Reducing SOS problem (21) to COP (7)

To derive a COP of the form (7) equivalent to SOS problem (21), we need to convert the SOS condition

$$f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^{m} y_k \sum_{i\in I_k} \theta^k_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 \in \sum_{k=1}^{m} \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega] \tag{27}$$

to a linear matrix inequality. By (23), we can represent the left hand side of (27) as

$$f^0(\boldsymbol{x}) - y_0 + \sum_{k=1}^{m} y_k \sum_{i\in I_k} \theta^k_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 = \left\langle \boldsymbol{Q}^0 - \boldsymbol{H}^0 y_0 + \sum_{k=1}^{m} \boldsymbol{Q}^k y_k,\, \boldsymbol{x}^{\square\mathcal{A}_\omega} \right\rangle. \tag{28}$$

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

$$\begin{aligned} f^0(\boldsymbol{x}) &= \left\langle \boldsymbol{Q}^0,\, \boldsymbol{x}^{\square\mathcal{A}_\omega} \right\rangle \in \sum_{k=1}^{m} \mathbb{R}[\boldsymbol{x}, \mathcal{A}^k_{\omega_{\min}} + \mathcal{A}^k_{\omega_{\min}}], \\ 1 &= \left\langle \boldsymbol{H}^0,\, \boldsymbol{x}^{\square\mathcal{A}_\omega} \right\rangle \in \bigcap_{k=1}^{m} \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega], \\ \sum_{i\in I_k} \theta^k_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 &= \left\langle \boldsymbol{Q}^k,\, \boldsymbol{x}^{\square\mathcal{A}_\omega} \right\rangle \in \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega] \ (k = 1, 2, \ldots, m). \end{aligned} \tag{29}$$

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

Suppose that $\omega = \omega_{\min} + 1 = 3$ in Example 4.1. By (24), we see that

$$\begin{aligned} \sum_{i\in I_1} \theta^1_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 &= (x_1 - x_3^2)^2 + \left(x_1(x_1 - x_3^2)\right)^2 + \left(x_3(x_1 - x_3^2)\right)^2, \\ \sum_{i\in I_2} \theta^1_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 &= (x_2 - x_4^2)^2 + \left(x_2(x_2 - x_4^2)\right)^2 + \left(x_4(x_2 - x_4^2)\right)^2, \\ \sum_{i\in I_3} \theta^1_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 &= (x_1 + x_2 + x_1x_2 - 1)^2 + (x_1(x_1 + x_2 + x_1x_2 - 1))^2 \\ &\quad + (x_2(x_1 + x_2 + x_1x_2 - 1))^2. \end{aligned}$$

Thus, each $\boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{A}_\omega}$ $(k = 1, 2, 3)$ can be represented as the sum of 3 rank-1 positive semidefinite matrices such that $Q^k_{\boldsymbol{\alpha\beta}} = 0$ if $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \notin \square\mathcal{A}^k_\omega = \mathcal{A}^k_\omega \times \mathcal{A}^k_\omega$.

We now begin to discuss the structured sparsity in the matrices $\boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{A}_\omega}$ $(k = 0, 1, \ldots, m)$ by introducing some notation and symbols. The discussion presented here is essentially an extension of the one for cones in the linear space $\mathbb{S}^{1+n}$ in Section 5 of Part I [5], where DNN and Lagrangian-DNN relaxations for a class of sparse QOPs were discussed, to cones in the linear space $\mathbb{S}^{\mathcal{A}_\omega}$. A finite subset $\mathcal{E}$ of $\square\mathcal{A}_\omega = \mathcal{A}_\omega \times \mathcal{A}_\omega$ is said to be symmetric if $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{E}$ implies $(\boldsymbol{\beta}, \boldsymbol{\alpha}) \in \mathcal{E}$. For every cone $\mathbb{K} \subset \mathbb{S}^{\mathcal{A}_\omega}$ and every symmetric subset $\mathcal{E}$ of $\square\mathcal{A}_\omega$, we use the following symbols and notation:

$$\begin{aligned} \mathcal{E}^c &= \{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \square\mathcal{A}_\omega : (\boldsymbol{\alpha}, \boldsymbol{\beta}) \notin \mathcal{E}\}, \\ \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}, 0) &= \left\{ \boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \square\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega} : X_{\boldsymbol{\alpha\beta}} = 0 \text{ if } (\boldsymbol{\alpha}, \boldsymbol{\beta}) \notin \mathcal{E} \right\}, \\ \mathbb{K}(\mathcal{E}, 0) &= \mathbb{K} \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}, 0), \\ \mathbb{K}(\mathcal{E}, ?) &= \mathbb{K} + \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c, 0). \end{aligned}$$

See Section 5 of Part I [5] for the meaning of the cone $\mathbb{K}(\mathcal{E}, ?)$. In the subsequent discussion, the following identities $\left(\mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}, 0)\right)^* = \left(\mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}, 0)\right)^\perp = \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c, 0)$ will be frequently used. These identities can be verified easily.

Let $\mathcal{E}_\omega = \bigcup_{k=1}^{m} \Box\mathcal{A}^k_\omega$. Then $\Box\mathcal{A}^k_\omega$ $(k = 1, 2, \ldots, m)$ and $\mathcal{E}_\omega$ form symmetric sets in $\Box\mathcal{A}_\omega$. By the relation (29), we can choose matrices $\boldsymbol{Q}^0$ and $\boldsymbol{Q}^k$ $(k = 1, \ldots, m)$ from $\mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)$ and $\mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0) \subset \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0)$ $(k = 1, 2, \ldots, m)$, respectively. On the other hand, by Lemma 2.2, each cone $\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega]$ on the right hand side of (27) can be expressed as

$$\mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega] \quad = \quad \left\{\langle \boldsymbol{W}^k, \boldsymbol{x}^{\Box\mathcal{A}_\omega}\rangle : \boldsymbol{W}^k \in \mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0)\right\}$$

$(k = 1, 2, \ldots, m)$. Hence, we can rewrite the inclusion relation (27) as

$$\begin{aligned} &\left\langle \boldsymbol{Q}^0 - \boldsymbol{H}^0 y_0 + \sum_{k=1}^{m} \boldsymbol{Q}^k y_k, \, \boldsymbol{x}^{\Box\mathcal{A}_\omega} \right\rangle = \left\langle \sum_{k=1}^{m} \boldsymbol{W}^k, \, \boldsymbol{x}^{\Box\mathcal{A}_\omega} \right\rangle, \\ &\boldsymbol{W}^k \in \mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0) \ (k = 1, 2, \ldots, m). \end{aligned}$$

or

$$\boldsymbol{Q}^0 - \boldsymbol{H}^0 y_0 + \sum_{k=1}^{m} \boldsymbol{Q}^k y_k \in \sum_{k=1}^{m} \mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0) + (\mathbb{L}^{\mathcal{A}_\omega})^\perp$$

by Lemma 2.1. Furthermore, since all matrices $\boldsymbol{H}^0$, $\boldsymbol{Q}^k$ $(k = 0, 1, \ldots, m)$ and all cones $\mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0)$ $(k = 1, 2, \ldots, m)$ are included in the linear subspace $\mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)$ of $\mathbb{S}^{\mathcal{A}_\omega}$, $(\mathbb{L}^{\mathcal{A}_\omega})^\perp$ can be replaced by $(\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)$. Therefore, the previous inclusion relation is equivalent to:

$$\boldsymbol{Q}^0 - \boldsymbol{H}^0 y_0 + \sum_{k=1}^{m} \boldsymbol{Q}^k y_k \in \sum_{k=1}^{m} \mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0) \ + \ (\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0). \tag{30}$$

Lemma 4.4. *Suppose that Condition (A) holds. Then,*

(i) $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) = \bigcap_{k=1}^{m} \left\{\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \Box\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega} : (X_{\boldsymbol{\alpha\beta}} : \Box\mathcal{A}^k_\omega) \in \mathbb{S}^{\mathcal{A}^k_\omega}_+\right\}$.

(ii) $\sum_{k=1}^{m} \mathbb{S}^{\mathcal{A}_\omega}_+(\Box\mathcal{A}^k_\omega, 0) = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0) = \left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)\right)^*, \quad \left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0)\right)^* = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)$.

(iii) $\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?) = \left\{\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \Box\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega} : \begin{array}{l} X_{\boldsymbol{\alpha\beta}} = X_{\boldsymbol{\gamma\delta}} \textit{ if } \boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\gamma} + \boldsymbol{\delta}, \\ (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{E}_\omega \textit{ and } (\boldsymbol{\gamma}, \boldsymbol{\delta}) \in \mathcal{E}_\omega \end{array}\right\}$.

(iv) $\left(\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)\right)^* = (\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0), \quad \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?) = \left((\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)\right)^*$.

Proof. (i) By definition, we know that $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) = \mathbb{S}^{\mathcal{A}_\omega}_+ + \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}^c_\omega, 0)$. Thus, $\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \Box\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega}$ lies in $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)$ if and only if there exists an $\overline{\boldsymbol{X}} = (\overline{X}_{\boldsymbol{\alpha\beta}} : \Box\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega}_+$ such that $\overline{X}_{\boldsymbol{\alpha\beta}} = X_{\boldsymbol{\alpha\beta}}$ $((\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{E}_\omega)$. The problem of finding such an $\overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{A}_\omega}_+$ or such values $\overline{X}_{\boldsymbol{\alpha\beta}}$ $((\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathcal{E}^c_\omega)$ is called the positive semidefinite matrix completion in the literature [14, 13, 23]. We consider an undirected graph $G(\mathcal{A}_\omega, \mathcal{E}^o_\omega)$ with the node set $\mathcal{A}_\omega$ and the edge set $\mathcal{E}^o_\omega = \{(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathcal{E}_\omega : \boldsymbol{\alpha} \neq \boldsymbol{\beta}\}$, where each edge $(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathcal{E}^o_\omega$ is identified with $(\boldsymbol{\beta},\boldsymbol{\alpha}) \in \mathcal{E}^o_\omega$. By using Condition (A), we can prove that $G(\mathcal{A}_\omega, \mathcal{E}^o_\omega)$ forms a chordal graph and that $\mathcal{A}^1_\omega, \mathcal{A}^2_\omega, \ldots, \mathcal{A}^m_\omega$ provide the family of maximal cliques of the graph. See Lemma 6.1 of [18]. In this case, $\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \square\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega}$ lies in $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)$ if and only if $(X_{\boldsymbol{\alpha\beta}} : \square\mathcal{A}^k_\omega)$ $(k = 1, 2, \ldots, m)$ are all positive semidefinite [14]. Therefore, assertion (i) follows. This assertion also follows from assertion (i) of Lemma 5.2 of Part I [5].

(ii) In assertion (i), we have shown that $G(\mathcal{A}_\omega, \mathcal{E}^o_\omega)$ forms a chordal graph and that $\mathcal{A}^1_\omega, \mathcal{A}^2_\omega, \ldots, \mathcal{A}^m_\omega$ provide the family of maximal cliques of the graph. By Theorem 2.3 of [1], we obtain the identity $\sum_{k=1}^{m} \mathbb{S}^{\mathcal{A}_\omega}_+(\square\mathcal{A}^k_\omega, 0) = \mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0) = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0)$. This identity may be regarded as the dual of assertion (i). (See also Theorem 4.2 of [15]). Since $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)$ and $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0)$ are closed cones by (i) and definition, respectively, we see that

$$\begin{aligned}\left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0)\right)^* &= \left(\mathbb{S}^{\mathcal{A}_\omega}_+ \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)\right)^* = \mathrm{cl}\ \left(\mathbb{S}^{\mathcal{A}_\omega}_+ + \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c_\omega, 0)\right) = \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?),\\ \left(\mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)\right)^* &= \left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0)\right)^{**} = \mathrm{cl}\ \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0) = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, 0).\end{aligned}$$

This assertion also follows from Lemmas 5.1 and 5.2 of Part I of [5].

(iii) Let

$$\mathbb{M} = \left\{ \boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \square\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega} : \begin{array}{l} X_{\boldsymbol{\alpha\beta}} = X_{\boldsymbol{\gamma\delta}} \text{ if } \boldsymbol{\alpha}+\boldsymbol{\beta} = \boldsymbol{\gamma}+\boldsymbol{\delta}, \\ (\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathcal{E}_\omega \text{ and } (\boldsymbol{\gamma},\boldsymbol{\delta}) \in \mathcal{E}_\omega \end{array} \right\}.$$

By definition, $\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?) = \mathbb{L}^{\mathcal{A}_\omega} + \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c_\omega, 0)$. Hence, if $\boldsymbol{X} \in \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$, then $\boldsymbol{X} \in \mathbb{M}$ obviously. Now suppose that $\boldsymbol{X} = (X_{\boldsymbol{\alpha\beta}} : \square\mathcal{A}_\omega) \in \mathbb{M}$. Then we can consistently define $\overline{\boldsymbol{X}} = (\overline{X}_{\boldsymbol{\alpha\beta}} : \square\mathcal{A}_\omega) \in \mathbb{S}^{\mathcal{A}_\omega}$ such that

$$\overline{X}_{\boldsymbol{\alpha\beta}} = \begin{cases} X_{\boldsymbol{\gamma\delta}} & \text{if there is a } (\boldsymbol{\gamma},\boldsymbol{\delta}) \in \mathcal{E}_\omega \text{ such that } \boldsymbol{\alpha}+\boldsymbol{\beta} = \boldsymbol{\gamma}+\boldsymbol{\delta}, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that $\overline{\boldsymbol{X}} \in \mathbb{L}^{\mathcal{A}_\omega}$ and $\boldsymbol{X} - \overline{\boldsymbol{X}} \in \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c_\omega, 0)$. Thus, $\boldsymbol{X} = \overline{\boldsymbol{X}} + (\boldsymbol{X} - \overline{\boldsymbol{X}}) \in \mathbb{L}^{\mathcal{A}_\omega} + \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c_\omega, 0) = \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$.

(iv) Since both $\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$ and $(\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)$ are linear subspaces of $\mathbb{S}^{\mathcal{A}_\omega}$, we see that

$$\begin{aligned}\left((\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)\right)^* &= \left((\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)\right)^\perp \\ &= \mathbb{L}^{\mathcal{A}_\omega} + \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}^c_\omega, 0) = \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?), \\ \left(\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)\right)^* &= \left((\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0)\right)^{\perp\perp} = (\mathbb{L}^{\mathcal{A}_\omega})^\perp \cap \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0).\end{aligned}$$

□

By Lemma 4.4, the inclusion relation (30) can be rewritten as

$$\boldsymbol{Q}^0 - \boldsymbol{H}^0 y_0 + \sum_{k=1}^{m} \boldsymbol{Q}^k y_k \in \left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)\right)^* + \left(\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)\right)^*.$$

Consequently, we obtain the following primal-dual pair of COPs:

$$\zeta^p_\omega \ := \ \inf \left\{ \langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle \ \middle| \ \begin{array}{l} \boldsymbol{X} \in \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?), \\ \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle = 1, \ \langle \boldsymbol{Q}^k, \boldsymbol{X} \rangle = 0 \ (k = 1, 2, \ldots, m) \end{array} \right\}. \tag{31}$$

$$\zeta^d_\omega \ := \ \sup \left\{ y_0 \in \mathbb{R} \ \middle| \ \begin{array}{l} \boldsymbol{Q}^0 - \boldsymbol{H}^0 y_0 + \sum_{k=1}^m \boldsymbol{Q}^k y_k \\ \in \left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)\right)^* + \left(\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)\right)^* \\ (y_0, y_1, \ldots, y_m) \in \mathbb{R}^{1+m} \end{array} \right\}. \tag{32}$$

The problem (32) is equivalent to the SOS problem (21).

Theorem 4.1. *Suppose that the feasible region of POP*(14) *is nonempty and that Conditions (A) and (B) hold. Let* $\mathbb{Z}_+ \ni \omega \geq \omega_{\min}$. *Then* $\zeta^d_\omega \leq \zeta^p_\omega \leq \zeta^*$ *and* $\zeta^d_\omega \uparrow \omega = \zeta^*$.

Proof. The inequality $\zeta^d_\omega \leq \zeta^p_\omega$ follows from the standard weak duality. To prove the inequality $\zeta^p_\omega \leq \zeta^*$, suppose that $\boldsymbol{x} \in \mathbb{R}^n$ is a feasible solution of (14) with the objective value $f^0(\boldsymbol{x})$. Let $\boldsymbol{X} = \boldsymbol{x}^{\square \mathcal{A}_\omega}$. By (29), $\boldsymbol{X}$ is a feasible solution of (31) with the same objective value $\langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle = f^0(\boldsymbol{x})$. As a result, the inequality $\zeta^p_\omega \leq \zeta^*$ holds. The relation $\zeta^d_\omega \uparrow \omega = \zeta^*$ follows from the equivalence between (32) and (23) and Lemmas 4.1, 4.1 and 4.3. □

If the previous discussion is applied to Example 4.1, the following is obtained:

$$\begin{array}{rcl} C^1 & = & \{1, 3\}, \ C^2 = \{2, 4\}, \ C^3 = \{1, 2\}, \\ \mathcal{A}_\omega & = & \{\boldsymbol{\alpha} \in \mathbb{Z}^4_+ : |\boldsymbol{\alpha}| \leq \omega\}, \ |\mathcal{A}_\omega| = \begin{pmatrix} n + \omega \\ \omega \end{pmatrix}, \\ \mathcal{A}^k_\omega & = & \{\boldsymbol{\alpha} \in \boldsymbol{A}_\omega : \alpha_i = 0 \ (i \notin C^k)\}, \ |\mathcal{A}^k_\omega| = \begin{pmatrix} 2 + \omega \\ \omega \end{pmatrix} \ (k = 1, 2, 3), \\ \square \mathcal{A}^k_\omega & = & \mathcal{A}^k_\omega \times \mathcal{A}^k_\omega, \ \boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{A}_\omega}_+(\square \mathcal{A}^k_\omega, 0) \ (k = 1, 2, 3), \\ \mathcal{E}_\omega & = & \bigcup_{k=1}^3 \square \mathcal{A}^k_\omega, \ |\mathcal{E}_\omega| \leq \sum_{k=1}^3 \left(|\mathcal{A}^k_\omega|\right)^2, \ \boldsymbol{Q}^0 \in \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, 0), \\ \boldsymbol{H}^0 & \in & \mathbb{S}^{\mathcal{A}_\omega}_+ \text{ with } H^0_{\boldsymbol{00}} = 1 \text{ and } H^0_{\boldsymbol{\alpha\beta}} = 0 \ ((\boldsymbol{\alpha}, \boldsymbol{\beta}) \neq (\boldsymbol{0}, \boldsymbol{0})), \end{array}$$

If we take $\omega = 3$, then the size of the matrices $\boldsymbol{Q}^k$ $(k = 0, 1, 2, 3)$ and $\boldsymbol{H}^0$ is 35×35, and the nonzero elements of the matrices correspond to the elements with indices $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{E}_\omega$, where $|\mathcal{E}_\omega|$ is bounded by 300. If $\omega = 4$, then the size and the bound are 70×70 and 675, respectively. This shows that it is computationally meaningful to exploit sparsity even for this small example.

Since $\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$ is described in terms of linear equalities as shown in Lemma 4.4, COP (31) forms an SDP, which has been constructed to satisfy a structured sparsity characterized by the maximal cliques $C^1, C^2, \ldots, C^m$ of the chordal graph $G(N, E)$. For such an SDP, we could use a primal-dual interior-point method combined with the technique for exploiting sparsity via positive semidefinite matrix completion [13, 23]. An additional important feature of COP (31) is that it satisfies Condition (I), which is used to derive the Lagrangian-SDP relaxation of POP (14) in the next subsection.

4.6 Lagrangian-SDP relaxations of COPs (31) and (32)

If $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$, then the problems (31) and (32) coincide with the primal-dual pair of COPs (6) and (7), respectively. We are now ready to apply the general discussions on COPs (6) and (7) given in Sections 2, Part I [5]. Let $\boldsymbol{H}^1 = \sum_{k=1}^{m} \boldsymbol{Q}^k$, and $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$ for their Lagrangian-conic relaxation problems (8) and (9). Then we obtain:

$$\eta^p_\omega(\lambda) \;:=\; \inf\left\{ \langle \boldsymbol{Q}^0 + \lambda \boldsymbol{H}^1,\, \boldsymbol{X}\rangle \;\middle|\; \begin{array}{l} \boldsymbol{X} \in \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?), \\ \langle \boldsymbol{H}^0,\, \boldsymbol{X}\rangle = 1 \end{array} \right\}. \tag{33}$$

$$\eta^d_\omega(\lambda) \;:=\; \sup\left\{ y_0 \in \mathbb{R} \;\middle|\; \begin{array}{l} \boldsymbol{Q}^0 + \lambda \boldsymbol{H}^1 - \boldsymbol{H}^0 y_0 \\ \in \left(\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)\right)^* + \left(\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)\right)^*,\ y_0 \in \mathbb{R} \end{array} \right\}. \tag{34}$$

Theorem 4.2. *Let $\omega_{min} \leq \omega \in \mathbb{Z}_+$. The following results hold.*

(i) $\eta^d_\omega(\lambda) \uparrow \lambda = \zeta^d_\omega$.

(ii) $\eta^d_\omega(\lambda) = \eta^p_\omega(\lambda)$ *for every* $\lambda \in \mathbb{R}$. *The problem* (33) *attains the optimal value at a feasible solution.*

Proof. (i) It suffices to show that Condition (I) holds for $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$. Then, the desired result follows from Lemma 2.3 of Part I [5]. By construction, we know that $\boldsymbol{H}^0,\ \boldsymbol{Q}^k \in \mathbb{S}^{\mathcal{A}_\omega}_+$ $(k = 1, 2, \ldots, m)$. On the other hand, $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega} \subset \mathbb{S}^{\mathcal{A}_\omega}_+$. Therefore $\mathbb{K}^* \supset \mathbb{S}^{\mathcal{A}_\omega}_+$, and Condition (I) follows.

(ii) By (i) and (ii) of Lemma 4.4, we see that $\mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?)$ and $\mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$ are both closed convex cones. And, so is their intersection. Therefore, Condition (II) holds for $\mathbb{K} = \mathbb{S}^{\mathcal{A}_\omega}_+(\mathcal{E}_\omega, ?) \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$, and assertion (ii) follows from Lemma 2.4 of Part I [5]. □

By Theorems 4.1 and 4.2, we can conclude under Conditions (A) and (B) that the optimal value ζ^* can be bounded from below by the optimal value $\eta^d_\omega(\lambda)$ of its Lagrangian-SDP relaxation (34) with any accuracy if we take sufficiently large ω and λ.

Remark 4.1. We have derived the Lagrangian-SDP relaxation problem (34) of (14) from SDP (31). By the same argument as above, we can prove directly that (34) is equivalent to SOS problem:

$$\eta^d_\omega(\lambda) \;:=\; \sup\left\{ y_0 \in \mathbb{R} \;\middle|\; \begin{array}{l} f^0(\boldsymbol{x}) - y_0 + \lambda \displaystyle\sum_{k=1}^{m} \sum_{i \in I_k} \theta^k_{\tau^i}(\boldsymbol{x})(f^i(\boldsymbol{x}))^2 \\ \in \displaystyle\sum_{k=1}^{m} \mathbb{SOS}[\boldsymbol{x}, \mathcal{A}^k_\omega],\ y_0 \in \mathbb{R} \end{array} \right\}.$$

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 $\lambda \geq 0$.

4.7 A brief discussion on the applications of the bisection and 1-dimensional Newton methods

In order to solve the primal-dual pair of COPs (33) and (34), it is possible to apply the bisection and 1-dimensional Newton methods proposed in Part I [5]. See also the numerical method in [16], consisting of a bisection method (Algorithm A of [16]), a proximal alternating direction multiplier method [11] (Algorithm B) and an accelerated proximal gradient method [6] (Algorithm C). In fact, we can transform (33) and (34) to a primal-dual pair of COPs with a cone $\mathbb{K} = \mathbb{K}_1 \cap \mathbb{K}_2$ in some linear space $\mathbb{V}$ such that the metric projections from $\mathbb{V}$ onto the cones $\mathbb{K}_1$ and $\mathbb{K}_2$ are easily computed. See also Remark 4.1 of Part I [5]. This transformation is essentially the same as the one called the conversion method in [13, 23]. Let

$$\begin{aligned} \mathbb{V} &= \prod_{k=1}^{m} \mathbb{S}^{\mathcal{A}^k_\omega} = \left\{ \boldsymbol{Y} = (\boldsymbol{Y}^1, \boldsymbol{Y}^2, \ldots, \boldsymbol{Y}^m) : \boldsymbol{Y}^k \in \mathbb{S}^{\mathcal{A}^k_\omega} \ (k = 1, 2, \ldots, m) \right\}, \\ \mathbb{K}_1 &= \left\{ \boldsymbol{Y} = (\boldsymbol{Y}^1, \boldsymbol{Y}^2, \ldots, \boldsymbol{Y}^m) \in \mathbb{V} : \ \boldsymbol{Y}^k \in \mathbb{S}^{\mathcal{A}^k_\omega}_+ \ (k = 1, 2, \ldots, m) \right\}, \\ \mathbb{K}_2 &= \left\{ \boldsymbol{Y} = (\boldsymbol{Y}^1, \boldsymbol{Y}^2, \ldots, \boldsymbol{Y}^m) \in \mathbb{V} : \begin{array}{l} Y^j_{\boldsymbol{\alpha}\boldsymbol{\beta}} = Y^k_{\boldsymbol{\gamma}\boldsymbol{\delta}} \text{ if } (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \Box\mathcal{A}^j_\omega, \\ (\boldsymbol{\gamma}, \boldsymbol{\delta}) \in \Box\mathcal{A}^k_\omega \ \text{ and } \boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\gamma} + \boldsymbol{\delta} \end{array} \right\}. \end{aligned}$$

Each $\boldsymbol{Y} \in \mathbb{V}$ can be identified with a symmetric block diagonal matrix. We use the notation $\langle \boldsymbol{U}, \boldsymbol{Y} \rangle = \sum_{k=1}^m \langle \boldsymbol{U}^k, \boldsymbol{Y}^k \rangle$ for the inner product of $\boldsymbol{U} = (\boldsymbol{U}^1, \boldsymbol{U}^2, \ldots, \boldsymbol{U}^m)$, $\boldsymbol{Y} = (\boldsymbol{Y}^1, \boldsymbol{Y}^2, \ldots, \boldsymbol{Y}^m) \in \mathbb{V}$. The cone $\mathbb{K}_1$ is described as the Cartesian product of positive semidefinite cones of smaller dimensions, and the cone $\mathbb{K}_2$ is actually a linear subspace of $\mathbb{V}$. Hence, it is straightforward to implement the metric projections from $\mathbb{V}$ onto both cones.

We now associate each $\boldsymbol{X} \in \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$ with $\boldsymbol{Y} = (\boldsymbol{Y}^1, \boldsymbol{Y}^2, \ldots, \boldsymbol{Y}^m) \in \mathbb{V}$ through $\boldsymbol{Y}^k = (X_{\boldsymbol{\alpha}\boldsymbol{\beta}} : \Box\mathcal{A}^k_\omega)$ $(k = 1, 2, \ldots, m)$. With this correspondence, we see that $\boldsymbol{X} \in \mathbb{S}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)_+ \cap \mathbb{L}^{\mathcal{A}_\omega}(\mathcal{E}_\omega, ?)$ if and only if $\boldsymbol{Y} \in \mathbb{K}_1 \cap \mathbb{K}_2$. Furthermore, we can choose $\widetilde{\boldsymbol{Q}}^0$, $\widetilde{\boldsymbol{H}}^0$, $\widetilde{\boldsymbol{H}}^1 \in \mathbb{V}$ such that

$$\langle \widetilde{\boldsymbol{Q}}^0, \boldsymbol{Y} \rangle = \langle \boldsymbol{Q}^0, \boldsymbol{X} \rangle, \ \langle \widetilde{\boldsymbol{H}}^0, \boldsymbol{Y} \rangle = \langle \boldsymbol{H}^0, \boldsymbol{X} \rangle \ \text{ and } \langle \widetilde{\boldsymbol{H}}^1, \boldsymbol{Y} \rangle = \langle \boldsymbol{H}^1, \boldsymbol{X} \rangle$$

hold. Consequently, we obtain the following primal-dual pair of COPs which are equivalent to the primal dual pair of COPs (33) and (34):

$$\begin{aligned} \eta^p_\omega(\lambda) &:= \inf \left\{ \langle \widetilde{\boldsymbol{Q}}^0 + \lambda \widetilde{\boldsymbol{H}}^1, \boldsymbol{X} \rangle \ \middle| \ \boldsymbol{Y} \in \mathbb{K}_1 \cap \mathbb{K}_2, \ \langle \widetilde{\boldsymbol{H}}^0, \boldsymbol{X} \rangle = 1 \right\}. \\ \eta^d_\omega(\lambda) &:= \sup \left\{ y_0 \in \mathbb{R} \ \middle| \ \widetilde{\boldsymbol{Q}}^0 + \lambda \widetilde{\boldsymbol{H}}^1 - \widetilde{\boldsymbol{H}}^0 y_0 \ \in \mathbb{K}^*_1 + \mathbb{K}^*_2, \ y_0 \in \mathbb{R} \right\}. \end{aligned}$$

For details, see [13, 16, 23].

5 Concluding remarks

For POP (1) with $\mathbb{J} = \mathbb{R}^n_+$, two different approaches, both based on the discussions in Section 4, can be used. The first one is the hierarchy of sparse Lagrangian-DNN relaxations, obtained by replacing SDP cones with DNN cones in the construction of the hierarchy of sparse Lagrangian-SDP relaxations in Section 4. The second one is the hierarchy of sparse Lagrangian-SDP relaxation for the reformulated equality constrained POP over $\mathbb{R}^{2n}$, obtained from adding $x_i - x_{i+n}^2 = 0$ $(i = 1, 2, \ldots, n)$ and replacing the cone $\mathbb{R}^n_+$ by $\mathbb{R}^{2n}$. We have taken the second approach in Example 4.1. 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.

There remain some issues to be investigated for an efficient implementation of the hierarchy of sparse Lagrangian-SDP relaxations. We can utilize some of the first-order algorithms in [16], where the Lagrangian-DNN relaxation was implemented to solve QOPs by applying a bisection method with the proximal alternating direction multiplier method [11] and the accelerated proximal gradient method [6]. We can also utilize the 1-dimensional Newton method proposed in Part I [5]. In addition, how to handle sparsity, which was not considered in [16], is an important issue. We hope to report some numerical results in the near future.

References

[1] J. Agler, J. W. Helton, S. McCullough, L. Rodman. Positive semidefinite matrices with a given sparsity pattern. *Linear Algebra Appl.*, 107, 101–149 (1988).

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

[3] N. Arima, S. Kim and M. Kojima. Simplified Copositive and Lagrangian Relaxations for Linearly Constrained Quadratic Optimization Problems in Continuous and Binary Variables. Research Report B-469, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, October (2012).

[4] N. Arima, S. Kim and M. Kojima. Extension of Completely Positive Cone Relaxation to Polynomial Optimization. Research Report B-471, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, February (2013).

[5] 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, Research Report B-475, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, January (2014).

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

[7] J. R. S. Blair and B. Peyton. *An introduction to chordal graphs and clique trees, in Graph Theory and Sparse Matrix Computation.* A. George, J. R. Gilbert and J. W. H. Liu, eds., Springer-Verlag, New York, 1–29 (1993).

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

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

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

[11] 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, 946–977 (2013).

[12] 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).

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

[14] R. Gron, C. R. Johnson, E. M. Sá and H. Wolkowicz. Positive definite completions of partial hermitian matrices. *Linear Algebra Appl.*, 58, 109–124 (1984).

[15] 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, 33–68 (2011).

[16] 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. Research Report B-472, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152-8552, Japan, October (2013).

[17] S. Kim, M. Kojima and H. Waki. Generalized Lagrangian duals and sums of squares relaxations of sparse polynomial optimization problems. *SIAM J. Optim.*, 15, 697–719 (2005).

[18] K. Kobayashi, S. Kim and M. Kojima. Correlative sparsity in primal-dual interior-point methods for LP, SDP, and SOCP. *Appl. Math. Optim.*, 58, 69–88 (2008).

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

[20] M. Kojima and M. Muramatsu. A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones. *Comput. Optim. Appl.*, 42, 31–41 (2009).

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

[22] J. B. Lasserre. Convergent SDP-relaxation in polynomial optimization with sparsity. *SIAM J. Optim.*, 17, 822–843 (2006).

[23] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima, and K. Murota. Exploiting sparsity in semidefinite programming via matrix completion. II: Implementation and numerical results. *Math. Program.*, 95, 303–327 (2003).

[24] 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).

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

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

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

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