RUTCOR
RESEARCH
REPORT

Semidefinite Characterization of Sum-Of-Squares Cones in Algebras

Dávid Papp [a]   Farid Alizadeh [b]

RRR 14-2011, August 2, 2011

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]RUTCOR, Rutgers–the State University of New Jersey, 640 Bartholomew Rd, Piscataway, NJ 08854, USA. E-mail: alizadeh@rutcor.rutgers.edu
[b]Department of Management Science and Information Systems, and RUTCOR, Rutgers–the State University of New Jersey, 640 Bartholomew Rd, Piscataway, NJ 08854, USA. E-mail: alizadeh@rutcor.rutgers.edu

Rutcor Research Report
RRR 14-2011, August 2, 2011

Semidefinite Characterization of Sum-Of-Squares Cones in Algebras

Dávid Papp, Farid Alizadeh

Abstract. We extend Nesterov's semidefinite programming characterization of squared functional systems to cones of sum-of-squares elements in general abstract algebras. Using algebraic techniques such as isomorphism, linear isomorphism, tensor products, sums and direct sums, we show that many concrete cones are in fact sum-of-squares cones with respect to some algebra, and thus representable by the cone of positive semidefinite matrices. We also consider nonnegativity with respect to a proper cone $\mathcal{K}$, and show that in some cases $\mathcal{K}$-nonnegative cones are either sum-of-squares, or are semidefinite representable. For example we show that some well-known Chebyshev systems, when extended to Euclidean Jordan algebras, induce cones that are either Sum-of-Squares cones or are semidefinite representable. Finally we will discuss some concrete examples and applications, including minimum ellipsoid enclosing given space curves, minimization of eigenvalues of polynomial matrix pencils, approximation of functions by shape-constrained functions, and approximation of combinatorial optimization problems by polynomial programming.

Acknowledgements: The authors gratefully acknowledge support of the National Science Foundation through Grant # NSF-CMMI-0935305.

1 Introduction

Consider a *general algebra* $(A, B, \diamond)$, where A and B are real, finite dimensional linear spaces. The binary operation $\diamond\colon A \times A \to B$ is *bilinear* on A, and the product of the operation is in B. In this paper we first show that the *sum-of-squares (SOS) cone* in B, namely

$$\Sigma_\diamond = \Big\{ \sum_{i=1}^{N} a_i \diamond a_i \mid a_i \in A, N \text{ a positive integer} \Big\},$$

is representable by the cone of real symmetric positive semidefinite bilinear forms on A, that is they are *SD-representable*, see [30, 3]. As a result, both the problem of deciding whether a vector $b \in B$ is also in $\Sigma_\diamond$, and the conic optimization problems over $\Sigma_\diamond$, are reducible to semidefinite programming.

We use common algebraic tools such as direct sums, tensor products, homomorphism, and isomorphism to show that many well-know cones are in fact SOS cones with respect to some general algebra, and thus representable by positive semidefinite matrices. We also show that there is a close connection between formally real property of algebras and "*properness*" of their SOS cone (see below for the definition of proper cones).

Of particular interest are algebras induced by a set of linearly independent, real-valued functions $\mathcal{F} = \{f_1, \dots, f_m\}$, all defined on a set Δ. Such algebras are named *squared functional systems (SFS)* by Nesterov [29], who also analyzed their SOS cone and showed that they are SD-representable. Indeed, our results in this paper are direct generalization of Nesterov's to abstract algebras.

One of our contributions in this paper is to examine tensor products of a squared functional systems with arbitrary algebras. This approach results in new and rich classes of functional systems, and corresponding SOS cones. In particular we examine some cases of $\mathcal{K}$-*nonnegative* functions (for a proper cone $\mathcal{K}$) which are also SOS cones or SD-representable. To achieve this we use the fundamental result of Youla [42], where he showed that matrix polynomial pencils which are symmetric and positive semidefinite for all $t \in \mathbb{R}$ are SOS. For instance, by taking tensor product of univariate real polynomials $\mathbb{R}[t]$, and a *Euclidean Jordan algebra* (see [11] and [17]), we can show that the cone of functions $p(t) = \sum_i p_i t^i$ which are in a *symmetric cone* $\mathcal{K}_\mathbb{J}$ for all $t \in \mathbb{R}$ or t in some closed or semi-closed interval, is SD-representable. In a similar manner, several other functional systems over Euclidean Jordan algebras are also shown to be SD-representable. In fact, symmetric cones are also SOS cones with respect to the Jordan algebra multiplication. Kojima and Muramatsu [18] consider this problem in the special case when f is a polynomial whose coefficients are taken from a formally real Jordan algebra, and derive a semidefinite characterization for them. As we show in Section 6.4, polynomials in this construction can be replaced by other functional systems, and the coefficients can be chosen from arbitrary (and not just Jordan) algebras. Multivariate SOS matrix polynomials are also considered in [19].

We close the paper by showing some concrete applications of our results.

Before we end this section let us review some background. Optimization models involving *nonnegative* functions in linear functional spaces (such as polynomials, for instance) have raised significant interest in recent years. A special subset of nonnegative functions are those which can be expressed as sums of squares of other functions in a (possibly different) linear function space.

Both nonnegative, and sum-of-squares (SOS) functions in a linear functional space are convex cones. However, there is no general method known to decide whether a given function is nonnegative over its domain, even in the most fundamental functional systems. For example, recognizing nonnegative (multivariate) polynomials of degree four is known to be NP-hard, by a simple reduction from the Partition problem. On the other hand, by Nesterov's result as mentioned earlier, the constraint that a function belong to a specific SOS functional system can always be cast as a semidefinite programming constraint.

The SD-representability of the SOS cones of functional systems presents an opportunity to approximate "nonnegative" functions in these systems by the SOS functions. One area where this approach has yielded interesting results is *polynomial programming (POP)*, that is, optimization models involving polynomial minimization, or equivalently, positive polynomial constraints. For example, POP has been applied to combinatorial optimization problems [22, 25, 40]. A key underlying result in all POP models is that a univariate polynomial is nonnegative everywhere if and only if it is a sum of squared polynomials. In multivariate polynomials, sums of squared polynomials form a proper subset of nonnegative polynomials. On the other hand, Hilbert's seventeenth problem, resolved by Artin, asserts that any nonnegative multivariate polynomial is sum of squares of *rational* functions. However, the bounds on dimensions of such linear spaces of rational functions may be extremely large, and in fact in many cases, no satisfactory bound is even known, see the texts [35, 24, 28] for details and further references. An even more dramatic result holds for trigonometric multivariate polynomials: Each such polynomial may be expressed as sum of squares of other polynomials. However, again, the degree of such polynomials is usually very large, see for example, [10, Chapters 3 & 4].

One way to cope with intractability of nonnegative functions is to give inner approximations by replacing them with SOS functions instead, or by SOS function in *truncated spaces*, if the required dimensions are too large. In shape constrained statistical estimation problems, for instance, consider an infinite dimensional function space $\mathcal{F}$ and the cone of nonnegative functions $\mathcal{P}$ in $\mathcal{F}$. In [31, 32] we have given some weak necessary conditions for a nested sequence of SOS cones Σ_d, each in a nested sequence of finite dimensional function spaces F_d (such as polynomials of degree d, or polynomial splines with d knots) such that $\bigcup_d F_d$ is dense in $\mathcal{F}$ and $\bigcup_d \Sigma_d$ is dense in $\mathcal{P}$. Nested hierarchies of SOS models, such as those proposed by Lasserre [23] and Parrilo [33] define nested sequences of SOS polynomial cones to approximate the cone of nonnegative multivariate polynomials.

Our paper is primarily motivated by shape constrained optimization problems, but the general theory we develop also has immediate applications to combinatorial optimization and POP (Section 7.4). As the following examples show, not all shape constraints can be translated naturally to nonnegativity of some linear transform of the shape constrained real valued function.

1. (Convexity of a multivariate function.) A twice continuously differentiable real valued function f is convex over $\Delta \subset \mathbb{R}^n$ if and only if its Hessian H is positive semidefinite over S, which we denote by $H(x) \succcurlyeq 0$, $\forall\, x \in \Delta$. Magnani et. al [27] consider this problem in the special case when f is a multivariate polynomial, and suggest the following approach: $H(x) \succcurlyeq 0$ for every $x \in S$ if and only if $h(x, y) = y^\top H(x)y \geq 0$ for every $x \in \Delta$ and $y \in \mathbb{R}^n$. Since h is also a polynomial, the problem is reduced to POP. This reduction, however, is not entirely satisfying for two reasons. First, it calls for doubling the number of variables at the outset. Second, if f is not a polynomial, and H is not a polynomial matrix, then H is still a linear transform of f, however, $h(x, y)$ may belong to an entirely different functional system, in which it is generally difficult to establish a connection between nonnegativity and sum of squares properties of functions.

2. (Space paths with bounded curvature.) Consider a twice differentiable curve given by its parametric representation $x(s) \in \mathbb{R}^n$, where $s \in [0, S]$ is the arc-length parameter. Suppose we wish to design such a path $x(s)$ under the constraint that its curvature must be bounded above by some constant $C \geq 0$. This constraint can be written as

$$\|x''(s)\| \leq C \quad \forall\, s \in (0, S),$$

where x'' is the component-wise second derivative of the vector-valued function x, and $\|\cdot\|$ is the Euclidean norm [39, Section 1-4]. Equivalently, the constraint can be written as

$$\left(\begin{smallmatrix} C \\ x''(s) \end{smallmatrix}\right) \in \mathcal{Q}_{n+1} \quad \forall\, s \in (0, S),$$

where $\mathcal{Q}_{n+1}$ is the $(n+1)$-dimensional second order cone, or the Lorentz cone [1]. This is an example of a problem in which a vector-valued function $x(\cdot)$ which is required to be $\mathcal{Q}$-nonnegative.

3. (Estimating positive semidefinite matrix functions) Let $A : \Delta \to \mathbb{S}_m$ be a function defined on a subset $\Delta \subseteq \mathbb{R}$ which only takes positive semidefinite values in $\mathbb{S}_m$, the set of symmetric $m \times m$ matrices. This type of problem arises for instance, in statistical applications in time series. Consider a series of observed data for a number of functions $f_i(t)$ in the form of (t_j, y_{ij}) where y_{ij} is the observed (and noisy) value of $f_i(t_j)$. Our goal is to estimate the variance-covariance matrix of these functions from these observations The variance-covariance matrix in this context is itself a matrix-valued time series which should be estimated from data, and needs to be positive semidefinite for all t.

In all these examples we are using some notion of "nonnegativity" or "positivity" beyond real numbers, and extended to linear spaces. However, in general there is no natural way of extending these concepts to higher dimensional linear spaces. The most common approach is to use the partial order induced by a *proper cone*, that is a closed, pointed, convex and full-dimensional cone $\mathcal{K}$. In this case our constraints are of the form $f(x) \in \mathcal{K}, \forall x \in \Delta$, where f is a (perhaps multivariate) vector-valued function. Such requirements are called *cone-nonnegativity* or *$\mathcal{K}$-nonnegativity* constraints. As in the motivating one-dimensional case, characterizing $\mathcal{K}$-nonnegative multivariate functions is in general intractable. In this paper we examine some special cases where these $\mathcal{K}$-nonnegative functions are either SOS functions with respect to some algebra, or are SD-representable.

2 Algebraic preliminaries

An algebraic system $(A, B, \diamond)$ is called a *general algebra* when A and B are two linear spaces, and $\diamond\colon A \times A \to B$ is a bilinear binary operation. In this paper we are primarily interested in finite dimensional linear spaces over the field of real numbers $\mathbb{R}$. Note that bilinearity is equivalent to the fact that $\diamond$ follows the distributive law over addition:

$$a_0 \diamond (\beta_1 a_1 + \beta_2 a_2) = \beta_1 a_0 \diamond a_1 + \beta_2 a_0 \diamond a_2$$
$$(\beta_1 a_1 + \beta_2 a_2) \diamond a_0 = \beta_1 a_1 \diamond a_0 + \beta_2 a_2 \diamond a_0$$

for all $\beta_1, \beta_2 \in \mathbb{R}$ and $a_0, a_1, a_2 \in A$. The algebraic system $(A, B, \diamond)$ is called an *algebra* when $A = B$. In this paper it is necessary to allow A and B to be different. As a result we use the term algebra also for general algebras. In many instances it is customary to use no symbol for the binary operation at all, for instance write ab. In such cases we write (A, B) for the underlying algebra.

Let $\dim(A) = m$, $\dim(B) = n$, and $\mathbb{M}_{A,B}$ be the set of linear operators from A to B. The *left multiplication operator* $L_\diamond$ for the algebra $(A, B, \diamond)$ is defined as $L_\diamond\colon A \to \mathbb{M}_{A,B}$ where for all $y \in A$, we have $L_\diamond(x)y = x \diamond y$. Since the binary operation "$\diamond$" is bilinear, it follows that $L_\diamond$ is a linear operator; indeed, the representation of $L_\diamond(x)$ is an $n \times m$ matrix in any basis consisting of linear forms in x.

We make a few assumptions which will prove convenient without any loss in generality.

Assumption 1: Let us define $A \diamond A = \{a \diamond b \mid a, b \in A\}$. Then there is no loss in generality to assume that $B = \mathrm{Span}(A \diamond A)$. When $B = \mathrm{Span}(A \diamond A)$ we say that the algebra $(A, B, \diamond)$ is *generated* by A. More generally, if $\mathcal{A}$ is a set in a linear space, then we say the algebra $(A, B, \diamond)$ is *generated* by $\mathcal{A}$, if $A = \mathrm{Span}(\mathcal{A})$, and $B = \mathrm{Span}(\mathcal{A} \diamond \mathcal{A})$. All this assumption does is to rule out extraneous elements in B, that is those that are not obtained as linear combinations of products in A.

Another way to view algebras generated by a linear space A is to note that if $\{e_i\}$, for $i = 1, \dots, m$ is a basis for A, the $B = \mathrm{Span}(e_i \diamond e_j)$ for $i, j = 1, \dots m$. In this way one can manufacture new algebras from subsets of algebraic structures.

Assumption 2: It will be convenient to assume that "$\diamond$" is commutative, that is for any $x, y \in A$, we have $x \diamond y = y \diamond x$. We will see shortly that this assumption will not cause any loss of generality for our purposes. In fact, for any binary operation "$\diamond$" we may define its *commutative version* $\bar{\diamond}$ as $x \bar{\diamond} y = \frac{x \diamond y + y \diamond x}{2}$. Thus, unless specifically stated otherwise, we assume that all algebras mentioned in this paper are either commutative, or that their binary operation is replaced by their commutative version.

Assumption 3: If L is not injective operator, then for some elements $x_1, x_2 \in A$ we may have $L_\diamond(x_1) = L_\diamond(x_2)$. This implies that for all $y \in A$, $x_1 \diamond y = x_2 \diamond y$. Clearly such a relationship defines an equivalence relation on A. Let $[x]$ be the equivalence class of elements $x_i \in A$ where $(x - x_i) \in \operatorname{Ker}(L_\diamond)$. In this case if we replace A by the set of equivalence classes, that is the linear space $A/\operatorname{Ker}(L_\diamond)$, then we can define a new algebra $\big(A/\operatorname{Ker}(L_\diamond), B, \diamond_1\big)$, that is we may define $[x] \diamond_1 [y] = x \diamond y$. From commutativity of "$\diamond$" it is easily verified that this definition is consistent, and that $L_{\diamond_1}$ is an injective operator. As a result, unless specifically stated otherwise, we assume that $L_\diamond$ is injective. Equivalently, we assume that if for all $y \in A$ $x_1 \diamond y = x_2 \diamond y$, then $x_1 = x_2$.

Let $(A_1, B_1, \diamond_1)$ and $(A_2, B_2, \diamond_2)$ be two algebras. Let $F\colon A_1 \to A_2$ and $G\colon B_1 \to B_2$ be two linear mappings. Then we say the pair (F, G) is a *homomorphism* if

$$G(x \diamond_1 y) = F(x) \diamond_2 F(y) \quad \text{for all } x, y \in A_1.$$

When both F and G are bijective, then the pair (F, G) is called an *isomorphism*.

When $A_1 = A_2$ and F is the identity map I, then the linear mapping $G\colon B_1 \to B_2$ is called a *linear homomorphism* if (I, G) is a homomorphism. Furthermore, if G is bijective, then G is a *linear isomorphism*. Note that in case of linear homomorphisms, the left multiplication operators are related by $L_{\diamond_2} = GL_{\diamond_1}$. In fact, if $(A, B, \diamond)$ is an algebra, and $G\colon B \to B_1$ is a linear transformation to another space B_1, then we may define a new algebra $(A, B_1, \diamond_1)$, with $L_{\diamond_1} = GL_\diamond$. If G is bijective, then the new algebra is linearly isomorphic to the former one.

Finally, the usual homomorphism (respectively isomorphism) of (ordinary) algebras $(A_1, \diamond)$ and $(A_2, \circ)$ is the special case when $F = G$.

We now recall some notions and notations from linear algebra. Let B be a finite dimensional linear space. Let $\langle \cdot, \cdot \rangle_B$ be an inner product on B, and $\langle \cdot, \cdot \rangle_A$ an inner product on A, (such inner products always exist, since A and B are assumed to be finite dimensional).

Let $\mathbb{M}_A$ be the set of all bilinear forms on the space A. Then $T^\top$ is the *transpose* of a bilinear form T if $T^\top(a_1, a_2) \stackrel{\text{def}}{=} T(a_2, a_1)$. The set $\mathbb{S}_A$ of *symmetric* bilinear forms (also called *quadratic forms*) is the set of forms S such that for all $a_1, a_2 \in A$ we have $S(a_1, a_2) = S(a_2, a_1)$; clearly for a symmetric form S we have $S = S^\top$. A symmetric form $S \in \mathbb{S}_A$ is *positive semidefinite*–written $S \succcurlyeq 0$–if and only if $S(a, a) \geq 0$ for all $a \in A$.

Any pair of vectors $a, b \in A$ induces a bilinear form $B_{a,b}$ defined as follows: $\big(B_{a,b}\big)(x, y) \stackrel{\text{def}}{=} \langle a, x \rangle_A \langle b, y \rangle_A$. In particular if $a = b$ then $Q(a) \stackrel{\text{def}}{=} B_{a,a}$ is a symmetric, positive semidefinite form. We note that the sets $\mathbb{M}_A$ and $\mathbb{S}_A$ themselves are finite dimensional linear spaces, and therefore are endowed with some inner product. In addition we note that the evaluation functional of bilinear forms, that is $T(a,b)$ induces an inner product on $\mathbb{M}_A$ as follows. First for each arbitrary form T and each induced form $B_{a,b}$ define $\langle T, B_{a,b}\rangle_{\mathbb{M}_A} \stackrel{\text{def}}{=} T(a,b)$. Then since each arbitrary form R can be expressed as a sum $R = \sum_i B_{a_i,b_i}$, the inner product can now be extended by $\langle R, T\rangle_{\mathbb{M}_A} \stackrel{\text{def}}{=} \sum_i T(a_i,b_i)$. It can be shown that this definition is consistent and results in the same value regardless of how the form R is expressed as $\sum_i B_{a_i,b_i}$. We use the same inner product for the subspace $\mathbb{S}_A$ of symmetric forms, only we use the notation $\langle\cdot,\cdot\rangle_{\mathbb{S}_A}$. Note that if S_1 and S_2 are both symmetric and positive semidefinite, then $\langle S_1, S_2\rangle_{\mathbb{S}_A} \geq 0$.

For two linear spaces A and B, as stated earlier, $\mathbb{M}_{A,B}$ is the set of linear transformations form A to B; when $A = B$, we simply write $\mathbb{M}_A$. Recall that there is a representation of linear transformations as bilinear forms. Given a linear transformation $T : A \to A$, we can uniquely define a bilinear form $T(a,b) \stackrel{\text{def}}{=} \langle a, T(b)\rangle_A$. In this paper $\mathbb{M}_A$ represents linear transformations on A, bilinear forms on A, or their matrix representation, as all these three spaces are isomorphic.

3 Semidefinite Characterization of Sums of Squares in Algebras

For the algebra $(A, B, \diamond)$ we define the *sum of squares (SOS) cone*, $\Sigma_\diamond \subseteq B$, as the convex cone generated by *square elements* $a \diamond a$:

$$\Sigma_\diamond \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{N} a_i \diamond a_i \;\middle|\; \text{for some } N \geq 1;\; a_1, \ldots, a_N \in A \right\}.$$

Note that for a non-commutative "$\diamond$" the SOS cones $\Sigma_\diamond = \Sigma_{\bar\diamond}$, where "$\bar\diamond$" is the commutative version of "$\diamond$". Therefore, as mentioned earlier, there is no loss of generality in assuming that $\diamond$ is commutative. In what follows, when we give an example of an operation which is not commutative, the reader may assume that it is replaced by its commutative version when necessary.

Lemma 1. *The cone $\Sigma_\diamond$ is a closed and convex cone, where every element is sum of at most $n = \dim B$ squares. Furthermore, $\Sigma_\diamond$ is full-dimensional in the space* $\operatorname{Span}(A \diamond A)$.

Proof. Convexity follows from definition. To show closedness, note that $\Sigma_\diamond$ is generated by $\{a \diamond a \mid \|a\| \leq 1\}$, which is the image of the closed unit ball under the continuous mapping $x \to x \diamond x$. Since the unit ball is compact, then this set is also compact. Thus, since $\Sigma_\diamond$ is generated by a compact set, it must be closed. The extreme rays of $\Sigma_\diamond$ are among perfect squares $a \diamond a$, and by Carathéodory's theorem for cones, each element can be written as a sum of at most n extreme rays. To prove full-dimensional property, note that every element of $\mathrm{Span}(A \diamond A)$ is sum of elements of the form $a \diamond b$. But by the "completing the square" technique, and noting the commutativity assumption, we can write

$$a \diamond b = \left(\tfrac{a+b}{2}\right)^{\diamond 2} - \left(\tfrac{a-b}{2}\right)^{\diamond 2}.$$

We have shown that $\mathrm{Span}(A \diamond A) = \Sigma_\diamond - \Sigma_\diamond$, that is $\Sigma_\diamond$ generates $\mathrm{Span}(A \diamond A)$, proving that it is full-dimensional. □

Recall our assumption earlier that $B = \mathrm{Span}(A \diamond A)$. Thus, $\Sigma_\diamond$ is full-dimensional in B. However, $\Sigma_\diamond$ still may not be a proper cone (that is a closed, convex, pointed, and full-dimensional in the space B), as it may contain lines.

Let $(A, B, \diamond)$ be an algebra, with $\Sigma_\diamond$ its SOS cone. Define the operator $\Lambda_\diamond \colon B \to \mathbb{S}_A$, as:

$$\big(\Lambda_\diamond(w)\big)(a_1, a_2) \stackrel{\text{def}}{=} \langle w, a_1 \diamond a_2 \rangle_B. \tag{1}$$

Note that since "$\diamond$" is commutative, $\Lambda(w)$ is always a symmetric form.

Our main theorem is the characterization of the sum of squares cone $\Sigma_\diamond$ as a linear image of the cone of positive semidefinite forms. Such sets are called *SD-representable*. More precisely, following Nesterov and Nemirovski [30], we say that a set X is SD-representable if

$$X = \left\{\mathbf{x} \in \mathbb{R}^n \mid \exists\, \mathbf{u} \in \mathbb{R}^k \text{ satisfying } C(\mathbf{x}) + D(\mathbf{u}) \succcurlyeq F\right\}$$

for some linear mappings $C : \mathbb{R}^n \to \mathbb{S}_A$, and $D : \mathbb{R}^k \to \mathbb{S}_A$, and for some form $F \in \mathbb{S}_A$. The notation $S_1 \succcurlyeq S_2$ means that the form $S_1 - S_2 \succcurlyeq 0$. In particular, affine images and affine pre-images of positive semidefinite forms are SD-representable.

Our development of SD-representability of $\Sigma_\diamond$ is a direct extension of Nesterov's for sum-of-squares functional systems in [29]. We first characterize the dual cone $\Sigma_\diamond^* = \{z \mid \langle x, z\rangle_B \geq 0 \text{ for all } x \in \Sigma_\diamond\}$.

Theorem 2. *For the dual cone of* $\Sigma_\diamond$ *we have:* $\Sigma_\diamond^* = \{w \mid \Lambda_\diamond(w) \succcurlyeq 0\}$.

Proof. If $v \in \Sigma_\diamond^*$, then for all $a \in A$, we have

$$0 \leq \langle v, a \diamond a\rangle_B = \Lambda_\diamond(v)(a, a)$$

and thus $\Lambda_\diamond(v) \succcurlyeq 0$.

Conversely, if $\Lambda_\diamond(v) \succcurlyeq 0$ then for all $a \in A$,

$$0 \leq \Lambda_\diamond(v)(a, a) = \langle v, a \diamond a\rangle_B$$

and therefore, $v \in \Sigma_\diamond^*$. □

As has been observed by Nesterov and Nemirovski [30], if a cone is SD-representable, then so is its dual. Since Theorem 2 shows that $\Sigma_\diamond^*$ is SD-representable, it follows that $\Sigma_\diamond$ is also SD-representable. The following theorem makes this representation explicit. As $\Lambda_\diamond$ is a linear operator from B to $\mathbb{S}_A$, its adjoint operator $\Lambda^* : \mathbb{S}_A \to B$ is defined by the relation $\langle X, \Lambda(w)\rangle_{\mathbb{S}_A} = \langle w, \Lambda^*(X)\rangle_B$ for all $X \in \mathbb{S}_A$ and $w \in A$.

Theorem 3. *$u \in \Sigma_\diamond$ if and only if there exists a quadratic form $Y \succcurlyeq 0$ such that $u = \Lambda_\diamond^*(Y)$.*

Proof. Suppose $Y \succcurlyeq 0$ is a form on A such that $\Lambda_\diamond^*(Y) = u$. Then for all $v \in \Sigma_\diamond^*$ we have

$$\langle v, u\rangle_B = \langle v, \Lambda_\diamond^*(Y)\rangle_B = \langle Y, \Lambda_\diamond(v)\rangle_{\mathbb{S}_A} \geq 0.$$

The last inequality is due to fact that by Theorem 2, $\Lambda_\diamond(v) \succcurlyeq 0$. Therefore, $u \in \Sigma_\diamond^{**}$, and since $\Sigma_\diamond$ is closed, $\Sigma_\diamond^{**} = \Sigma_\diamond$.

Conversely, let $u \in \Sigma_\diamond$; therefore, there are $a_i \in A$ where $u = \sum_i a_i \diamond a_i$. Let $v \in B$ be any vector. Then

$$\langle u, v\rangle_B = \Big\langle \sum_i a_i \diamond a_i, v\Big\rangle_B = \sum_i \Lambda_\diamond(v)(a_i, a_i) = \sum_i \big\langle \Lambda_\diamond(v), Q(a_i)\big\rangle_{\mathbb{S}_A} = \big\langle \Lambda_\diamond^*(Y), v\big\rangle_B,$$

where $Y = \sum_i Q(a_i)$ and thus $Y \succcurlyeq 0$. Comparing the first and the last terms, and noting that this is true for all $v \in B$, we conclude that $u = \Lambda_\diamond^*(Y)$. $\square$

The definition of the cone $\Sigma_\diamond$ is applicable to *any* algebra $(A, B, \diamond)$ with virtually no restrictions. However, not all such algebras result in interesting cones. We start with two trivial examples to see what the issues are.

Example 1. Let $A = B = \mathbb{C}$, the set of complex numbers, viewed as a two-dimensional algebra over $\mathbb{R}$, with $\diamond$ being the usual complex number multiplication: $\left(\begin{smallmatrix} a_1 \\ b_1 \end{smallmatrix}\right) \diamond \left(\begin{smallmatrix} a_2 \\ b_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a_1a_2 - b_1b_2 \\ a_1b_2 + a_2b_1 \end{smallmatrix}\right)$. Using the standard basis for $\mathbb{C}$, the matrix representation of $\Lambda(w)$ is given by $\left(\begin{smallmatrix} w_1 & w_2 \\ w_2 & -w_1 \end{smallmatrix}\right)$, which can be positive semidefinite only if $w_1 = w_2 = 0$. Thus, $\Sigma_\diamond^* = \{0\}$, and $\Sigma_\diamond = \mathbb{C}$. This, of course, is expected as all complex numbers are perfect squares. A similar situation holds for quaternions (identified with $\mathbb{R}^4$ with quaternion multiplication), and octonions (identified with $\mathbb{R}^8$ with octonion multiplication). Indeed for any algebra $(A, B, \diamond)$ where every element $y \in B$ is either a perfect square, or sum of squares, $\Sigma_\diamond = B$ and $\Sigma_\diamond^* = \{0\}$. $\square$

Example 2. Consider the algebra of square matrices $(\mathbb{R}^{n\times n}, \circ)$, where $M \circ N = \frac{MN+NM}{2}$. Then, for matrices of the form

$$M = \begin{pmatrix} 0 & A & 0 \\ 0 & 0 & -A \\ 0 & 0 & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} 0 & A & 0 \\ 0 & 0 & A \\ 0 & 0 & 0 \end{pmatrix},$$

we have $M^{\circ 2} + N^{\circ 2} = 0$. Thus, the cone $\Sigma_\circ$ contains non-trivial subspaces and, as a result, is not pointed. $\square$

To avoid trivial situations, we focus on cases where $\Sigma_\diamond$ is a pointed cone. Recall that a convex cone $\mathcal{K}$ is *pointed* if it does not contain a line, or equivalently, if $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$. As the following theorem shows, a condition sufficient to obtain a pointed SOS cone is that the multiplication $\diamond$ be formally real: $\diamond$ is said to be *formally real* if for any integer k, and every $a_1, \ldots, a_k \in A$, $\sum_{i=1}^k (a_i \diamond a_i) = 0$ implies that each $a_i = 0$.

Theorem 4. *Assume $(A, B, \diamond)$ has no nilpotent elements of degree 2. Then the following statements are equivalent:*

i. $(A, B, \diamond)$ *is formally real.*

ii. $\Sigma_\diamond$ *is a proper cone.*

iii. $\Sigma_\diamond^*$ *is a proper cone.*

iv. $\mathrm{Ker}(\Lambda_\diamond) = \{0\}$.

Proof. To show equivalence of *i.* and *ii.*, suppose that $\diamond$ is formally real and that for some vector x, both x and $-x$ are in $\Sigma_\diamond$. Then $0 = x + (-x) \in \Sigma_\diamond$ is sum of squares: $0 = \underbrace{\sum_i a_i \diamond a_i}_{x} + \underbrace{\sum_i b_i \diamond b_i}_{-x}$, implying that each a_i and b_i is zero. Consequently $x = 0$ confirming that $\Sigma_\diamond$ is pointed.

Conversely, if $\Sigma_\diamond$ is pointed and $\sum_i a_i \diamond a_i = 0$ for some a_i, then each $a_i \diamond a_i = 0$, and thus each $a_i = 0$, since it is not nilpotent of degree two. This proves that $\diamond$ is formally real.

To show equivalence of *ii.* and *iii.*, note that in general a closed convex cone $\mathcal{K}$ is full-dimensional if and only if its dual cone $\mathcal{K}^*$ is pointed. Therefore, $\mathcal{K}$ is both pointed and full-dimensional if and only if its dual cone $\mathcal{K}^*$ is pointed and full-dimensional. Since by our assumptions $\Sigma_\diamond$ is full-dimensional in B, its being pointed is equivalent to $\Sigma_\diamond^*$ being also full-dimensional and pointed.

Finally to show equivalence of *iii.* and *iv.* note that for $w \in B$, we have $\Lambda_\diamond(w) \succcurlyeq 0$ and $\Lambda_\diamond(-w) \succcurlyeq 0$ if and only if $\Lambda_\diamond(w) = 0$. Since both w and $-w$ are in $\Sigma_\diamond^*$, $\Sigma_\diamond^*$ is pointed if and only if $w = 0$, that is $\mathrm{Ker}(\Lambda_\diamond) = \{0\}$. ∎

3.1 Some Examples of Formally Real Algebras and their SOS Cones

Squared Functional Systems (SFS)

Let $\mathcal{F}(x) = \{f_1(x), \ldots, f_m(x)\}$ be a set of linearly independent functions, each mapping a set Δ to $\mathbb{R}$, (Δ can be any set, including subsets of $\mathbb{R}^n$). Let $F(x) = \mathrm{Span}\big(\mathcal{F}(x)\big)$, and $S(x) = \mathrm{Span}(\{f_i(x) f_j(x) \mid i, j = 1, \ldots, m\})$, and the binary operation be the product of functions in $F(x)$: $\big(f_i f_j\big)(x) = f_i(x) f_j(x)$ for $x \in \Delta$. Then the semidefinite characterization of the SOS cone, that is the set of sum-of-squares functions in S, obtained from Theorem 3 is identical to the one developed by Nesterov in [29]. We will write $\Sigma_{\mathcal{F}(x)}$ for this cone; this algebra along with its associated cone $\Sigma_{\mathcal{F}(x)}$ is called a *squared functional system (SFS)* and is written as $(F(x), S(x), \Sigma_{\mathcal{F}(x)})$.

For squared functional systems one can also consider $\mathcal{P}_{\mathcal{F}(x)}$, the cone of nonnegative functions, that is $\mathcal{P}_{\mathcal{F}(x)} = \{g(x) \in S(x) \mid g(x) \geq 0 \text{ for all } x \in \Delta\}$. Then it is clear that $\Sigma_{\mathcal{F}(x)} \subseteq \mathcal{P}_{\mathcal{F}(x)}$. Except for some important cases these cones are in general different.

As a special case consider the SFS induced by the set of multivariate polynomials of degree at most d over real or complex variables $t_1, \ldots, t_k$. In this case we use a more conventional notation, and write $\mathbb{R}_d[t_1, \ldots, t_k]$ for the k-variate polynomials of degree at most d over real numbers. Also the SOS cone for the real polynomial functional system is written as $\Sigma_{2d}[t_1, \ldots, t_k]$; it is composed of polynomials of degree $2d$ over variables $t_1, \ldots, t_k$ which can be written as sum-of-squares of polynomials of degree d. The cone $\Sigma_{2d}[t_1, \ldots, t_k]$ is a proper subset of $\mathcal{P}_{2d}[t_1, \ldots, t_k]$, the cones of *nonnegative* k-variate polynomials, except for some well-known special cases.

As a further specialization, consider the SFS induced by univariate polynomials: $\big(\mathbb{R}_d[t], \mathbb{R}_{2d}[t], \Sigma_{2d}[t]\big)$. This functional system is isomorphic to the algebra $(\mathbb{R}^{d+1}, \mathbb{R}^{2d+1}, *)$ where "$*$" is the *convolution* of vectors: $a * b$. If we represent polynomials by their coefficient vectors using the standard basis $\{1, t, \ldots, t^d\}$ for $\mathbb{R}_d[t]$ and $\{1, t, \ldots, t^{2d}\}$ for $\mathbb{R}_{2d}[t]$, then $\Lambda^*(w)$ may be represented by the Hankel form derived from w, that is, in its matrix representation we have: $\big[\Lambda^*(w)\big]_{ij} = w_{i+j}$. For univariate polynomials the cones $\Sigma_{2d}[t]$ of SOS polynomials and $\mathcal{P}_{2d}[t]$ of nonnegative polynomials are identical, see for example [15, 34].

Symmetric Cracovian algebra

Let A and B be linear spaces. Then for linear transformations $M, N \in \mathbb{M}_{A,B}$, the $A-B$-*Cracovian* multiplication is defined by: $M \diamond N = MN^\top$, see [16]. With this multiplication, the algebra $(\mathbb{M}_{A,B}, \mathbb{M}_A, \diamond)$ is formally real. However, since $\diamond$ is not commutative in general, following our convention we replace it with its commutative version: $M \star N \stackrel{\text{def}}{=} \frac{MN^\top + NM^\top}{2}$. In this case, note that the product $M \star N$ is actually a symmetric linear transformation. We summarize these observations in the following definition.

Definition 5. *Let A and B be finite dimensional linear spaces, $\mathbb{M}_{A,B}$ the set of linear transformations from A to B, and $\mathbb{S}_A$, the set of symmetric linear transformations on A. Then $(\mathbb{M}_{A,B}, \mathbb{S}_A, \star)$, under the multiplication $M \star N = \frac{MN^\top + NM^\top}{2}$ is called the* symmetric Cracovian algebra*.*

If $\dim(A) = m$ and $\dim(B) = n$, we may also write $\mathbb{M}_{m \times n}$ and $\mathbb{S}_m$ for general and symmetric linear transformations, respectively.

In this case it is clear that $\Sigma_\star$ is identical to the cone of positive semidefinite (real, symmetric) $m \times m$ matrices, where $\dim(A) = m$. In this special case we write $\mathcal{P}_{\mathbb{R},m}$ or simply $\mathcal{P}_m$ for $\Sigma_\star$. This example can be extended to complex numbers (with $\mathcal{P}_{\mathbb{C},m}$ for its SOS cone), and quaternions (with $\mathcal{P}_{\mathbb{H},m}$ as the SOS cone). In fact, consider any associative algebra A endowed with a linear, involutive anti-isomorphism '$\prime$', that is, $(x+y)' = x' + y'$, $(x')' = x$ and $(xy)' = y'x'$. Then the generalized Cracovian multiplication is simply $x \diamond y = xy'$, and its symmetric version is $x \star y = \frac{xy'+yx'}{2}$. This multiplication is neither associative, nor even power-associative. However, it is formally real, and its SOS cone $\Sigma_\star$ is a section of the cone of positive semidefinite matrices. For example, consider the complex numbers $\mathbb{C}$ with the multiplication $x \star y = \frac{xy^*+yx^*}{2}$, where y^* is the conjugate of y. Then $(\mathbb{C}, \mathbb{R}, \star)$ is a special case of formally real Cracovian algebra whose SOS cone is the set of nonnegative real numbers.

Another special case is $(\mathbb{R}^n, \mathbb{S}_n, \star)$, which also has the cone of positive semidefinite matrices as its SOS cone. Thus, the same cone may be simultaneously the SOS cone of many different algebras. This gives us some flexibility. For example, for the sake of efficiency we may wish to find the smallest dimension linear space A that generate an SOS cone.

The set of nonnegative real numbers $\mathbb{R}_+$, for instance, is generated in many different ways. The field of real numbers, viewed as an algebra, generates it; so does the *inner product algebra* $\big(A, \mathbb{R}, \langle\cdot,\cdot\rangle_A\big)$. Of course, $\mathbb{R}$ under real number multiplication, is the smallest dimension linear space generating $\mathbb{R}_+$.

Euclidean Jordan Algebras and Symmetric Cones

We now consider Euclidean Jordan algebras and their well-studied SOS cone. Recall that an (ordinary) algebra $(\mathbb{J}, \circ)$ is a *Jordan algebra* if it is commutative, and the *Jordan identity* is satisfied for all $x, y \in \mathbb{J}$:

$$x^{\circ 2}(x \circ y) = x \circ (x^{\circ 2} \circ y).$$

Also, recall that a Jordan algebra is *Euclidean* if every inner product in $(\mathbb{J}, \circ)$ is associative, that is every inner product in $\mathbb{J}$ satisfies $\langle a, b \circ c\rangle_{\mathbb{J}} = \langle a \circ b, c\rangle_{\mathbb{J}}$. Finally, recall that Euclidean Jordan algebras are exactly the formally real Jordan algebras, see [11] and [17] for details. Note that in this case $\mathbb{J} = A = B$, and that these algebras are not necessarily associative. If $(\mathbb{J}, \circ)$ is a Euclidean Jordan algebra then its *cone of squares* $\mathcal{K}_{\mathbb{J}} = \{x \circ x \mid x \in \mathbb{J}\}$ is a *symmetric cone*, that is a cone which is self-dual, and whose automorphism group acts transitively on its interior. Also $\mathcal{K}_{\mathbb{J}} = \Sigma_\circ$, as symmetric cones are convex. In the case of Euclidean Jordan algebras it turns out that $\Lambda_\circ(w)$ is essentially the same as $L_\circ(w)$.

Lemma 6. *In a Euclidean Jordan algebra* $(\mathbb{J}, \circ)$*,* $\Lambda_\circ(w)$ *is the bilinear form representation of* $L_\circ(w)$*.*

Proof. Since every inner product $\langle x, y\rangle_{\mathbb{J}}$ is associative, and $\circ$ is commutative, for each $w \in \mathbb{J}$, and for all $a, b \in \mathbb{J}$, we have:

$$\Lambda_\circ(w)(a, b) = \langle w, a \circ b\rangle_{\mathbb{J}} = \langle w \circ a, b\rangle_{\mathbb{J}} = \langle a, w \circ b\rangle_{\mathbb{J}} = \langle a, L_\circ(w) b\rangle_{\mathbb{J}}. \qquad \square$$

Since $\mathcal{K}_{\mathbb{J}}$ is self-dual, $w \in \mathcal{K}_{\mathbb{J}}$ if and only if $\Lambda_\circ(w) \succcurlyeq 0$, and by the previous lemma this means that $L_\circ(w) \succcurlyeq 0$. This is again well-known in the context of Jordan algebras, and will be used in the following sections.

4 Building SOS Cones from Other SOS Cones

Using familiar operations such as direct sums and tensor products, we may build new algebras, and thus, new SOS cones. In this section, and the next one, we explore some of these operations. We will also go in the other direction, and show that some operations on SOS cones results in other SOS cones. To do so, we need to build their underlying algebras.

4.1 Direct Sums of Algebras and SOS Cones

For algebras $(A_i, B_i, \diamond_i)$ for $i = 1, \ldots, \ell$, and corresponding SOS cones $\Sigma_{\diamond_i}$, and operators $\Lambda_i = \Lambda_{\diamond_i}$, the *direct sum* operation is defined by $\diamond : (A_1 \times \cdots \times A_\ell) \times (A_1 \times \cdots \times A_\ell) \to B_1 \times \cdots \times B_\ell$, where

$$(x_1, \ldots, x_\ell) \diamond (y_1, \ldots, y_\ell) \stackrel{\text{def}}{=} (x_1 \diamond_1 y_1, \ldots, x_\ell \diamond_\ell y_\ell).$$

In this case it is easily seen that $\Sigma_\diamond = \Sigma_1 \times \cdots \times \Sigma_\ell$, and $\Lambda_\diamond\big((w_1, \ldots, w_\ell)\big) = \Lambda_{\diamond_1}(w_\ell) \oplus \cdots \oplus \Lambda_{\diamond_\ell}(w_\ell)$, where $\oplus$ is the direct sum of bilinear forms defined as $(A_1 \oplus A_2)\big((a_1, a_2), (b_1, b_2)\big) = A_1(a_1, b_1) + A_2(a_2, b_2)$.

An example of direct sum SOS cone is the nonnegative orthant, which is direct sum of n copies of $\mathbb{R}_+$, and thus is an SOS cone.

4.2 Minkowski Sums of SOS Cones

For algebras $(A_i, B, \diamond_i)$ the Minkowski sum of their respective SOS cones $\Sigma_{\diamond_i}, i = 1, \ldots, \ell$ is: $\Sigma_{\diamond_1} + \cdots + \Sigma_{\diamond_\ell}$. This cone is a convex cone, and by results of Nesterov and Nemirovski [30] it is SD-representable. However, this Minkowski sum itself is actually an SOS cone for the algebra $(\bigoplus_i A_i, B, \diamond)$, where $(a_1, \ldots, a_\ell) \diamond (b_1, \ldots, b_\ell) = a_1 \diamond_1 b_1 + \cdots + a_\ell \diamond_\ell b_\ell$.

Results of Nesterov and Nemirovski also show that the intersection cone $\bigcap_i \Sigma_{\diamond_i}$ is SD-representable. But it is not known if this cone is SOS.

5 Tensor Products of Algebras and Their Associated SOS and $\mathcal{K}$-nonnegative Cones

5.1 Basic definitions and results

First, let us recall the definition of tensor product of algebras. For two linear spaces A_1 and A_2, their tensor product space is denoted by $A_1 \otimes A_2$; recall that $\dim(A_1 \otimes A_2) = \dim(A_1) \cdot \dim(A_2)$, see Jacobson [14] or Shafarevich [38] for introduction to tensor product spaces, or refer to the readable tutorials by K. Conrad [8, 9]. The *tensor product of algebras* $(A_1, B_1, \diamond_1)$ and $(A_2, B_2, \diamond_2)$, is a new algebra $(A_1 \otimes A_2, B_1 \otimes B_2, \diamond_1 \otimes \diamond_2)$; to have a more concise notation, when possible we write $\diamond$ for $\diamond_1 \otimes \diamond_2$.

To define the operation "$\diamond$" on the tensor product spaces, we first define it for *elementary tensors*, that is the tensor products of two vectors. Let $u_1, v_1 \in A_1$ and $u_2, v_2 \in A_2$:

$$(u_1 \otimes u_2) \diamond (v_1 \otimes v_2) \stackrel{\text{def}}{=} (u_1 \diamond_1 v_1) \otimes (u_2 \diamond_2 v_2).$$

The $\diamond$ operation is then extended to all elements of $A_1 \otimes A_2$ by taking its closure under the distributive law.

For bilinear forms T_1 on A_1 and T_2 on A_2, recall that the tensor product form $T_1 \otimes T_2$ on $A_1 \otimes A_2$ is first defined on elementary tensors as $(T_1 \otimes T_2)(a_1 \otimes a_2, b_1 \otimes b_2) \stackrel{\text{def}}{=} T_1(a_1, b_1) T_2(a_2, b_2)$; this relation also extends to all elements of $A_{\otimes} A_2$ in a unique way.

Next, if $\langle \cdot, \cdot \rangle_{A_1}$ and $\langle \cdot, \cdot \rangle_{A_2}$ are inner products on A_1 and A_2 spaces, then they induce $\langle \cdot, \cdot \rangle_{A_1 \otimes A_2}$ by

$$\langle a_1 \otimes a_2, b_1 \otimes b_2 \rangle_{A_1 \otimes A_2} = \langle a_1, b_1 \rangle_{A_1} \langle a_2, b_2 \rangle_{A_2}$$

and then the inner product is distributively extended to all elements in $A_1 \otimes A_2$.

Finally, note that

$$\big(\operatorname{Span}(A_1 \diamond_1 A_1)\big) \otimes \big(\operatorname{Span}(A_2 \diamond_2 A_2)\big) = \operatorname{Span}\big((A_1 \otimes A_1) \diamond (A_2 \otimes A_2)\big)$$

Therefore, if $\Sigma_{\diamond_1}$ is full-dimensional in B_1 and $\Sigma_{\diamond_2}$ is full-dimensional in B_2, then $\Sigma_\diamond$ is full-dimensional in $B_1 \otimes B_2$.

Lemma 7. *If $w_1 \in B_1$ and $w_2 \in B_2$, then*

$$\Lambda_\diamond(w_1 \otimes w_2) = \Lambda_{\diamond_1}(w_1) \otimes \Lambda_{\diamond_2}(w_2).$$

Thus, for the linear operators $\Lambda_\diamond, \Lambda_{\diamond_1}, \Lambda_{\diamond_2}$ we have, $\Lambda_\diamond = \Lambda_{\diamond_1} \otimes \Lambda_{\diamond_2}$.

Proof.

$$\begin{aligned}\left[\Lambda_\diamond(w_1 \otimes w_2)\right](a_1 \otimes a_2, b_1 \otimes b_2) &= \left\langle w_1 \otimes w_2, (a_1 \otimes a_2) \diamond (b_1 \otimes b_2)\right\rangle_{B_1 \otimes B_2} \\ &= \left\langle w_1 \otimes w_2, (a_1 \diamond_1 b_1) \otimes (a_2 \diamond_2 b_2)\right\rangle_{B_1 \otimes B_2} \\ &= \left\langle w_1, a_1 \diamond_1 b_1\right\rangle_{B_1} \left\langle w_2, a_2 \diamond_2 b_2\right\rangle_{B_2} \\ &= \left[\Lambda_{\diamond_1}(w_1)\right](a_1, b_1)\left[\Lambda_{\diamond_2}(w_2)\right](a_2, b_2)\end{aligned}$$

Noting that, as an operator, $\Lambda_\diamond$ maps $B_1 \otimes B_2$ on to $\mathbb{S}_{B_1 \otimes B_2} = \mathbb{S}_{B_1} \otimes \mathbb{S}_{B_2}$, the assertion of the lemma is proved. □

The tensor product of formally real algebras results in a formally real algebra.

Theorem 8. *The tensor product algebra* $(A_1 \otimes A_2, B_1 \otimes B_2, \diamond_1 \otimes \diamond_2)$ *is formally real if and only if both* $(A_1, B_1, \diamond_1)$ *and* $(A_2, B_2, \diamond_2)$ *are formally real. Consequently,* $\Sigma_{\diamond_1 \otimes \diamond_2}$ *is a proper cone if and only if both* $\Sigma_{\diamond_1}$ *and* $\Sigma_{\diamond_2}$ *are proper.*

Proof. From Theorem 4 we need only to observe that $\mathrm{Ker}(\Lambda_{\diamond_1 \otimes \diamond_2}) = \{0\}$ if and only if $\mathrm{Ker}(\Lambda_{\diamond_1}) = \{0\}$ and $\mathrm{Ker}(\Lambda_{\diamond_2}) = \{0\}$. However, by Lemma 7 $\Lambda_{\diamond_1 \otimes \diamond_2} = \Lambda_{\diamond_1} \otimes \Lambda_{\diamond_2}$. On the other hand, from the properties of tensor product, it immediately follows that the kernel of the tensor product of two linear transformations is trivial, if the kernel of each of those linear transformations is trivial. □

Now, for the linear spaces B_1 and B_2 let $\mathcal{K}_1 \subseteq B_1$ and $\mathcal{K}_2 \subseteq B_2$ be two proper cones. Then we define the tensor product of these two cones as:

$$\mathcal{K}_1 \otimes \mathcal{K}_2 \stackrel{\text{def}}{=} \operatorname{cone}\{u \otimes v \mid u \in \mathcal{K}_1, v \in \mathcal{K}_2\},$$

where, in general, $\operatorname{cone} A$ is the conic hull of A. It is easily verified that for SOS cones, $\Sigma_{\diamond_1} \otimes \Sigma_{\diamond_2} \subseteq \Sigma_{\diamond_1 \otimes \diamond_2}$. However, the inclusion is usually proper.

Through tensor product, new and rich classes of algebras may be constructed. These new algebras, in turn, induce new and rich classes of SOS cones. We first examine some concrete and basic cases.

Tensor product with complex numbers and quaternions

The field of complex numbers $\mathbb{C}$ under ordinary complex number multiplication is a two dimensional algebra which, as noted earlier, is not formally real. The tensor product algebra $(\mathbb{C} \otimes A, \mathbb{C} \otimes B, \diamond_\mathbb{C})$ is sometimes called the *complexification* of the algebra $(A, B, \diamond)$, see Conrad's tutorial [7] for more information. The elements of $\mathbb{C} \otimes A$ and $\mathbb{C} \otimes B$ may be thought of as vectors over the field of complex numbers; they may be represented by $a + bi$

where $a, b \in A$ (respectively, $a, b \in B$), and $i = \sqrt{-1}$. Then the product in the complexified algebra is defined as

$$(a + bi) \diamond_{\mathbb{C}} (c + di) = (a \diamond c - b \diamond d) + (a \diamond d + b \diamond c)i$$

In the complexified algebra, one can also define the *conjugation* operator: $\overline{x + iy} = x - iy$. Recall that the algebra $(\mathbb{C}, \mathbb{R}, \star)$, where $z_1 \star z_2 = \frac{z_1 \bar{z}_2 + z_2 \bar{z}_1}{2}$ is formally real and $\mathbb{R}_+$ is its SOS cone. Now in the tensor product algebra $(\mathbb{C} \otimes A, B, \star \otimes \diamond)$ the tensor algebra product of two such vectors is given by

$$(a + bi) \star \otimes \diamond (c + di) = (a + bi) \diamond_{\mathbb{C}} (\overline{c + di}) = (a \diamond c + b \diamond d)$$

Note that the product is always in $\mathbb{R} \otimes B$ which we identify with B. In particular, due to commutativity of $\diamond$, if $(A, B, \diamond)$ is formally real this tensor product algebra is also formally real.

As a further example, consider two sets of functions $\mathcal{F} = \{f_1(\cdot), \dots, f_n(\cdot)\}$ and $\mathcal{G} = \{g_1(\cdot), \dots, g_n(\cdot)\}$ of real valued functions, where each of f_j and g_j are defined on some set Δ. Consider the algebra where $F = \text{Span}(f_j + ig_j)$ and $S = \text{Span}\big((f_j + ig_j) \star (f_k + ig_k)\big)$. Then the cone of squares of this algebra is given by

$$\Sigma_{\mathcal{F}+i\mathcal{G}} = \Big\{ \sum_j |u_j + iv_j|^2 \mid u_j \in F \text{ and } v_j \in G \Big\},$$

where $| \cdot |$ is the modulus of complex numbers. In this way we recover Nesterov's SOS characterization of sum of squares of moduli of complex-valued functional systems, [29, section 2.3].

A similar construction can be made with quaternions. We may define *quaternionification* and conjugation of tensor products of quaternions and arbitrary algebras (the quaternionification step). Then in the quaternified algebra define $q_1 \star q_2 = \frac{q_1 \overline{q_2} + q_2 \overline{q_1}}{2}$, where, in general, $\overline{q}$ is the quaternionic conjugate of q. This multiplication results in a formally real algebra. Finally, we may give an SOS characterization of the sums of squares of moduli of quaternion-valued functions on Δ.

Tensor product with matrices

We now consider the tensor product of the space of $m \times n$ matrices and an arbitrary algebra $(A, B, \diamond)$. Elements of $\mathbb{R}^{m \times n} \otimes A$ may be thought of as matrices whose entries are in A; we call A-*matrices*, and write $A^{m \times n}$ for $\mathbb{R}^{m \times n} \otimes A$, and $\mathbb{S}_{A,n}$ for symmetric $n \times n$ A-matrices. Let M be an $m \times r$ A-matrix, and N an $r \times n$ A-matrix. Then from definition of tensor products it follows immediately that their product $M \diamond N$ is an $m \times n$ B-matrix whose i, j entry is $\sum_k M_{ik} \diamond N_{kj}$.

We may also consider, instead of the ordinary product, the tensor Cracovian product. In other words we take the tensor product $(A^{m\times n}, \mathbb{S}_{B,m}, \star_\diamond)$, where $\star_\diamond = \star \otimes \diamond$. Therefore, If $M, N \in A^{m\times n}$,

$$[M \star_\diamond N]_{ij} = \frac{\sum_k M_{ik} \diamond N_{kj} + \sum_k N_{ik} \diamond M_{kj}}{2}.$$

Then the SOS cone of the tensor product algebra is given by $\Sigma_{\star_\diamond} = \{\sum a_i \star_\diamond a_i^\top \mid a_i \in A^n\}$. Thus, just like real-valued vectors, a_i is a "column vector" whose entries are from A, and $a_i^\top$ is the corresponding row vector. Again, it is easily seen that this cone can also be generated by $(A^n, \mathbb{S}_{B,m}, \star_\diamond)$.

Tensor product with linear function spaces

We now focus on tensor product of a squared functional system $(F(x), S(x), \Sigma_{\mathcal{F}})$, generated by a set of functions $\mathcal{F} = \{f_1(x), \dots, f_m(x)\}$, and an arbitrary algebra $(A, B, \diamond)$. Note that in this case elements of $F(x) \otimes A$ are functions of the form $\sum_i a_i f_i(x)$, where $a_i \in A$. Likewise, elements of $S(x) \otimes B$ are functions of the form $\sum_i b_i v_i(x)$ where $v_i(x)$ are a basis of $S(x)$. Also, in this case the spaces $A \otimes F(x)$ and $F(x) \otimes A$ are equal. In the tensor product algebra $\big(F(x) \otimes A, S(x) \otimes B, \diamond\big)$,

$$\Sigma_{\diamond\otimes\mathcal{F}} = \Big\{f(x) = \sum_i \Big(\sum_j a_{ij} f_j(x)\Big)^{\diamond 2} \mid a_{ij} \in A\Big\},$$

is the associated SOS cone[1].

The tensor product of a linear space A with a space of real-valued functions $F(x) = \mathrm{Span}\big(\{f_i(x)\}\big)$, forms a linear space of A-valued functions, called *A-functions*, for short. Likewise, the tensor product algebra $(F(x)\otimes A, S(x)\otimes B, \diamond)$ is called an "*A-squared functional systems*". In the case of univariate polynomials, we use the notation $\big(A_d[t], B_{2d}[t], \diamond\big)$ for the tensor product algebra, and $\Sigma_{\diamond\otimes 2d}[t]$ for its SOS cone. In particular, if $\mathcal{F}$ is a basis for (either univariate or multivariate) polynomials, then the corresponding tensor product is called the space of *A-polynomials*.

A well-known example is the tensor product of two copies of univariate polynomials, $(\mathbb{R}_d[t_1], \mathbb{R}_{2d}[t_1])$ and $(\mathbb{R}_d[t_2], \mathbb{R}_{2d}[t_2])$, which results in bivariate polynomials $(\mathbb{R}_d[t_1, t_2], \mathbb{R}_{2d}[t_1, t_2])$. Similarly, multivariate polynomials may be obtained by repeated tensor products of univariate polynomials.

We will examine the SOS nature, or SDP-representability of some cones associated with A-polynomials shortly.

[1] For a list of SOS and nonnegative cones, and the notation used to represent them, refer to the Appendix at the end of the paper.

5.2 $\mathcal{K}$-nonnegative functional systems

It is possible to extend the notion of "nonnegativity" beyond real numbers to linear spaces. However, for many useful applications the price paid is to replace the total order in the real numbers with a partial order, usually induced by a proper cone. Using the notation of the end of previous section (and those listed in the Appendix), let $\mathcal{K}$ be such a cone in B, and let the induced partial order be $\succcurlyeq_{\mathcal{K}}$, that is, $a \succcurlyeq_{\mathcal{K}} b$ when $a - b \in \mathcal{K}$. Then the set of $\mathcal{K}$-*nonnegative functions* in the tensor product space $S(x) \otimes B$ is defined by

$$\mathcal{P}^{\mathcal{F}}_{\mathcal{K},B} \stackrel{\text{def}}{=} \{u(x) \in S(x) \otimes B \mid u(x) \succcurlyeq_{\mathcal{K}} 0 \text{ for all } x \in \Delta\}. \tag{2}$$

When $\mathcal{K} = \Sigma_\diamond$ then $\Sigma_{\diamond\otimes\mathcal{F}(x)} \subseteq \mathcal{P}^{\mathcal{F}}_{\Sigma_\diamond,B}$, that is the set of sum-of-squares functions in $S(x) \otimes A$ is a subset of the set of functions in $S(x) \otimes A$ which are SOS (in $\Sigma_\diamond$) for each value of x. However, this inclusion is generally proper. The relation between $\Sigma_{\diamond\otimes\mathcal{F}(x)}$ and $\mathcal{P}^{\mathcal{F}}_{\Sigma_\diamond,B}$ is somewhat analogous to the relationship between point-wise convergence and uniform convergence of sequences of functions in mathematical analysis. Functions in $\mathcal{P}^{\mathcal{F}}_{\Sigma_\diamond,B}$ are *point-wise* SOS, while those in $\Sigma_{\diamond\otimes\mathcal{F}(x)}$ are SOS as a result of identities which hold *for all* $x \in \Delta$. This situation remains true even if the squared functional system is the one induced by univariate polynomials: $\Sigma_{\diamond\otimes 2d}[t] \subseteq \mathcal{P}_{\Sigma_\diamond,d,\Delta}[t]$ where in general, $\Delta \subseteq \mathbb{R}$ and,

$$\mathcal{P}_{\mathcal{K},d,\Delta}[t] = \big\{p(t) \in B_d[t] \mid p(t) \succcurlyeq_{\mathcal{K}} 0 \text{ for all } t \in \Delta\big\}.$$

As usual, when $\Delta = \mathbb{R}$, we drop it from the subscript and write $\mathcal{P}_{\mathcal{K},d}[t]$.

An important exception is given by Youla's Theorem.

Proposition 9 ([42, Theorem 2 and Corollary 2])**.** *Let $P(t)$ be an $m \times m$ complex Hermitian polynomial matrix, and let r be the largest integer such that $P(t)$ has at least one minor of order r that does not vanish identically. Then there exists an $m \times r$ complex polynomial matrix $Q(t)$ satisfying the identity $P(t) = Q(t)Q(t)^*$, where, in general, M^* is the conjugate transpose of matrix M.*

(Note that even if $P(t)$ is a real symmetric polynomial matrix, $Q(t)$ may still have to be complex.) Thus, Youla's Theorem states that the set of Hermitian matrix valued univariate polynomials which are positive semidefinite for all $t \in \mathbb{R}$, is an SOS cone. In algebraic terms, Youla's Theorem may be expressed in the following form:

Corollary 10. *Let $\mathbb{C}^{m\times n}_d[t]$ be the complex polynomial pencil of $m \times n$ matrices of degree d, and let $\mathbb{H}_{m,2d}[t]$ be the complex polynomial pencil of Hermitian matrices of degree $2d$. Then for the tensor product algebra $(\mathbb{C}^{m\times n}_d[t], \mathbb{H}^m_{2d}[t], \star)$ with $M \star N = \frac{MN^* + NM^*}{2}$, we have*

$$\Sigma_{\star\otimes 2d}[t] = \mathcal{P}_{\mathcal{P}_{\mathbb{C},m},2d}[t],$$

where $\mathcal{P}_{\mathbb{C},m}$ is the cone of positive semidefinite complex Hermitian matrices.

Another special case is the formally real Jordan algebra $(\mathbb{J}, \circ)$, where Youla's Theorem may be used to show that $\mathcal{P}_{\mathcal{K}_\mathbb{J},2d}[t]$ is SD-representable.

Theorem 11. *Let $\mathbb{J}_d[t]$ be the set of $\mathbb{J}$-polynomials $p_0 + p_1 t + \cdots + p_d t^d$, that is each p_i is in the Euclidean Jordan algebra $(\mathbb{J}, \circ)$, and let $\mathcal{K}_\mathbb{J}$ be the symmetric cone of $\mathbb{J}$. Then the cone*

$$\mathcal{P}_{\mathcal{K}_\mathbb{J},2d}[t] \stackrel{\text{def}}{=} \big\{p(t) \mid p(t) \in \mathcal{K}_\mathbb{J} \text{ for all } t \in \mathbb{R}\big\}$$

is SD-representable.

Proof. The polynomial $p(t) \in \mathcal{P}_{\mathcal{K}_\mathbb{J},2d}[t]$ if, and only if $L_\circ\big(p(t)\big) \succcurlyeq 0$ for all $t \in \mathbb{R}$. By Youla's theorem, $L_\circ\big(p(t)\big) \succcurlyeq 0$ for all t if, and only if there is a (possibly complex) polynomial matrix $Q(t)$ where $L_\circ\big(p(t)\big) = Q(t)Q(t)^*$, proving that $\mathcal{P}_{\mathcal{K}_\mathbb{J},2d}[t]$ is SD-representable. □

This proof is actually applicable to a more general setting.

Theorem 12. *Let $\Sigma_\diamond$ be the SOS cone of the algebra $(A, B, \diamond)$. Then the cone of $\Sigma_\diamond^*$-nonnegative B-polynomials, $\mathcal{P}_{\Sigma_\diamond^*,2d}[t]$, is SD-representable.*

Proof. $p(t) \in \Sigma_\diamond^*$ for all t if and only if $\Lambda\big(p(t)\big) \succcurlyeq 0$ for t. This condition, by Youla's theorem is SD-representable. □

We do not know if the cones $\mathcal{P}_{\mathcal{K}_\mathbb{J},2d}[t]$ and $\mathcal{P}_{\Sigma_\diamond^*,d}[t]$ are themselves SOS cones. For instance, in the case of Jordan algebras, we may wonder if a $\mathbb{J}$-polynomial $p(t) \succcurlyeq_{\mathcal{K}_\mathbb{J}} 0$ for all $t \in \mathbb{R}$, then whether there are $\mathbb{J}$-polynomials $q_i(t)$ where $p(t) = \sum_i q_i^2(t)$. However, an elementary argument shows that even for the Jordan algebra of symmetric 2×2 matrices, such $q_i(t)$ do not exist in general.

6 Sums of Squares Representations by Linear Isomorphisms

6.1 Linear isomorphism through homomorphism and function composition

Let B_1 and B_2 be two linear spaces of equal dimensions, and let $\mathcal{K}_1 \subseteq B_1$ and $\mathcal{K}_2 \subseteq B_2$ be two convex cones. We say that $\mathcal{K}_1$ and $\mathcal{K}_2$ are *linearly isomorphic*, and write $K_1 \simeq K_2$, if there is a bijective linear transformation $F : B_1 \to B_2$ such that $\mathcal{K}_2 = F(\mathcal{K}_1)$. If $B_1 = B_2 = B$ then linear isomorphism defines an equivalence relation on convex cones in B.

Theorem 13. *Let (F, G) be a homomorphism from $(A_1, B_1, \diamond_1)$ to $(A_2, B_2, \diamond_2)$. If F is surjective then $\Sigma_{\diamond_2} = G\big(\Sigma_{\diamond_1}\big)$. Furthermore, if G is bijective, then $\Sigma_{\diamond_1}$ and $\Sigma_{\diamond_2}$ are linearly isomorphic.*

Proof. Let $y \in \Sigma_{\diamond_2}$. Then

$$\begin{aligned} y = \sum_i y_i \diamond_2 y_i &= \sum_i F(x_i) \diamond_2 F(x_i) \quad \text{(for some } x_i \in A\text{, since } F \text{ is surjective)} \\ &= \sum_i G(x_i \diamond_1 x_i) \quad \text{(by definition of homomorphism)} \\ &= G\Big(\sum_i x_i \diamond_1 x_i\Big) \in G\big(\Sigma_{\diamond_1}\big) \quad \text{(by linearity).} \end{aligned}$$

The sequence of implications above goes through in both directions, establishing that $\Sigma_{\diamond_2} = G\big(\Sigma_{\diamond_1}\big)$. By definition, if G is bijective, then it is also a linear isomorphism between $\Sigma_{\diamond_1}$ and $\Sigma_{\diamond_2}$. ☐

Lemma 14. *Let $(A, B, \diamond)$ be an algebra and $\Sigma_\diamond$ its SOS cone, and let B_1 be a linear space. Then for every linear transformation F mapping B to B_1 there is an algebra $(A, B_1, \circ)$ such that $F\big(\Sigma_\diamond\big) = \Sigma_\circ$. Furthermore, if F is bijective, then the following statements are true:*

1. *$\Sigma_\diamond$ and $\Sigma_\circ$ are linearly isomorphic.*
2. *$\Sigma_\diamond$ is pointed if and only if $\Sigma_\circ$ is.*
3. *$(A, B, \diamond)$ is formally real if and only if $(A, B_1, \circ)$ is.*

Proof. That $\Sigma_\diamond$ and $\Sigma_\circ$ are linearly isomorphic has already been established in section 2: The algebras $(A, B, \diamond)$ and $(A, B_1, \circ)$ are linearly homomorphic if $L_\circ = F L_\diamond$, and F is the linear homomorphism. If F is bijective then these algebras are linearly isomorphic. Now each element of $\Sigma_\circ$ is the image of an element of $\Sigma_\diamond$ under F, that is $\Sigma_\circ = F\Sigma_\diamond$. Thus, these cones are linearly homomorphic (isomorphic if F is bijective).

To show $\circ$ is formally real if $\diamond$ is note that if for some $a_1, \dots, a_k$,

$$0 = a_1 \circ a_1 + \cdots + a_k \circ a_k = F(a_1 \diamond a_1 + \cdots + a_k \diamond a_k),$$

then since F is bijective we conclude that $\sum_i a_i \diamond a_i = 0$, and thus $a_i = 0$. Finally, under linear isomorphism the property that a cone is pointed is preserved. All the arguments above go in both directions since linear isomorphism is a symmetric relation: F^{-1} is a linear isomorphism from $\Sigma_\circ$ to $\Sigma_\diamond$. ☐

We now show that one can get linear isomorphism through function composition.

Lemma 15. *Let $\mathcal{F} = \{f_1, \dots, f_m\}$, and $F = \operatorname{Span} \mathcal{F}$, where each f_i is a real-valued function on a set Δ, and let $(F, S, \Sigma_{\mathcal{F}})$ be the induced squared functional system. Let Ω be a set and $H : \Omega \to \Delta$ be a surjective map. Define functions $g_i \colon \Omega \to \mathbb{R}$ by $g_i(y) = f_i\big(H(y)\big)$, $\mathcal{G} = \{g_1, \dots, g_m\}$, $G = \operatorname{Span} \mathcal{G}$, and let $(G, T, \Sigma_{\mathcal{G}})$ be their induced squared functional system. Then the algebras $(F, S, \Sigma_{\mathcal{F}})$ and $(G, T, \Sigma_{\mathcal{G}})$ are isomorphic, and the corresponding SOS cones $\Sigma_{\mathcal{F}}$ and $\Sigma_{\mathcal{G}}$ are also linearly isomorphic.*

Proof. Let $\{v_1(\cdot), \dots, v_n(\cdot)\}$ be a basis for $S = \operatorname{Span}\{f_i f_j\}$. Then the set of functions $u_i(\cdot) = v_i\big(H(\cdot)\big)$ is linearly independent and in fact, is a basis for $T = \operatorname{Span}\{g_i g_j\}$. To see this note that for a set of real numbers $\alpha_1, \dots, \alpha_n$, if we have

$$0 = \alpha_1 u_1(y) + \cdots + \alpha_n u_n(y) = \alpha_1 v_1\big(H(y)\big) + \cdots + \alpha_n v_n\big(H(y)\big) \quad \forall\, y \in \Omega,$$

then since $H(\cdot)$ is surjective, we must have that for all $x \in \Delta$: $\alpha_1 v_1(x) + \cdots + \alpha_n v_n(x) = 0$. This last equality implies $\alpha_i = 0$, and thus u_i are linearly independent. To show that the u_i form a basis for T, let $u_{n+1}(y) \in \operatorname{Span}\{g_i g_j\}$. Then there are real numbers β_{ij} such that $u_{n+1}(y) = \sum_{ij} \beta_{ij} g_i(y) g_j(y)$. Now consider the function $v_{n+1}(x) \stackrel{\text{def}}{=} \sum_{ij} \beta_{ij} f_i(x) f_j(x)$, defined on Δ. Then there are real numbers α_i such that for all $x \in \Delta$ we have, $v_{n+1}(x) = \sum_i \alpha_i v_i(x)$. Thus, for all $y \in \Omega$ we have:

$$\sum_i \alpha_i u_i(y) = \sum_i \alpha_i v_i\big(H(y)\big) = v_{n+1}\big(H(y)\big) = \sum_{ij} \beta_{ij} f_i\big(H(y)\big) f_j\big(H(y)\big) = \\ \sum_i \beta_{ij} g_i(y) g_j(y) = u_{n+1}(y),$$

proving that $u_i(y)$ form a basis. We have established an isomorphism (P, Q) between $(F, S, \Sigma_{\mathcal{F}})$ and $(G, T, \Sigma_{\mathcal{G}})$, defined by relations $P : f_i(x) \leftrightarrow g_i(y)$ mapping F to G, and $Q : v_i(x) \leftrightarrow u_i(y)$ mapping S to T. By Theorem 13, $\Sigma_{\mathcal{G}} = Q\big(\Sigma_{\mathcal{F}}\big)$, and the two SOS cones are linearly isomorphic. □

The following special case of this theorem will be used below to establish linear isomorphisms among several concrete and well-known functional systems and their cones of nonnegative functions.

Corollary 16. *Let $H : \Omega \to \Delta$, with $\Delta \subseteq \mathbb{R}$, be a surjective mapping. Then the cone*

$$\mathcal{P}_d[H] \stackrel{\text{def}}{=} \{p(x) \mid p(x) = p_0 + p_1 H(x) + \cdots + p_d H^d(x) \geq 0 \text{ for all } x \in \Omega\}$$

is linearly isomorphic to the cone of ordinary polynomials $p(t)$ where $p(t) \geq 0$ for all $t \in \Delta$, that is $\mathcal{P}_{\Delta,d}[t]$.

Note that here Ω can be any arbitrary set. For instance, Ω may be a subset of $\mathbb{R}^k$; it could be a (possibly infinite dimensional) functional space like a Hilbert or Banach space on which H is a real-valued functional; it could even be a collection of measurable sets, and H could be a measure. We should state that variants of Corollary 16 has appeared implicitly or explicitly in different forms. For instance, Karlin and Studden use isomorphism through composition to establish similarities between Chebyshev systems of functions defined on a closed interval $[a, b]$ and semi infinite interval $[a, \infty)$, see [15, Chapters V & VI]. Nesterov [29] also uses similar arguments to show SOS nature of various nonnegative polynomial spaces.

6.2 Weighted sums of squares cones

Using linear isomorphism we can extend the SD-representability results to *weighted* SOS (WSOS) cones.

Lemma 17. *Let $(A, B, \diamond)$ and $(B, C, \circ)$ be two algebras. For $w \in B$ define $\Sigma^w = w \circ \Sigma_\diamond$. Then Σ^w is an SOS cone. Furthermore, the cone Σ^w is linearly isomorphic to $\Sigma_\diamond$.*

Proof. Define the algebra $(A, C, \triangledown)$ by $a \triangledown b = w \circ (a \diamond b)$. Recalling our assumption that L_w is injective, it is clear that $w \circ \Sigma_\diamond = L_w \Sigma_\diamond = \Sigma_\triangledown$. Hence, Σ^w is an SOS cone. Also, $L_\circ(w)$ defines a bijective linear transformation from $\Sigma_\diamond$ to Σ^w proving their linear isomorphism. □

Now let $\mathbf{w} = (w_1, \ldots, w_r) \in B^r$, and let $\Sigma^{w_i} = w_i \circ_i \Sigma_{\diamond_i}$, where $\diamond_i : A \times A \to B$ and $\circ_i : B \times B \to C$, for some linear spaces A, B, and C. Define

$$\Sigma^{\mathbf{w}} = \Sigma^{w_1} + \cdots + \Sigma^{w_r}$$

where $+$ is the Minkowski sum of cones. Since a Minkowski sum of SOS cones is also an SOS cone, we have:

Lemma 18. *Every WSOS cone $\Sigma^{\mathbf{w}}$ is also an SOS cone.*

This result is a generalization of Nesterov's characterization of weighted squared functional systems in [29]. Nesterov showed that a WSOS functional system is an SOS with respect to another system that consists of square roots of the weight functions. Our development is more general in that, in the case of functional systems, it does not require the weights w_i be perfect squares with respect to $\diamond_i$.

6.3 Linear isomorphism in functional systems

We start this section by recalling the well-known Lukács-Markov Lemma about WSOS representation of univariate polynomials which are nonnegative over an interval. These results will be extended to more general settings in the following sections.

Proposition 19 (The Lukács-Markov Theorem ([26, 20]))**.** *Let $\mathcal{P}_{I,d}[t]$ be the cone of univariate polynomials of degree at most d which are nonnegative over the (possibly infinite) interval I, and let $\mathcal{P}_{2d}[t] = \mathcal{P}_{\mathbb{R},2d}[t]$. Then*

1. *If d is odd then $\mathcal{P}_{[a,b],d}[t] = (t-a)\mathcal{P}_{d-1}[t] + (b-t)\mathcal{P}_{d-1}[t]$ and $\mathcal{P}_{[a,\infty),d}[t] = \mathcal{P}_{d-1}[t] + (t-a)\mathcal{P}_{d-1}[t]$.*
2. *If d is even then $\mathcal{P}_{[a,b],d}[t] = \mathcal{P}_d[t] + (t-a)(b-t)\mathcal{P}_{d-2}[t]$ and $\mathcal{P}_{[a,\infty),d}[t] = \mathcal{P}_d[t] + (t-a)\mathcal{P}_{d-2}[t]$.*

Therefore, $\mathcal{P}_{[a,b],d}[t]$ and $\mathcal{P}_{[a,\infty),d}[t]$ are WSOS–and thus, SOS–cones.

Some of the following isomorphisms have already been noted in [36, 37] in slightly more restricted form. We include them here again to make the paper self-contained, and to prepare the ground for further generalizations to tensor product algebras.

We first consider nonnegative functions over an interval obtained by multiplying polynomials by a given function $g(\cdot)$.

Definition 20. *Let $\Delta \subseteq \mathbb{R}$, and $g\colon \Delta \to \mathbb{R}$ a function. We say that g* changes sign *at t_0 if $t_0 \in \operatorname{Int}\Delta$ and for some $\delta > 0$, if $g(t_1)g(t_2) < 0, \forall\, t_1 \in (t_0-\delta, t_0)$ and $\forall\, t_2 \in (t_0, t_0+\delta)$. (We do not make any assumptions on the sign of g, nor on its continuity at the point of a sign change.)*

Lemma 21. *Let $\Delta \subseteq \mathbb{R}$ be an interval, and $g : \Delta \to \mathbb{R}$. Assume g has only isolated roots, and that it changes sign at exactly k different points in Δ. Then the cone*

$$^{g}\mathcal{P}_{\Delta,d}[t] = \Big\{p(t) = \sum_i p_i t^i \mid g(t)p(t) \geq 0 \text{ for all } t \in \Delta\Big\}.$$

is linearly isomorphic to $\mathcal{P}_{\Delta,d-k}[t]$.

Proof. Let $p(t) = \sum_{i=1}^{d} p_i t^i$. Suppose g changes signs at the points $t_1 < \cdots < t_k$, then the sign of g agrees with the sign of either $p_1(t) = \prod_{i=1}^{k}(t-t_i)$, or of $-p_1(t)$ at every interior point of Δ, except for the zeros of g and the points t_i. Let us assume that the first case holds, the other case is similar. Then, since the signs of g and p also agree at every point except for the roots of g and p, p has a root at every root of p_1. Thus, $p(t) = p_1(t)q(t)$ for some degree $d-k$ polynomial q which is nonnegative on Δ. This gives a linear bijection between the elements p of $^{g}\mathcal{P}_{\Delta,d}[t]$ and q of $\mathcal{P}_{\Delta,d-k}[t]$. ☐

Consider the function space $V = \operatorname{Span}\{t^{-k}, \ldots, t^{-1}, 1, t, \ldots, t^d\}$ $(k \geq 1)$ defined on an interval $\Delta \subset \mathbb{R}$, and its cone of nonnegative functions $\mathcal{P}_{\Delta}^{-k,n}[t]$. In this case Δ cannot contain 0, therefore observing that t^{-k} does not change sign on Δ, from Lemma 21 we get the following result.

Corollary 22. *Define the cone*

$$\mathcal{P}_{\Delta,-k,d}[t] = \{p(t) = p_{-k}t^{-k} + \cdots + p_0 + \cdots p_d t^d \mid p(t) \geq 0 \text{ for all } t \in \Delta, \} = {}^{t^{-k}}\mathcal{P}_{\Delta,d+k}[t]$$

for the interval $\Delta \not\ni 0$. Then, $\mathcal{P}_{\Delta,-k,d}[t] \simeq \mathcal{P}_{\Delta,k+d}[t]$.

We now give several consequences of the preceding observations by applying them to concrete functional systems. The list below is only a partial list and many more functional systems can be added to it. Later, we extend these results to functions taking values in algebras. First let us introduce a notation. Let $U = \{u_1(t), \ldots, u_d(t)\}$. Then, $\mathcal{P}_{\Delta}^{U} = \{p(t) = \sum_i p_i u_i(t) \mid p(t) \geq 0 \text{ for all } t \in \Delta\}$. As usual, if Δ is not indicated, then $\Delta = \mathbb{R}$.

Nonnegative polynomials over a half-line: The Lukáacs-Markov Lemma (Lemma 19) shows that the cones $\mathcal{P}_{[a,\infty)}[t]$ and $\mathcal{P}_{[a,b)}[t]$ are WSOS (and thus SOS) cones. In fact, these cones are actually linearly isomorphic. In Corollary 16 set $H(t) = \frac{t}{1-t}$ and note that $p \in \mathcal{P}_{[0,\infty),d}[t]$ if and only if the polynomial $(1-t)^d p\left(\frac{t}{1-t}\right) \in \mathcal{P}_{[0,1),d}[t]$. Using Lemma 21 with $g(t) = (1-t)^d$ we obtain $\mathcal{P}_{[0,\infty),d}[t] \simeq \mathcal{P}_{[0,1),d}[t]$. Finally, $\mathcal{P}_{[a,b]}[t] \simeq \mathcal{P}_{[0,1)]}[t]$, and $\mathcal{P}_{[a,\infty)]}[t] \simeq \mathcal{P}_{[0,\infty]}[t]$ for any real numbers $a < b$.

Cosine polynomials: Let cos $= \{1, \cos(t), \dots, \cos(dt)\}$. The following classic and well-known identities can be obtained, for example, by applying the binomial theorem to the de Moivre formula:

$$\cos(kt) = \sum_{j=0}^{\lfloor k/2 \rfloor} \binom{k}{2j} \cos^{k-2j}(t)\left(\cos^2(t) - 1\right)^j, \tag{3}$$

$$\sin(kt) = \sin(t) \sum_{j=0}^{\lfloor (k-1)/2 \rfloor} \binom{k}{2j+1} \cos^{k-2j-1}(t)\left(\cos^2(t) - 1\right)^j. \tag{4}$$

From (3) it is clear that $\cos(kt) = T_k(\cos(t))$ where T_k are the *Chebyshev polynomials of the first kind*. Thus, setting $H(\cdot) = \cos(\cdot)$ in Corollary 16, and with a change of basis (from the standard monomial basis to $\{T_k\}$) we conclude that $\mathcal{P}_d^{\cos}[t] \simeq \mathcal{P}_d[\cos(t)] \simeq \mathcal{P}_{[-1,1],d}[t] \simeq \mathcal{P}_{[0,1],d}[t]$, where $\mathcal{P}_d^{\cos}[t] = \{p(t) = p_0 + p_1 \cos(t) + \cdots + p_d \cos(dt) \mid p(t) \geq 0 \text{ for all } t \in \mathbb{R}\}$. Indeed, $\mathcal{P}_{\Delta,d}^{\cos}[t] \simeq \mathcal{P}_{[0,1],d}[t]$ for any closed or semi-closed interval Δ.

Trigonometric polynomials: Consider the cone

$$\begin{aligned}
\mathcal{P}_{2d+1}^{\text{trig}} &= \left\{ r(t) = r_0 + \sum_{k=1}^{d} \left(r_{2k-1}\cos(kt) + r_{2k}\sin(kt)\right) \mid r(t) \geq 0 \text{ for all } t \in \mathbb{R} \right\} \\
&= \left\{ r(t) = r_0 + \sum_{k=1}^{d} \left(r_{2k-1}\cos(kt) + r_{2k}\sin(kt)\right) \mid r(t) \geq 0 \text{ for all } t \in (-\pi, \pi) \right\}
\end{aligned}$$

To transform a trigonometric polynomial $r(t) = r_0 + \sum_{k=1}^{d}\left(r_{2k-1}\cos(kt) + r_{2k}\sin(kt)\right)$ into ordinary polynomials we make the change of variables $t = 2\arctan(s)$. With this transformation we have

$$\sin(t) = \frac{2s}{1+s^2} \quad \text{and} \quad \cos(t) = \frac{1-s^2}{1+s^2}.$$

Using (3-4) we can write

$$r(t) = p_1\left(\tfrac{1-s^2}{1+s^2}\right) + \tfrac{2s}{1+s^2} p_2\left(\tfrac{1-s^2}{1+s^2}\right) \tag{5}$$

where p_1 and p_2 are ordinary polynomials of degree d, and $d-1$, respectively; $p_1(\cdot)$ is obtained from (3) and $p_2(\cdot)$ is obtained from (4). Multiplying by $(1+s^2)^d$ we see that

$$r(t) = (1+s^2)^{-d} p(s)$$

for some ordinary polynomial p. Substituting (3) and (4), the polynomial p can be expressed in the following basis:

$$\Big\{(1+s^2)^d, (1+s^2)^{d-1}(1-s^2), \dots, (1-s^2)^d\Big\} \cup \Big\{s(1+s^2)^{d-1}, s(1-s^2)^{d-2}(1-s^2), \dots, s(1-s^2)^{d-1}\Big\}$$

Since the function $\tan(t/2)$ maps $(-\pi, \pi)$ to $\mathbb{R}$, and noting that $(1+s^2)^{-d}$ is strictly positive, Corollary 16 and Lemma 21 imply that the cone of nonnegative trigonometric polynomials is isomorphic to the cone of nonnegative ordinary polynomials. More generally, If Δ contains the interval $(k\pi, (k+2)\pi)$ for any integer k, then again $\mathcal{P}^{trig}_{\Delta,2d+1}[t] \simeq \mathcal{P}_{2d}[t]$. If, on the other hand Δ is such that under the mapping $t \rightarrow \tan(t/2)$ is mapped to a closed, or a semi-closed interval, then $\mathcal{P}^{\text{trig}}_{\Delta,d}[t] \simeq \mathcal{P}_{[0,1],d}[t]$.

Exponential polynomials: Similar results hold for hyperbolic trigonometric polynomials, but the situation is slightly simpler. Given integers $k \le d$ define $\mathcal{P}_{-k,d}[\exp] \stackrel{\text{def}}{=} \Big\{e(t) = \sum_{j=-k}^{d} e_j \exp(jt) \mid e(t) \ge 0 \text{ for all } t \in \mathbb{R}\Big\}$. Then it is clear that replacing H with exp in Corollary 16, and using Lemma 21, we get that $\mathcal{P}_{-k,d}[\exp] \simeq \mathcal{P}_{[0,\infty),k+d+1}[t] \simeq \mathcal{P}_{[0,1],k+n+1}$. Again, when Δ is such that under the $t \rightarrow \exp(t)$, it is mapped to a closed, or semi-closed interval, for $\mathcal{P}_{\Delta,-k,d}[\exp] \simeq \mathcal{P}_{[0,1],k+d+1}[t]$.

Rational Chebyshev systems. Let $\alpha_1 < \alpha_2 < \cdots < \alpha_d$ be real numbers, and consider the linear functional space $C = \text{Span}\{\frac{1}{t-\alpha_1}, \dots, \frac{1}{t-\alpha_d}\}$. We assume that the common domain of these functions Δ is a (closed or semiclosed) interval that does not contain any of the α_i. Define $g(t) = \prod_{i=1}^{d}(t-\alpha_i)$, and note that with our assumptions g is either strictly negative or strictly positive on Δ. Therefore, the cone of nonnegative functions in C, $\mathcal{P}^C_\Delta \simeq \mathcal{P}_{\Delta,d-1}[t]$, the cone of degree $d-1$ polynomials nonnegative over Δ. Elements of Span(C) are sometimes called *Cauchy polynomials* since they are generated by the Cauchy kernel, see [15].

Note that in the squared functional system induced by $\mathcal{C} = \big\{\frac{1}{t-\alpha_i} \mid i = 1, \dots d\big\}$, the corresponding SOS cone $\Sigma_{\mathcal{C}}$ is a proper superset of $\mathcal{P}^C_\Delta$, since functions in the SOS cone are in $B = \text{Span}(\mathcal{B})$, where $\mathcal{B} = \big\{\frac{1}{t-\alpha_i}, \frac{1}{(t-\alpha_i)^2} \mid i = 1, \dots d\big\}$. However, a function $f(t) = \sum_i \frac{a_i}{t-\alpha_i} + \sum_i \frac{b_i}{(t-\alpha_i)^2}$ in B is nonnegative, if and only if the polynomial

$$p(t) = \sum_i a_i(t-\alpha_i)\prod_{j\neq i}(t-\alpha_j)^2 + \sum_i b_i \prod_{j\neq i}(t-\alpha_j)^2$$

is nonnegative. This is so, because the rational function $f(t)$ has $p(t)$ as its numerator, and $\prod_i (t-\alpha_i)^2$ as its denominator. On the other hand, the set

$$\bigcup_i \Big\{\big\{(t-\alpha_i)\prod_{j\neq i}(t-\alpha_j)^2\big\} \cup \big\{\prod_{j\neq i}(t-\alpha_j)^2\big\}\Big\}$$

is a basis for polynomials of degree $2d$. Thus, $\Sigma_{\mathcal{C}} \simeq \mathcal{P}_{\Delta,2d}[t]$. With appropriate modification to d this result can be extended to Chebyshev systems induced by $\frac{1}{(t-\alpha_i)^{n_i}}$ for any integers $0 \le n_i \le d_i$ and distinct real numbers α_i.

6.4 Nonnegative Cones in Polynomial Tensor Product Spaces

We now turn our attention to the tensor product of an arbitrary algebra $(A, B, \diamond)$ and univariate functional systems. Recall that we use the notation $\big(A_d[t], B_{2d}[t], \diamond\big)$ for the tensor product algebra of univariate polynomials and $(A, B, \diamond)$, and we define

$$\mathcal{P}_{\mathcal{K},\Delta,d}[t] = \{p(t) \in B_{2d}[t] \mid p(t) \in \mathcal{K} \text{ for all } t \in \Delta\}$$

where $\mathcal{K} \subseteq B$ is a proper cone.

We first present a generalization of the Lukács-Markov Theorem to symmetric matrix valued polynomials, that is the tensor product of univariate polynomials $(\mathbb{R}_d[t], \mathbb{R}_{2d}[t], \mathcal{P}_{2d}[t])$ and the symmetric Cracovian algebra $(\mathbb{M}_{m,n}, \mathbb{S}_m, \star)$.

Theorem 23. *For every $d \geq 1, m \geq 1$ the cones $\mathcal{P}_{\mathcal{P}_m,[a,b],d}[t]$ and $\mathcal{P}_{\mathcal{P}_m,[a,\infty),d}[t]$ are weighted SOS cones, specifically,*

$$\begin{aligned}
\mathcal{P}_{\mathcal{P}_m,[a,\infty),d}[t] &= \mathcal{P}_{\mathcal{P}_m,d-2}[t] + (t-a)\mathcal{P}_{\mathcal{P}_m,d-1}[t], && \textit{when } d \textit{ is odd} && (6)\\
\mathcal{P}_{\mathcal{P}_m,[a,\infty),d}[t] &= \mathcal{P}_{\mathcal{P}_m,d}[t] + (t-a)\mathcal{P}_{\mathcal{P}_m,d-1}[t], && \textit{when } d \textit{ is even} && (7)\\
\mathcal{P}_{\mathcal{P}_m,[a,b],d}[t] &= (t-a)\mathcal{P}_{\mathcal{P}_m,d-1}[t] + (b-t)\mathcal{P}_{\mathcal{P}_m,d-1}[t] && \textit{if } n \textit{ is odd,} && (8)\\
\mathcal{P}_{\mathcal{P}_m,[a,b],d}[t] &= (t-a)(b-t)\mathcal{P}_{\mathcal{P}_m,d-2}[t] + \mathcal{P}_{\mathcal{P}_m,d}[t] && \textit{if } n \textit{ is even.} && (9)
\end{aligned}$$

Proof. In all of equations (6) through (9) the fact that the right-hand-sides are included in the left-hand side is trivial. So we focus on showing that the left-hand-sides are included in the right-hand-sides.

First, let us consider (6) and (7). For simplicity we assume $(a, b) = (0, 1)$, the argument for general intervals is essentially identical. (The general results can also be obtained from the special cases by a change of variables.) We will use the notation $M^{\star 2} = MM^*$ for squaring a matrix M with respect to the Cracovian multiplication.

Suppose $P(t) \in \mathcal{P}_{\mathcal{P}_m,[0,\infty),d}[t]$, the cone of symmetric $m \times m$ polynomial matrix pencils which are positive semidefinite for all $t \in [0, \infty)$. Define the polynomial Q by $Q(t) = P(t^2)$. By assumption, $Q(t)$ is nonnegative everywhere, hence, by Youla's theorem (Proposition 9) $Q(t) = \sum_i Q_i^{\diamond 2}(t)$. Grouping the terms of the $Q_i(t)$ by the parity of their degrees we write $Q_i(t) = R_i(t^2) + tS_i(t^2)$, and since $Q(t)$ has no odd degree terms,

$$P(t^2) = Q(t) = \sum_i \big(R_i(t^2)\big)^{\star 2} + t^2 \sum_i \big(S_i(t^2)\big)^{\star 2}.$$

Taking $R(t) = \sum_i R_i^{\star 2}(t)$ and $S(t) = \sum_i S_i^{\star 2}(t)$ we have $P(t^2) = R(t^2) + t^2 S(t^2)$, implying $P(t) = R(t) + tS(t)$ with sum-of-squares matrices $R(t)$ and $S(t)$, as claimed. The bounds on the degrees of $R(t)$ and $S(t)$ follow from the fact that the degree of each $R_i(t)$ is at most d.

Next, suppose $P(t) \in \mathcal{P}_{\mathcal{P}_m,[0,1],d}[t]$, and consider the polynomial $Q(t) = (1+t)^d P\big(\frac{t}{1+t}\big)$: by assumption it is nonnegative for all $t \geq 0$, and so by the first claim of our theorem, $Q(t) =$

$R(t) + tS(t)$ identically for some $R(t) = \sum_i R_i^2(t)$ of degree at most d and $S(t) = \sum_i S_i^2(t)$ of degree at most $d-1$.

Observe that $P(t) = (1-t)^d Q(t/(1-t))$. If $d = 2k+1$, then this yields

$$\begin{aligned} P(t) &= (1-t)^{2k+1} \sum_i R_i^{\star 2}(t/(1-t)) + (1-t)^{2k+1} \frac{t}{1-t} \sum_i S_i^{\star 2}(t/(1-t)) \\ &= (1-t) \sum_i \big((1-t)^k R_i(t/(1-t))\big)^{\star 2} + t \sum_i \big((1-t)^k S_i(t/(1-t))\big)^{\star 2}. \end{aligned}$$

If $d = 2k$, then

$$\begin{aligned} P(t) &= (1-t)^{2k} \sum_i R_i^{\star 2}(t/(1-t)) + (1-t)^{2k} \frac{t}{1-t} \sum_i S_i^{\star 2}(t/(1-t)) \\ &= \sum_i \big((1-t)^k R_i(t/(1-t))\big)^{\star 2} + t(1-t) \sum_i \big((1-t)^{k-1} S_i(t/(1-t))\big)^{\star 2}. \end{aligned}$$

The degree bounds come from the degree bound in the first claim of the theorem. □

From here we can show that some cones in A and B-polynomials are either SOS or are SD-representable. This complements Theorem 12.

Theorem 24. *Let $(A, B, \diamond)$ be a formally real algebra with $\Sigma_\diamond$ its SOS cone. Define $\mathcal{P}_{\Sigma_\diamond^*,d}[t] = \{u(t) \in B_d[t] \mid u(t) \in \Sigma_\diamond^* \text{ for all } t \in \mathbb{R}\}$. Then the cone $\mathcal{P}_{\Sigma_\diamond^*,\Delta,d}[t]$ is SD-representable whenever $\Delta = [a.b]$, or $\Delta = [a, \infty)$.*

Proof. The function $u(t) \in B_d[t]$ is in $\Sigma_\diamond^*$ for all $t \in \Delta$ if and only if $\Lambda\big(u(t)\big) \succcurlyeq 0$ for all $t \in \Delta$. Since Λ is a linear operator, from Theorem 23, the Theorem follows. □

We can show that in the following tensor product functional systems, various cones of $\Sigma_\diamond^*$-nonnegative functions are SD-representable. The cosine and trigonometric

Corollary 25. *Let $(A, B, \diamond)$ be an algebra. Then the following cones are SD-representable, whenever Δ is either $\mathbb{R}$, or a closed, or a semi-closed interval. (Below, the coefficient p_i are all in B.)*

1. $\mathcal{P}_{\Sigma_\diamond^*,\Delta,d}[\cos] = \big\{p(t) = p_0 + p_1 \cos(t) + \cdots + p_d \cos(dt) \mid p(t) \in \Sigma_\diamond^* \text{ for all } t \in \Delta\big\}$,
2. $\mathcal{P}^{trig}_{\Sigma_\diamond^*\Delta,d}[t] = \big\{p(t) = p_0 + p_1 \sin(t) + p_2 \cos(t) + \cdots + p_{d-1} \sin(dt) + p_d \cos(t) \mid p(t) \in \Sigma_\diamond^* \text{ for all } t \in \Delta\big\}$,
3. $\mathcal{P}_{\Sigma_\diamond^*,\Delta,-k,d}[\exp] = \big\{p(t) = p_{-k} \exp(-kt) + \cdots + p_d \exp(dt) \mid p(t) \in \Sigma_\diamond^* \text{ for all } t \in \mathbb{R}\big\}$.
4. $\mathcal{P}^C_{\Sigma_\diamond^*\Delta,d}[t] = \big\{p(t) = \frac{p_1}{t-\alpha_1} + \cdots + \frac{p_d}{t-\alpha_d} \mid p(t) \in \Sigma_\diamond^* \text{ for all } t \in \Delta\big\}$, *where α_i are distinct real numbers, and Δ does not contain any of α_i,*

Again the results can be extended to other Chebyshev functional systems.

Corollary 26. *Let $(\mathbb{J}, \circ)$ be a formally real Jordan algebra ofwith $\mathcal{K}_{\mathbb{J}}$ its cone of squares. Then the following cones are SD-representable, whenever Δ is either $\mathbb{R}$, or a closed or a semiclosed interval, and each of coefficients p_i are in $\mathbb{J}$.*

1. $\mathcal{P}_{\mathcal{K}_{\mathbb{J}},\Delta,d}[\cos] = \big\{p(t) = p_0 + p_1\cos(t) + \cdots + p_d\cos(dt) \mid p(t) \in \mathcal{K}_{\mathbb{J}} \text{ for all } t \in \Delta\big\}$,
2. $\mathcal{P}^{trig}_{\mathcal{K}_{\mathbb{J}},d}[t] = \big\{p(t) = p_0 + p_1\sin(t) + p_2\cos(t) + \cdots + p_{d-1}\sin(dt) + p_d\cos(dt) \mid p(t) \in \mathcal{K}_{\mathbb{J}} \text{ for all } t \in \Delta\big\}$,
3. $\mathcal{P}_{\mathcal{K}_{\mathbb{J}},\Delta,-k,d}[\exp] = \big\{p(t) = p_{-k}\exp(-kt) + \cdots + p_d\exp(dt) \mid p(t) \in \mathcal{K}_{\mathbb{J}} \text{ for all } t \in \mathbb{R}\big\}$.
4. $\mathcal{P}^{C}_{\mathcal{K}_{\mathbb{J}},\Delta,d}[t] = \big\{p(t) = \frac{p_1}{t-\alpha_1} + \cdots + \frac{p_d}{t-\alpha_d} \mid p(t) \in \mathcal{K}_{\mathbb{J}} \text{ for all } t \in \Delta\big\}$, *where α_i are distinct real numbers, and Δ does not contain any of α_i,*

7 Further applications

In this section we sketch a few applications of the SOS functional systems in algebras. Of course, direct applications of Youla's Theorem in control theory are well-known and are some the earliest of such applications. The paper of Genin et al. [12], for instance, shows how SD-representability of polynomial matrix pencils can be used to design filters. Here we sketch a few other applications.

7.1 The Smallest Enclosing Ellipsoids of Curves

The smallest enclosing ball (SEB) and the smallest enclosing ellipsoid (SEE) (also often called minimum enclosing ball and ellipsoid) problems ask for the sphere or ellipsoid of minimum volume that contains a finite set of given points. These problems have been thoroughly studied, partly because of their important applications in machine learning, see for example [5, 6, 21]. Both of them admit simple second order cone programming (SOCP) and semidefinite programming (SDP) formulations: denoting the input points by $Y = \{y_1, \ldots, y_m\}$, the smallest sphere, with center x and radius r, containing Y is determined by the SOCP

$$\text{minimize } r \quad \text{subject to } \|x - y_i\| \leq r, i = 1, \ldots m,$$

while the ellipsoid of smallest volume is determined by the SDP

$$\text{maximize } (\det A)^{(1/n)} \quad \text{subject to } A \succcurlyeq 0; \quad \|x - Ay_i\| \leq 1, i = 1, \ldots m.$$

Equivalently, the objective function can be replaced by $\log\det A$, which results in a convex optimization problem with an SD-representable feasible set and a convex and SD-representable objective function, see [30] and [3].

We consider the following generalization of the SEB and SEE problems: given a closed parametric curve $p(t)$, $t \in [a, b]$ find the sphere or ellipsoid of minimum volume that contains all points of the curve. Replacing the finite set of constraints involving y_i by the constraints involving $p(t)$ we obtain optimization problems with a continuum of constraints over the cone $\mathcal{Q}_{n+1} = \{(x_0, \bar{x}) : x_0 \in \mathbb{R}, \bar{x} \in \mathbb{R}^n, \text{ and } x_0 \geq \|\bar{x}\|\}$. It is well known that $\mathcal{Q}_{n+1}$ is a symmetric cone; in fact it is the cone of squares of the Jordan algebra $(\mathbb{R}^{n+1}, \circ)$ with binary operation $(x_0, \bar{x}) \circ (y_0, \bar{y}) = (\langle \bar{x}, \bar{y} \rangle + x_0 y_0, x_0 \bar{y} + y_0 \bar{x})$. Now, if $p(t)$ is in any of the classes of functions discussed in Section 6.4, then the set $\{p \mid p(t) \in \mathcal{Q}\ \forall t \in \Delta\}$ is SD-representable for appropriate intervals Δ.

Example 3. Figure 1 shows a parametric curve

$$(x(t), y(t)) = \sum_{k=1}^{n} p_{2k-1} \sin(kt) + p_{2k} \cos(kt), \quad t \in [0, 2\pi]$$

with $n = 3$ and coefficient vectors $p_1, \ldots, p_6$ that are independent random vectors chosen uniformly from the interval $[-1, 1]^2$. Since p is a trigonometric polynomial with coefficients in $\mathbb{R}^{n+1}$, by the preceding discussion the condition $\|x - Ap(t)\| \leq 1$ is SD-representable. The minimal circle and minimal ellipse containing $\{p(t) \mid t \in [0, 2\pi]\}$ is shown in Figure 1. $\square$

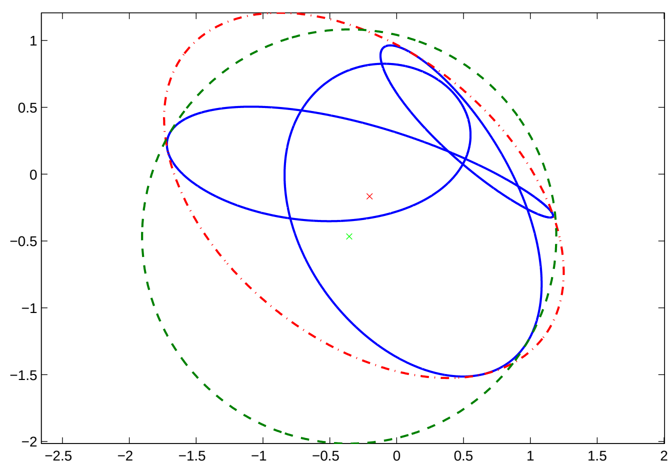

Figure 1: A trigonometric polynomial curve (blue, solid) together with its smallest enclosing circle (green, dashed) and its smallest enclosing ellipsoid (red, dot-dashed).

Note that this approach is very flexible and can be easily extended in several directions without difficulty. Clearly the dimension of the problem can be any number. We can also have a number of curves, and not just one, and can seek the ball or ellipsoid with the smallest volume containing all of them. All these problems can easily, and in an obvious manner be solved with our approach, as long as the parametric formula of each curve is from a linear functional system where nonnegative functions are SOS, or at least SD-representable. Also, our approach significantly generalizes the problem of finding the smallest area circle or ellipse containing a number of ellipses in the two dimensional plane; this problem is studied, for example, by Vandenberghe and Boyd [41, 4].

7.2 Eigenvalue optimization of polynomial matrix pencils

Let $\lambda_{[k]}(A)$ be the k^{th} largest eigenvalue of the symmetric (or complex Hermitian) matrix A. It is well-known that the functions $\sum_{i=1}^k \lambda_{[i]}(A)$, and $\sum_{i=1}^k m_i\lambda_{[i]}(A)$—where $m_1 \geq m_2 \geq \cdots \geq m_k$—are convex and SD-representable, see for instance, [2], [30] and [3]. In fact, for any symmetric, convex, and SD-representable function $s(y_1, \ldots, y_n)$, the function $s\big(\lambda_{[1]}(A), \ldots, \lambda_{[n]}(A)\big)$ is SD-representable, [3].

Now let $A = A(x,t) = \sum_i A_i(x)t^i$, where each $A_i(x)$ is a symmetric (or complex Hermitian) matrix-valued affine function of $x = (x_1, \ldots, x_r)$. Then the optimization problem

$$\min_{x,t} s\big(\lambda_{[1]}(A(x,t)), \ldots, \lambda_{[n]}(A(x,t))\big)$$

may be formulated as a semidefinite program, whenever s is a symmetric, convex and SD-representable. This generalizes Haeberly and Overton's work on linear matrix pencils, [13].

More generally, let $p(x,t) = \sum_i p_i(x)t^i$ where each p_i is an affine function whose range is a Euclidean Jordan algebra $\mathbb{J}$. Then the following optimization problem is solvable by semidefinite programming:

$$\min_{x,t} s\big(\lambda_{[1]}(p(x,t), \ldots, \lambda_{[r]}(p(x,t)\big),$$

where s is a symmetric, convex and SD-representable function, and $\lambda_{[k]}(a)$ is the k^{th} largest Jordan algebraic eigenvalue of a.

7.3 Shape-constrained approximation of functions

Let $\mathcal{F} = \{f_1(\cdot), \ldots, f_m(\cdot)\}$ be a set of linearly independent functions defined on a set Δ, and let $(F(x), S(x), \Sigma_{\mathcal{F}})$ be its corresponding induced SFS. Let $(A, B, \diamond)$ be an algebra and let $h(x) \in S(x) \otimes B$, that is $h(x) = \sum_i h_{ij} f_i(x) f_j(x)$ and each $h_{ij} \in B$. We wish to find a function $g(x) = \sum_{ij} g_{ij} f_i(x) f_j(x)$ with some constraints, and closest to $h(x)$. The constraints could be, for instance, that $g(x)$ be convex, or be in a proper cone $\mathcal{K}$. The "closeness" criteria could be, for instance, $d(h,g) = \int_\Delta \|h(x) - g(x)\| dx$ or $d(h,g) = \int_\Delta \|h(x) - g(x)\|^2 dx$, for some appropriately chosen norm $\|\cdot\|$ on B. The optimization problem, in general can be formulated as

$$\begin{array}{ll} \min & d(h,g) \\ \text{s.t.} & \mathcal{L}_i(g) \succcurlyeq_{\mathcal{K}} 0 \quad \text{for } i = 1, \ldots, k \end{array}$$

where each $\mathcal{L}_i$ is a linear operator on $S(x) \otimes B$.

Such a problem in general may be difficult. However, in certain circumstances, depending on the cone $\mathcal{K}$, and functions $f_i(x)$, the problem may become tractable. For instance, if

$\mathcal{K} = \Sigma_\diamond^*$ for some algebra $(A, B, \diamond)$, and the SFS is the one induced by univariate polynomials, then the constraint $\mathcal{L}_i(g) \succcurlyeq_{\Sigma_\diamond^*} 0$ is equivalent to the inequality $\Lambda_\diamond\Big(\mathcal{L}_i\big(g(t)\big)\Big) \succcurlyeq 0$; by Youla's theorem the set of functions $g(\cdot)$ satisfying such inequalities is SD-representable.

7.4 Polynomial Programming and Combinatorial Optimization

Polynomial multiplication with respect to a (fixed) polynomial ideal is a bilinear operation. The application of Theorem 3 to the space of polynomials of a given degree yields simplified SDP relaxations and hierarchies to polynomial programs with equality constraints such as those considered in [25]. Consider the POP

$$\text{maximize } f(x) \qquad \text{subject to } h_i(x) = 0, \quad i = 1, \dots m, \tag{10}$$

where f and each h_i are given n-variate polynomials. An equivalent formulation is

$$\text{minimize } c \;\text{ subject to } f(x) - c \geq 0 \;\forall x \in \{x\colon h_i(x) = 0, i = 1, \dots m\}.$$

This problem is NP-hard, but the following restriction may be tractable:

$$\text{minimize } c \;\text{ subject to } f(x) - c \in \Sigma, \tag{11}$$

where Σ is the sum of squares cone with respect to polynomial multiplication modulo the ideal I generated by $\{h_i\}$. Depending on $\{h_i\}$, the bottleneck in the solution of (11) may be the computation of Λ, which involves finding a Gröbner basis of I, and a large number of multivariate polynomial divisions.

This method can be very effective when h_i form a Gröbner basis, and division modulo I is simple, as in the following example.

Example 4. Consider a graph $G = (V, E)$ with $|V| = n$ vertices and $|E| = m$ edges. Letting $f(x) = \sum_{i=1}^n x_i$ and $h_{ij}(x) = x_i x_j$, $ij \in E$, the solution to POP (10) is the stability number of G. A sequence of semidefinite relaxations to the stable set problem is obtained if we replace (10) by (11), and constrain $f - c$ to be a sum of squares of polynomials of a given total degree. (Higher degrees give tighter relaxations.) Theorem 2 gives the dual SDP's of these relaxations, which are identical to Laurent's simplified characterization [25] of the Lasserre hierarchy [22] applied to the stable set problem. This also serves as a new, simpler, proof of [25, Lemma 20]. □

8 Conclusion and Further research

By building an abstract algebraic setting for the notion of sum-of-squares, and showing that such sets are SD-representable, we have demonstrated that rich classes of functional systems may be SOS. Connections to functional analysis, and to differential geometry may yield important applications in a variety of fields including shape-constrained statistical learning, optimal geometric design, and optimal control theory.

In statistical learning, one may be interested in multivariate learning of functions or distributions, which must follow some additional constraints. For instance, suppose we are interested in a function $F\colon \Delta \to \mathbb{R}^m$ from a set of possibly noisy observations. Note that here the response variable $y = F(x)$ itself is a vector. In addition suppose that some linear transformations $\mathcal{L}_i\big(F(x)\big)$ are restricted to be in a set $\mathcal{C}$ for all $x \in \Delta$. In some cases, it may be possible to approximate the condition $F(x) \in \mathcal{C}$ by one or more Σ_i or Σ_i^* nonnegativity constraints, for some SOS cones Σ_i. In these cases, such conditions may be represented by semidefinite inequalities.

In geometric optimization, we may be interested in designing paths or surfaces, parametrically represented by $\big(f_1(t), \ldots, f_n(t)\big)$, where the parameter t may consist of a single or multiple variables. In addition one may, for instance, have restrictions on the curvature, torsion, or higher order curvatures of the path or the surface. Such conditions often may be expressed in the form of nonnegativity condition on some functionals. For instance, the problem of designing a path that has to go through a sequence of points, while keeping its curvature below a certain threshold, is a problem that may be formulated using SOS inequalities over functional systems.

Our approach gives a unifying framework to formulate and solve these, and many other shape-restricted problems, either exactly or approximately.

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming Series B **95** (2003), 3–51.

[2] Farid Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. on Optim. **5** (1995), no. 1, 13–51.

[3] Aharon Ben-Tal and Arkadi Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[4] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[5] Christopher J. C. Burges, *A tutorial on support vector machines for pattern recognition*, Data Mining and Knowledge Discovery **2** (1998), no. 2, 121–167.

[6] Olivier Chapelle, Vladimir Vapnik, Olivier Bousquet, and Sayan Mukherjee, *Choosing multiple parameters for support vector machines*, Machine Learning **46** (2002), no. 1-3, 131–159.

[7] Keith Conrad, *Complexification*, Available at `http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/complexification.pdf`.

[8] ______, *Tensor Products*, Available at `http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/tensorprod.pdf`.

[9] ______, *Tensor Products, II*, Available at `http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/tensorprod2.pdf`.

[10] Bogdan Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications*, Springer, 2010.

[11] J. Faraut and A. Korányi, *Analysis on symmetric cones*, Oxford University Press, Oxford, UK, 1994.

[12] Y. Genin, Y. Hachez, Y. Nesterov, R. Stefan, P. Van Dooren, and S. Xu, *Positivity and Linear Matrix Inequalities*, Eur. J. Control **8** (2002), 275–298.

[13] J. P. Haeberley and M. L. Overton, *A Hybrid Algorithm for Optimizing Eigenvalues of Symmetric Definite Pencils*, SIAM J. Matr. Anal. Appl. **15** (1994), 1141–1156.

[14] N. Jacobson, *Basic algebra II*, Dover Publications, 2009, Republication of the original text in 1989.

[15] S. Karlin and W. J. Studden, *Tchebycheff Systems, with Applications in Analysis and Statistics*, Wiley Interscience Publishers, 1966.

[16] Jerzy Kociński, *Cracovian algebra*, Nova Science, 2004.

[17] M. Koecher, *The Minnesota Notes on Jordan Algebras and Their Applications*, Springer-Verlag, 1999, Edited by Kreig, Aloys and Walcher, Sebastian based on Lectures given in The University of Minnesota, 1960.

[18] M. Kojima and M. U. Muramats, *An extension of sums of squares relaxations to polynomial optimization problems over symmetric cones*, Mathematical Programming **110** (2007), no. 2, 315–336.

[19] Masakazu Kojima, *Sums of squares relaxations of polynomial semidefinite programs*, Tech. Report B-397, Dept. of Math and Comp. Sci, Tokyo Institute of Technology, November 2003.

[20] M.G. Krein and A. A. Nudel'man, *The Markov Moment Problem and Extremal Problems*, Translations of Mathematical Monographs, vol. 50, American Mathematical Socisty, 1973.

[21] Piyush Kumar, MitchelJoseph S. B. l, and E. Alper Yıldırım, *Computing core-sets and approximate smallest enclosing hyperspheres in high dimensions*, Proceedings of the Fifth Workshop on Algorithm Engineering and Experiments (Richard E. Ladner, ed.), SIAM, Baltimore, MD, 2003, pp. 45–55.

[22] Jean B. Lasserre, *An explicit exact SDP relaxation for nonlinear 0-1 programs*, Integer Programming and Combinatorial Optimization (Karen Aardal and A. M. H. Gerards, eds.), Lecture Notes in Computer Science, no. 2081, Springer, 2001, pp. 293–303.

[23] ———, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization **11** (2001), no. 3, 796–817.

[24] Jean-Bernard Lasserre, *Moments, Positive Polynomials and their Applications*, Imperial College Press Optimization Series, Imperial College of London, 2009.

[25] Monique Laurent, *A comparison of the Sherali–Adams, Lovász–Schrijver, and Lasserre relaxations for 0-1 programming*, Mathematics of Operations Research **28** (2003), no. 3, 470–496.

[26] Franz Lukács, *Verschärfung der ersten Mittelwersatzes der Integralrechnung für rationale Polynome*, Mathematische Zeitschrift **2** (1918), 229–305.

[27] Alessandro Magnani, Sanjay Lall, and Stephen Boyd, *Tractable fitting with convex polynomials via sum-of-squares*, 44th IEEE Conference on Decision and Control and 2005 European Control Conference. CDC–ECC '05. (Seville, Spain), December 2005, pp. 1672–1677.

[28] Murray Marshall, *Positive Polynomials and Sums of Squares*, Mathematical Surveys and Monographs, American Mathematical Society, 2008.

[29] Yu. Nesterov, *Squared functional systems and optimization problems*, High Performance Optimization (H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds.), Appl. Optim., Kluwer Acad. Publ., Dordrecht, 2000, pp. 405–440.

[30] Yu Nesterov and Arkadi Nemirovski, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.

[31] Dávid Papp and Farid Alizadeh, *Linear and second order cone programming approaches to statistical estimation problems*, Tech. Report RRR 13–2008, Rutgers Center for Operations Research (RUTCOR), Rutgers University, September 2008.

[32] ———, *Multivariate Arrival Rate Estimation by Sums-Of-Squares Polynomial Splines and Semidefinite Programming*, Tech. Report RRR 5–2011, Rutgers Center for Operations Research (RUTCOR), Rutgers University, April 2011.

[33] Pablo A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Mathematical Programming **96** (2003), no. 2, 293–320.

[34] George Pólya and Gabor Szegö, *Problems and theorems in analysis*, vol. 2, Springer, 1976.

[35] Alexander Pretsel and Charles N. Delzell, *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*, Springer Monographs in Mathematics, Springer, 2001.

[36] Gábor Rudolf, Nilay Noyan, Dávid Papp, and Farid Alizadeh, *Bilinearity Rank of the Cone of Positive Polynomials and Related Cones*, Tech. Report RRR 7-2009, Rutgers Center for Operations Research (RUTCOR), March 2009.

[37] ______, *Bilinearity Rank of the Cone of Positive Polynomials and Related Cones*, Mathematical Programming (2011).

[38] Igore R. Shafarevich, *Basic Notions of Algebra*, Encyclopedia of Mathematical Sciences, Vol. 11, Springer, 2005, Translated from Russian by M. Reids.

[39] Dirk J. Struik, *Lectures on classical differential geometry*, Dover, New York, 1988.

[40] H. van Maaren, L. van Norden, and M. J. H. Heule, *Sums of squares based approximation algorithms for MAX-SAT*, Discrete Applied Mathematics **156** (2008), no. 10, 1754–1779.

[41] Lieven Vandenberghe and Stephen Boyd, *Semidefinite programming*, SIAM Review **38** (1996), no. 1, 49–95.

[42] D. C. Youla, *On the factorization of rational matrices*, IRE Transactions on Information Theory (1961), 172–189.

Notation

Let $(A, B, \diamond)$ be an algebra, $\mathcal{F} = \{f_1(x), \ldots, f_n(x)\}$ where each $f_i\colon \Delta \to \mathbb{R}$, $(F(x), S(x), \Sigma_{\mathcal{F}})$ its squared functional system, $\mathcal{K} \subseteq B$ a convex cone, and Δ is either an interval or a some arbitrary set, depending on the context. We use the following $\mathcal{K}$-nonnegative cones:

1. $\mathcal{P}^{\mathcal{F}}_{\mathcal{K},B} = \big\{f(x) = \sum_i a_i, f_i(x_i) \mid a_i \in B, \text{ and } f(x) \in \mathcal{K} \text{ for all } x \in \Delta\big\}$. If B is understood from the context (e.g $B = \operatorname{Span}(\mathcal{K})$), then we may drop it and simply write $\mathcal{P}^{\mathcal{F}}_{\mathcal{K}}$.

   If $B = \mathbb{R}$ and $\mathcal{K} = \mathbb{R}_+$, the set of nonnegative real numbers, then we write $\mathcal{P}^{\mathcal{F}}$.

2. In the univariate case, if $\mathcal{U} = \{u_1(t), \ldots, u_n(t)\}$, then the cone of $\mathcal{K}$ nonnegative functions generated by $\mathcal{U}$ and the linear space B is

$$\mathcal{P}^{\mathcal{U}}_{\mathcal{K},B,\Delta} = \big\{p(t) = \sum_i a_i u_i(t) \mid a_i \in B, \text{ and for all } t \in \Delta\colon p(t) \in \mathcal{K}\big\}$$

In particular for $B = \mathbb{R}$, and $\mathcal{K}$ nonnegative real numbers, we write $\mathcal{P}_{\Delta}^{\mathcal{U}}$.

3. If the SFS is the algebra of multivariate polynomials of degree at most d, and $\Delta \subseteq \mathbb{R}^k$, we write

$$\mathcal{P}_{\mathcal{K},B,d,\Delta}[t_1,\ldots,t_k] = \Big\{p(t_1,\ldots,t_k) = \sum_{i_1,\ldots,i_k} a_{i_1,\ldots,i_k} t_1^{i_1}\cdots, t_k^{i_k} \mid$$
$$a_{i_1,\ldots,i_k} \in B, \text{ and } p(t_1,\ldots,t_k) \in \mathcal{K} \text{ for all } (t_1,\ldots,t_k) \in \Delta\Big\}.$$

   Again, we drop B if it is understood from the context.

   If $\Delta = \mathbb{R}^k$ we simply write $\mathcal{P}_{\mathcal{K},d}[t_1,\ldots,t_k]$. Also for univariate case we write $\mathcal{P}_{\mathcal{K},d,\Delta}[t]$, and $\mathcal{P}_{\mathcal{K},d}[t]$, if $\Delta = \mathbb{R}$.

4. If the functional system is mapped through composition with $H\colon \Omega \to \Delta$ to another system $(G(y), T(y), \Sigma_{\mathcal{G}})$ then we may either write $\mathcal{P}_{\mathcal{K},\Delta}^{\mathcal{G}}$ or $\mathcal{P}_{\mathcal{K},\Omega}^{\mathcal{F}(H)}$. In case the functional system is the space of univariate polynomials, we write $\mathcal{P}_{\mathcal{K},d,\Omega}[H]$.

5. Let $g(\cdot)$ be a real-valued function on Δ. Consider the set of functions $f(\cdot)$ defined on Δ such that $g(x)f(x) \in \mathcal{P}_{\mathcal{K},B}^{\mathcal{F}}$. Then the set of such functions f is denoted by ${}^{g}\mathcal{P}_{\mathcal{K},B}^{\mathcal{F}}$. In particular, note that ${}^{g}\mathcal{P}_{\mathcal{K},B,d,\Delta}[t_1,\ldots,t_k]$, ${}^{g}\mathcal{P}_{\mathcal{K},d,\Delta}[t]$ are the cone of functions $f(\cdot)$ where $g(\cdot)f(\cdot) \in \mathcal{P}_{\mathcal{K},B,\Delta,d}[t_1,\ldots,t_k]$, and $g(\cdot)f(\cdot) \in \mathcal{P}_{\mathcal{K},d}[t]$, respectively.

For SOS cones we use the following notation:

1. The SOS cone of $(A, B, \diamond)$ is $\Sigma_{\diamond}$.

2. The SOS cone of the squared functional system induced by $\mathcal{F}$ is $\Sigma_{\mathcal{F}}$.

3. If the SFS is the set of all real multivariate polynomials of degree at most d then we write $\Sigma_d[t_1,\ldots,t_n]$. If the coefficients and t can be complex, and the multiplication of complex numbers are defined by $z_1 \star z_2 = z_1\overline{z_2}$, we write $\Sigma_{\star,\mathbb{C},d}[t_1,\ldots,t_k]$, and for univariate polynomials $\Sigma_{\star,\mathbb{C},d}[t]$.

4. For the two algebras $(A, B, \diamond)$ and $(B, C, \circ)$, if $w \in B$, then $\Sigma^w = w \circ \Sigma_{\diamond}$. In general, if $\mathbf{w} \in B^r$, and $(B, C, \circ_i)$ are a sequence of algebras, then $\Sigma^{\mathbf{w}} = \sum_i w_i \circ_i \Sigma_{\diamond}$, where the sum here is the Minkowski sum of cones.

5. For the tensor product of $(A_1, B_1, \diamond_1)$ and $(A_2, B_2, \diamond_2)$, that is $(A_1 \otimes A_2, B_1 \otimes B_2, \diamond)$, we may write for the SOS cone either $\Sigma_{\diamond}$ or $\Sigma_{\diamond_1\otimes\diamond_2}$.

6. For the tensor product of $(A, B, \diamond)$ and the SFS (F, S) induced by $\mathcal{F}$ we write $\Sigma_{\diamond\otimes\mathcal{F}}$. If the SFS is the space of multivariate polynomials of degree at most d we write $\Sigma_{\diamond\otimes d}[t_1,\ldots,t_k]$ and for univariate polynomials of degree at most d we write $\Sigma_{\diamond\otimes d}[t]$.