



# Sum of squares generalizations for conic sets

Lea Kapelevich<sup>1</sup> · Chris Coey<sup>1</sup> · Juan Pablo Vielma<sup>2</sup>

Received: 11 December 2020 / Accepted: 8 May 2022

© The Author(s) 2022

## Abstract

Polynomial nonnegativity constraints can often be handled using the *sum of squares* condition. This can be efficiently enforced using semidefinite programming formulations, or as more recently proposed by Papp and Yildiz (Papp D in SIAM J O 29: 822–851, 2019), using the sum of squares cone directly in an interior point algorithm. Beyond nonnegativity, more complicated polynomial constraints (in particular, generalizations of the positive semidefinite, second order and  $\ell_1$ -norm cones) can also be modeled through structured sum of squares programs. We take a different approach and propose using more specialized cones instead. This can result in lower dimensional formulations, more efficient oracles for interior point methods, or self-concordant barriers with smaller parameters.

**Keywords** Polynomial optimization · Sum of squares · Interior point · Non-symmetric conic optimization

**Mathematics Subject Classification** 90-08 · 90C25 · 90C51

---

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

---

✉ Lea Kapelevich  
lkap@mit.edu

Chris Coey  
coey@mit.edu

Juan Pablo Vielma  
jvielma@google.com; jvielma@mit.edu

<sup>1</sup> Operations Research Center, MIT, Cambridge, US

<sup>2</sup> Google Research and Sloan School of Management, MIT, Cambridge, US

# 1 Introduction

The *sum of squares* (SOS) condition is commonly used as a tractable restriction of polynomial nonnegativity. While SOS programs have traditionally been formulated and solved using semidefinite programming (SDP), Papp and Yildiz [16] recently demonstrated the effectiveness of a nonsymmetric interior point algorithm in solving SOS programs without SDP formulations. In this note, we focus on structured SOS constraints that can be modeled using more specialized cones. We describe and give barrier functions for three related cones useful for modeling functions of dense polynomials, which we hope will become useful modeling primitives.

The first is the cone of *SOS matrices*, which was described by Coey et al. [4, Section 5.7] without derivation. Characterizations of univariate SOS matrix cones in the context of optimization algorithms have previously been given by Genin et al. [6, Section 6]. However, their use of monomial or Chebyshev bases complicates computations of oracles in an interior point algorithm [16, Section 3.1] and prevents effective generalizations to the multivariate case. The second is an *SOS  $\ell_2$ -norm* (SOS-L2) cone, which can be used to certify pointwise membership in the second order cone for a vector with polynomial components. The third is an *SOS  $\ell_1$ -norm* (SOS-L1) cone, which can be used to certify pointwise membership in the epigraph set of the  $\ell_1$ -norm function. Although it is straightforward to use SOS representations to approximate these sets, such formulations introduce cones of higher dimension than the constrained polynomial vector. We suggest new barriers, with lower barrier parameters than SOS formulations allow.

In what follows, we use  $\mathbb{S}^m$ ,  $\mathbb{S}_+^m$ , and  $\mathbb{S}_{++}^m$  to represent the symmetric, positive semidefinite and positive definite matrices respectively with side dimension  $m$ . For sets,  $\text{cl}$  denotes the closure and  $\text{int}$  denotes the interior.  $\llbracket a..b \rrbracket$  are the integers in the interval  $[a, b]$ .  $|A|$  denotes the dimension of a set  $A$ , and  $\text{sd}(m) = |\mathbb{S}^m| = m(m+1)/2$ . We use  $\langle \cdot, \cdot \rangle_A$  for the inner product on  $A$ . For a linear operator  $M : A \rightarrow B$ , the adjoint  $M^* : B \rightarrow A$  is the unique operator satisfying  $\langle x, My \rangle_A = \langle y, M^*x \rangle_B$  for all  $x \in A$  and  $y \in B$ .  $\mathbf{I}_m$  is the identity in  $\mathbb{R}^{m \times m}$ .  $\otimes_K : \mathbb{R}^{a_1 \times a_2} \times \mathbb{R}^{b_1 \times b_2} \rightarrow \mathbb{R}^{a_1 b_1 \times a_2 b_2}$  is the usual Kronecker product.  $\text{diag}$  returns the diagonal elements of a matrix and  $\text{Diag}$  maps a vector to a matrix with the vector on the diagonal. All vectors, matrices, and higher order tensors are written in bold font.  $s_i$  is the  $i$ th element of a vector  $\mathbf{s}$  and  $\mathbf{s}_{i \in \llbracket 1..N \rrbracket}$  is the set  $\{\mathbf{s}_1, \dots, \mathbf{s}_N\}$ . If  $A$  is a vector space then  $A^n$  is the Cartesian product of  $n$  spaces  $A$ .  $\mathbb{R}[\mathbf{x}]_{n,d}$  is the ring of polynomials in the variables  $\mathbf{x} = (x_1, \dots, x_n)$  with maximum degree  $d$ . Following Papp and Yildiz [16], we use  $L = \binom{n+d}{n}$  and  $U = \binom{n+2d}{n}$  to denote the dimensions of  $\mathbb{R}[\mathbf{x}]_{n,d}$  and  $\mathbb{R}[\mathbf{x}]_{n,2d}$  respectively, when  $n$  and  $d$  are given in the surrounding context.

## 1.1 The SOS polynomials cone and generic interior point algorithms

A polynomial  $p(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}$  is SOS if it can be expressed in the form  $p(\mathbf{x}) = \sum_{i \in \llbracket 1..N \rrbracket} q_i(\mathbf{x})^2$  for some  $N \in \mathbb{N}$  and  $q_i \in \llbracket 1..N \rrbracket(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,d}$ . We denote the set of SOS polynomials in  $\mathbb{R}[\mathbf{x}]_{n,2d}$  by  $K_{\text{SOS}}$ , which is a proper cone in  $\mathbb{R}[\mathbf{x}]_{n,2d}$  [13].

We also say that  $\mathbf{s} \in K_{\text{SOS}}$  for  $\mathbf{s} \in \mathbb{R}^U$  if  $\mathbf{s}$  represents a vector of coefficients of an SOS polynomial under a given basis. We use such *vectorized* definitions interchangeably with functional definitions of polynomial cones. To construct a vectorized definition for  $K_{\text{SOS}}$ , suppose we have a fixed basis for  $\mathbb{R}[\mathbf{x}]_{n,2d}$ , and let  $p_{i \in [1..L]}(\mathbf{x})$  be basis polynomials for  $\mathbb{R}[\mathbf{x}]_{n,d}$ . Let  $\lambda : [1..L]^2 \rightarrow \mathbb{R}^U$  be a function such that  $\lambda(i, j)$  returns the vector of coefficients of the polynomial  $p_i(\mathbf{x})p_j(\mathbf{x})$  using the fixed basis for  $\mathbb{R}[\mathbf{x}]_{n,2d}$ . Define the *lifting operator*  $\Lambda : \mathbb{R}^U \rightarrow \mathbb{S}^L$ , introduced by Nesterov [13], as:

$$\Lambda(\mathbf{s})_{i,j} = \langle \lambda(i, j), \mathbf{s} \rangle_{\mathbb{R}^U} \quad \forall i, j \in [1..L], \quad (1.1)$$

where  $\Lambda(\mathbf{s})_{ij}$  is a component in row  $i$  and column  $j$ . Now the cones  $K_{\text{SOS}}$  and  $K_{\text{SOS}}^*$  admit the characterization [13, Theorem 7.1]:

$$K_{\text{SOS}} = \{\mathbf{s} \in \mathbb{R}^U : \exists \mathbf{S} \in \mathbb{S}_+^L, \mathbf{s} = \Lambda^*(\mathbf{S})\}, \quad K_{\text{SOS}}^* = \{\mathbf{s} \in \mathbb{R}^U : \Lambda(\mathbf{s}) \in \mathbb{S}_+^L\}. \quad (1.2)$$

Equation (1.2) shows that the dual cone  $K_{\text{SOS}}^*$  is an inverse linear image of the positive semidefinite (PSD) cone, and therefore has an efficiently computable *logarithmically homogeneous self-concordant barrier* (LHSCB) (see [14, Definitions 2.3.1, 2.3.2]). In particular, by linearity of  $\Lambda$ , the function  $\mathbf{s} \mapsto -\log \det(\Lambda(\mathbf{s}))$  is an LHSCB for  $K_{\text{SOS}}^*$  [14, Proposition 5.1.1] with parameter  $L$  (an  $L$ -LHSCB for short). This makes it possible to solve optimization problems over  $K_{\text{SOS}}$  or  $K_{\text{SOS}}^*$  with a *generic* primal-dual interior point algorithm in polynomial time [17].<sup>1</sup>

In a generic primal-dual interior point algorithm, very few oracles are needed for each cone in the optimization problem. For example, the algorithm used by Coey et al. [4] only requires a membership check, an initial interior point, and evaluations of derivatives of an LHSCB for each cone *or* its dual. Therefore, there is no particular advantage to favoring either  $K_{\text{SOS}}$  or  $K_{\text{SOS}}^*$  formulations. Optimizing over  $K_{\text{SOS}}$  (or  $K_{\text{SOS}}^*$ ) directly instead of building SDP formulations is appealing because the dimension of  $K_{\text{SOS}}$  is generally much smaller than the cone dimension in SDP formulations that are amenable to more specialized algorithms [4, 16]. In later sections we describe efficient LHSCBs and membership checks for each cone we introduce.

The output of the lifting operator depends on the polynomial basis chosen for  $\mathbb{R}[\mathbf{x}]_{n,d}$  as well as the basis for  $\mathbb{R}[\mathbf{x}]_{n,2d}$ . Following Papp and Yildiz [16], we use a set of Lagrange polynomials that are interpolant on some points  $\mathbf{t}_{i \in [1..U]}$  as the basis for  $\mathbb{R}_{n,2d}[\mathbf{x}]$  and the multivariate Chebyshev polynomials [9] as the basis in  $\mathbb{R}_{n,d}[\mathbf{x}]$ . These choices give the particular lifting operator  $\Lambda_{\text{SOS}}(\mathbf{s})$ :

$$\Lambda_{\text{SOS}}(\mathbf{s})_{i,j} = \sum_{u \in [1..U]} p_i(\mathbf{t}_u) p_j(\mathbf{t}_u) s_u \quad \forall i, j \in [1..L]. \quad (1.3)$$

Equivalently,  $\Lambda_{\text{SOS}}(\mathbf{s}) = \mathbf{P}^\top \text{Diag}(\mathbf{s}) \mathbf{P}$ , where  $P_{u,\ell} = p_\ell(\mathbf{t}_u)$  for all  $u \in [1..U]$ ,  $\ell \in [1..L]$ . The adjoint  $\Lambda_{\text{SOS}}^* : \mathbb{S}^L \rightarrow \mathbb{R}^U$  is given by  $\Lambda_{\text{SOS}}^*(\mathbf{S}) = \text{diag}(\mathbf{PSP}^\top)$ . Papp and

<sup>1</sup> We direct the interested reader to Faybusovich [5], who obtained non-linear barriers for the cone of univariate polynomials generated by Chebyshev systems by computing the universal volume barrier of Nesterov and Nemirovskii [14], which is unrelated to the SDP representations of these polynomials.

Yildiz [16] show that the Lagrange basis gives rise to expressions for the gradient and Hessian of the barrier for  $K_{\text{SOS}}^*$  that are computable in  $\mathcal{O}(LU^2)$  time for any  $d, n \geq 1$ . Although we assume for simplicity that  $p$  is a dense basis for  $\mathbb{R}[\mathbf{x}]_{n,d}$ , this is without loss of generality. A modeler with access to a suitable sparse basis of  $\bar{L} < L$  polynomials in  $\mathbb{R}[\mathbf{x}]_{n,d}$  and  $\bar{U} < U$  interpolation points, could use Eq. (1.3) and obtain a barrier with parameter  $\bar{L}$ .

## 2 Polynomial generalizations for three conic sets

The first set we consider are the polynomial matrices  $\mathbf{Q}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}^{m \times m}$  (i.e.  $m \times m$  matrices with components that are polynomials in  $n$  variables of maximum degree  $2d$ )<sup>2</sup> satisfying the constraint:

$$\mathbf{Q}(\mathbf{x}) \geq 0 \quad \forall \mathbf{x}. \quad (2.1)$$

One of the first applications of matrix SOS constraints was by Henrion and Lasserre [8]. The moment-SOS hierarchy was extended from the scalar case to the matrix case, using a suitable extension of Putinar's Positivstellensatz studied by Hol and Scherer [10] and Kojima [11].

This constraint has various applications in statistics, control, and engineering (e.g. [2, 7]). A tractable restriction for Eq. (2.1) is given by the SOS formulation:

$$\mathbf{y}^\top \mathbf{Q}(\mathbf{x}) \mathbf{y} \in K_{\text{SOS}} \quad \forall \mathbf{y} \in \mathbb{R}^m. \quad (2.2)$$

This formulation is sometimes implemented in practice and requires an SOS cone of dimension  $U \text{sd}(m)$  (by exploiting the fact that all terms are bilinear in the  $\mathbf{y}$  variables). It is well known that Eq. (2.2) is equivalent to restricting  $\mathbf{Q}(\mathbf{x})$  to be an *SOS matrix* of the form  $\mathbf{Q}(\mathbf{x}) = \mathbf{M}(\mathbf{x})^\top \mathbf{M}(\mathbf{x})$  for some  $N \in \mathbb{N}$  and  $\mathbf{M}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,d}^{N \times m}$  [3, Definition 3.76]. To be consistent in terminology with the other cones we introduce, we refer to SOS matrices as *SOS-PSD* matrices, or belonging to  $K_{\text{SOSPSD}}$ . We show how to characterize  $K_{\text{SOSPSD}}$  and use it directly in an interior point algorithm in Sect. 3.

The second set we consider are the polynomial vectors  $\mathbf{q}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}^m$  satisfying:

$$q_1(\mathbf{x}) \geq \sqrt{\sum_{i \in [2..m]} (q_i(\mathbf{x}))^2} \quad \forall \mathbf{x}, \quad (2.3)$$

and hence requiring  $\mathbf{q}(\mathbf{x})$  to be in the epigraph set of the  $\ell_2$ -norm function (second order cone) pointwise (cf. Eq. (2.1) requiring the polynomial matrix to be in the PSD cone). A tractable restriction for this constraint is given by the SOS formulation:

$$\mathbf{y}^\top \text{Arw}(\mathbf{q}(\mathbf{x})) \mathbf{y} \in K_{\text{SOS}} \quad \forall \mathbf{y} \in \mathbb{R}^m, \quad (2.4)$$

<sup>2</sup> We assume that polynomial components in vectors and matrices involve the same variables and have the same maximum degree, to avoid more cumbersome notation.

where  $\text{Arw} : \mathbb{R}[\mathbf{x}]_{n,2d}^m \rightarrow \mathbb{R}[\mathbf{x}]_{n,2d}^{m \times m}$  is defined by:

$$\begin{aligned} \text{Arw}(\mathbf{p}(\mathbf{x})) &= \begin{bmatrix} p_1(\mathbf{x}) & \bar{\mathbf{p}}(\mathbf{x})^\top \\ \bar{\mathbf{p}}(\mathbf{x}) & p_1(\mathbf{x})\mathbf{I}_{m-1} \end{bmatrix}, \\ \mathbf{p}(\mathbf{x}) &= (p_1(\mathbf{x}), \bar{\mathbf{p}}(\mathbf{x})) \in \mathbb{R}[\mathbf{x}]_{n,2d} \times \mathbb{R}[\mathbf{x}]_{n,2d}^{m-1}. \end{aligned} \quad (2.5)$$

Due to the equivalence between Eq. (2.2) and membership in  $K_{\text{SOSPSD}}$ , Equation (2.4) is equivalent to requiring that  $\mathbf{q}(\mathbf{x})$  belongs to the cone we denote  $K_{\text{Arw SOSPSD}}$  defined by:

$$K_{\text{Arw SOSPSD}} = \{\mathbf{q}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}^m : \text{Arw}(\mathbf{q}(\mathbf{x})) \in K_{\text{SOSPSD}}\}. \quad (2.6)$$

Membership in  $K_{\text{Arw SOSPSD}}$  ensures Eq. (2.3) holds due to the SDP representation of the second order cone [1], and the fact that the SOS-PSD condition certifies pointwise positive semidefiniteness. An alternative restriction of Eq. (2.3) is described by the set we denote  $K_{\text{SOS } \ell_2}$ , which is not representable by the usual scalar polynomial SOS cone in general:

$$K_{\text{SOS } \ell_2} = \left\{ \begin{array}{l} \mathbf{q}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}^m : \exists N \in \mathbb{N}, \mathbf{p}_{i \in [1..N]}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,d}^m, \\ \mathbf{q}(\mathbf{x}) = \sum_{i \in [1..N]} \mathbf{p}_i(\mathbf{x}) \circ \mathbf{p}_i(\mathbf{x}) \end{array} \right\}, \quad (2.7)$$

where  $\circ : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  is defined by:

$$\mathbf{x} \circ \mathbf{y} = \begin{bmatrix} \mathbf{x}^\top \mathbf{y} \\ x_1 \bar{\mathbf{y}} + y_1 \bar{\mathbf{x}} \end{bmatrix}, \quad \mathbf{x} = (x_1, \bar{\mathbf{x}}), \quad \mathbf{y} = (y_1, \bar{\mathbf{y}}) \in \mathbb{R} \times \mathbb{R}^{m-1}, \quad (2.8)$$

and  $\circ : \mathbb{R}[\mathbf{x}]_{n,d}^m \times \mathbb{R}[\mathbf{x}]_{n,d}^m \rightarrow \mathbb{R}[\mathbf{x}]_{n,2d}^m$  on polynomial vectors is defined analogously. This set was also studied by Kojima and Muramatsu with a focus on extending Positivstellensatz results [12]. The validity of  $K_{\text{SOS } \ell_2}$  as a restriction of Eq. (2.3) follows from the characterization of the second order cone as a *cone of squares* [1, Section 4]. For this reason we will refer to the elements of  $K_{\text{SOS } \ell_2}$  as the *SOS- $\ell_2$  polynomials*. For a polynomial vector in  $\mathbb{R}[\mathbf{x}]_{n,2d}^m$ , the dimension of  $K_{\text{SOS } \ell_2}$  is  $Um$ , which is favorable to the dimension  $U \text{sd}(m)$  of  $K_{\text{SOS}}$  required for Eq. (2.4) or  $K_{\text{SOSPSD}}$  in Eq. (2.6). In addition, we show in Sect. 4.1 that  $K_{\text{SOS } \ell_2}$  admits an LHSCB with smaller parameter than  $K_{\text{Arw SOSPSD}}$ . However, we conjecture that for general  $n$  and  $d$ ,  $K_{\text{SOS } \ell_2} \subsetneq K_{\text{Arw SOSPSD}}$  (for example, consider the vector  $[1 + x^2, 1 - x^2, 2x]$ , which belongs to  $K_{\text{Arw SOSPSD}}$  but not  $K_{\text{SOS } \ell_2}$ ). A third formulation could be obtained by modifying the SDP formulation for  $K_{\text{Arw SOSPSD}}$  to account for all sparsity in the  $\mathbf{y}$  monomials (by introducing a specialized cone for the Gram matrix of  $\mathbf{y}^\top \text{Arw}(\mathbf{q}(\mathbf{x}))\mathbf{y}$ ). However, this approach suffers from requiring  $\mathcal{O}(L^2)$  conic variables for each polynomial in  $\mathbf{q}(\mathbf{x})$ .

The third and final set we consider is also described through a constraint on a polynomial vector  $\mathbf{q}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}^m$ . This constraint is given by:

$$q_1(\mathbf{x}) \geq \sum_{i \in [2..m]} |q_i(\mathbf{x})| \quad \forall \mathbf{x}, \quad (2.9)$$

**Table 1** Properties of new cones compared to SOS formulations

	SOS-PSD		SOS-L2		SOS-L1	
	$K_{\text{SOSPSD}}$	(2.2)	$K_{\text{SOS } \ell_2}$	(2.4)	$K_{\text{SOS } \ell_1}$	(2.10)
Cone dim.	$U \text{ sd}(m)$	$U \text{ sd}(m)$	$Um$	$U \text{ sd}(m)$	$Um$	$U(2m - 1)$
Parameter	$Lm$	$Lm$	$2L$	$Lm$	$Lm$	$L(2m - 1)$
Hessian flops	$\mathcal{O}(LU^2m^3)$	$\mathcal{O}(LU^2m^5)$	$\mathcal{O}(LU^2m^2)$	$\mathcal{O}(LU^2m^5)$	$\mathcal{O}(LU^2m)$	$\mathcal{O}(LU^2m)$
Conservatism	Equal	–	Greater	–	Equal	–
Equalities	0	0	0	0	0	$U(m - 1)$
Variables	0	0	0	0	0	$2U(m - 1)$

and hence requires the polynomial vector to be in the epigraph set of the  $\ell_1$ -norm function ( $\ell_1$ -norm cone) pointwise. A tractable restriction for this constraint is given by the SOS formulation:

$$q_1(\mathbf{x}) - \sum_{i \in \llbracket 2..m \rrbracket} (p_i(\mathbf{x})^+ + p_i(\mathbf{x})^-) \in K_{\text{SOS}}, \quad (2.10a)$$

$$q_i(\mathbf{x}) = p_i(\mathbf{x})^+ - p_i(\mathbf{x})^-, p_i(\mathbf{x})^+, p_i(\mathbf{x})^- \in K_{\text{SOS}} \quad \forall i \in \llbracket 2..m \rrbracket, \quad (2.10b)$$

which uses auxiliary polynomial variables  $p_{i \in \llbracket 2..m \rrbracket}^+(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}$  and  $p_{i \in \llbracket 2..m \rrbracket}^-(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}$ . We refer to the projection of Eq. (2.10) onto  $\mathbf{q}(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{n,2d}^m$  as  $K_{\text{SOS } \ell_1}$  and to its elements as the *SOS-L1* polynomials. Note that the dimension of  $K_{\text{SOS } \ell_1}$  is  $Um$ , while Eq. (2.10) requires  $2m - 1$  SOS cones of dimension  $U$  and  $U(m - 1)$  additional equality constraints. In Sect. 4.2 we derive an  $Lm$ -LHSCB that allows us to optimize over  $K_{\text{SOS } \ell_1}$  directly, while Eq. (2.10) would require an LHSCB with parameter  $L(2m - 1)$ .

We summarize some key properties of the new cones and SOS formulations in Table 1: the total dimension of cones involved, the parameter of an LHSCB for the conic sets, the time complexity to calculate the Hessian of the LHSCB, the level of conservatism of each new conic set compared to its alternative SOS formulation, and the number of auxiliary equality constraints and variables that need to be added in an optimization problem. A computational comparison of each pair of formulations using an example problem can be found at <https://github.com/chriscoey/Hypatia.jl/wiki>. The algorithmic advantages from the new cones usually translate to faster solve times in practice, and our experiments agree with the conjecture  $K_{\text{SOS } \ell_2} \subsetneq K_{\text{Arw SOSPSD}}$ .

### 3 SOS-PSD and SOS-L2 cones from general algebras

The ideas introduced by Papp and Alizadeh [15] relating to SOS cones in *general algebras* allow us to characterize  $K_{\text{SOSPSD}}$  and  $K_{\text{SOS } \ell_2}$  without auxiliary SOS polynomial constraints. As in Papp and Alizadeh [15], let us define  $(A, B, \diamond)$  as a general algebra if  $A, B$  are vector spaces and  $\diamond : A \times A \rightarrow B$  is a bilinear product that satisfies the

distributive property. For a general algebra  $(A, B, \diamond)$ , Papp and Alizadeh [15] define the SOS cone  $K_\diamond$ :

$$K_\diamond = \{b \in B : \exists N \in \mathbb{N}, a_i \in [1..N] \in A, b = \sum_{i \in [1..N]} a_i \diamond a_i\}. \quad (3.1)$$

For instance,  $\mathbb{S}_+$  is equal to the SOS cone of  $(\mathbb{R}^m, \mathbb{S}^m, \bar{\diamond})$  for  $\bar{\diamond}$  given by  $\mathbf{x} \bar{\diamond} \mathbf{y} = \frac{1}{2}(\mathbf{xy}^\top + \mathbf{yx}^\top)$ . The second order cone is equal to the SOS cone of  $(\mathbb{R}^m, \mathbb{R}^m, \circ)$ .  $K_{\text{SOS}}$  is equal to the SOS cone of  $(\mathbb{R}[\mathbf{x}]_{n,d}, \mathbb{R}[\mathbf{x}]_{n,2d}, \cdot)$  where  $\cdot$  is the product of polynomials. To obtain our vectorized representation of  $K_{\text{SOS}}$  we can redefine the function  $\lambda : \mathbb{R}^L \times \mathbb{R}^L \rightarrow \mathbb{R}^U$  so that for  $\mathbf{p}_i, \mathbf{p}_j \in \mathbb{R}^L$  representing coefficients of any polynomials in  $\mathbb{R}[\mathbf{x}]_{n,d}$ ,  $\lambda(\mathbf{p}_i, \mathbf{p}_j)$  returns the vector of coefficients of the product of the polynomials. Then  $K_{\text{SOS}}$  is equal to the SOS cone of  $(\mathbb{R}^L, \mathbb{R}^U, \lambda)$ .

As we describe in Sect. 3.1, Papp and Alizadeh [15] also show how to build lifting operators for general algebras. This allows us to construct membership checks and easily computable LHSCBs for  $K_{\text{SOSPSD}}^*$  and  $K_{\text{SOS}\ell_2}^*$  once we represent them as SOS cones of *tensor products* of algebras.

The tensor product of two algebras  $(A_1, B_1, \diamond_1)$  and  $(A_2, B_2, \diamond_2)$  is a new algebra  $(A_1 \otimes A_2, B_1 \otimes B_2, \diamond_1 \otimes \diamond_2)$ , where  $\diamond_1 \otimes \diamond_2$  is defined via its action on *elementary tensors*. For  $\mathbf{u}_1, \mathbf{v}_1 \in A_1$  and  $\mathbf{u}_2, \mathbf{v}_2 \in A_2$ :

$$(\mathbf{u}_1 \otimes \mathbf{u}_2) \diamond_1 \otimes \diamond_2 (\mathbf{v}_1 \otimes \mathbf{v}_2) = (\mathbf{u}_1 \diamond_1 \mathbf{v}_1) \otimes (\mathbf{u}_2 \diamond_2 \mathbf{v}_2). \quad (3.2)$$

The algebra we are interested in for a functional representation of  $K_{\text{SOSPSD}}$  is the tensor product of  $(\mathbb{R}[\mathbf{x}]_{n,d}, \mathbb{R}[\mathbf{x}]_{n,2d}, \cdot)$  with  $(\mathbb{R}^m, \mathbb{S}^m, \bar{\diamond})$ . We can think of elements in  $\mathbb{R}[\mathbf{x}]_{n,d} \otimes \mathbb{R}^m$  as polynomial vectors in  $\mathbb{R}[\mathbf{x}]_{n,d}^m$ , and  $\mathbb{R}[\mathbf{x}]_{n,2d} \otimes \mathbb{S}^m$  as the symmetric polynomial matrices in  $\mathbb{R}[\mathbf{x}]_{n,2d}^{m \times m}$ . The SOS cone of  $(\mathbb{R}[\mathbf{x}]_{n,d} \otimes \mathbb{R}^m, \mathbb{R}[\mathbf{x}]_{n,2d} \otimes \mathbb{S}^m, \cdot \otimes \bar{\diamond})$  corresponds to the polynomial matrices that can be written as  $\sum_{i \in [1..N]} \mathbf{m}_i(\mathbf{x}) \mathbf{m}_i(\mathbf{x})^\top$  with  $\mathbf{m}_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]^m$  for all  $i \in [1..N]$  [15, Section 4.3], which is exactly  $K_{\text{SOSPSD}}$ . Equivalently, a vectorized representation of  $K_{\text{SOSPSD}}$  can be characterized as the SOS cone of  $(\mathbb{R}^L \otimes \mathbb{R}^m, \mathbb{R}^U \otimes \mathbb{S}^m, \lambda \otimes \bar{\diamond})$ . We can think of  $\mathbb{R}^L \otimes \mathbb{R}^m$  as  $\mathbb{R}^{L \times m}$  and we can think of  $\mathbb{R}^U \otimes \mathbb{S}^m$  as a subspace of  $\mathbb{R}^{U \times m \times m}$  that represents the coefficients of symmetric polynomial matrices.

Likewise, the algebra we are interested in for a functional representation of  $K_{\text{SOS}\ell_2}$  is the tensor product of  $(\mathbb{R}[\mathbf{x}]_{n,d}, \mathbb{R}[\mathbf{x}]_{n,2d}, \cdot)$  with  $(\mathbb{R}^m, \mathbb{R}^m, \circ)$ . We can think of  $\mathbb{R}[\mathbf{x}]_{n,d} \otimes \mathbb{R}^m$  and  $\mathbb{R}[\mathbf{x}]_{n,2d} \otimes \mathbb{R}^m$  as  $\mathbb{R}[\mathbf{x}]_{n,d}^m$  and  $\mathbb{R}[\mathbf{x}]_{n,2d}^m$  respectively. The SOS cone of the tensor product of these algebras then corresponds to  $K_{\text{SOS}\ell_2}$  due to Eq. (2.7). A vectorized representation of  $K_{\text{SOS}\ell_2}$  may be characterized as the SOS cone of  $(\mathbb{R}^L \otimes \mathbb{R}^m, \mathbb{R}^U \otimes \mathbb{R}^m, \lambda \otimes \circ)$ . We can think of  $\mathbb{R}^U \otimes \mathbb{R}^m$  as the coefficients of polynomial vectors, represented in  $\mathbb{R}^{U \times m}$ .

### 3.1 Lifting operators for SOS-PSD and SOS-L2

The lifting operator of  $(A, B, \diamond)$ , when  $A$  and  $B$  are finite dimensional, is defined by Papp and Alizadeh [15] as the function  $\Lambda_\diamond : B \rightarrow \mathbb{S}^{|A|}$  satisfying  $\langle a_1, \Lambda_\diamond(b) a_2 \rangle_A$

$= \langle b, a_1 \diamond a_2 \rangle_B$  for all  $a_1, a_2 \in A, b \in B$ . This leads to the following descriptions of  $K_\diamond$  and  $K_\diamond^*$  [15, Theorem 3.2]:

$$K_\diamond = \{s \in B : \exists S \succeq 0, s = \Lambda_\diamond^*(S)\}, \quad K_\diamond^* = \{s \in B : \Lambda_\diamond(s) \succeq 0\}. \quad (3.3)$$

Recall that in order to use either  $K_\diamond$  or  $K_\diamond^*$  in a generic interior point algorithm, we require efficient oracles for a membership check and derivatives of an LHSCB of  $K_\diamond$  or  $K_\diamond^*$ . If  $\Lambda_\diamond(s)$  is efficiently computable, Equation (3.3) provides a membership check for  $K_\diamond^*$ . Furthermore, an LHSCB for  $K_\diamond^*$  is given by  $s \mapsto -\log \det(\Lambda_\diamond(s))$  with barrier parameter  $|A|$  due to the linearity of  $\Lambda_\diamond$  [14, Proposition 5.1.1]. The following lemma describes how to compute  $\Lambda_\diamond(s)$  for a tensor product algebra.

**Lemma 3.1** [15, Lemma 4.1]: *If  $w_1 \in B_1$  and  $w_2 \in B_2$ , then:*

$$\Lambda_{\diamond_1 \otimes \diamond_2}(w_1 \otimes w_2) = \Lambda_{\diamond_1}(w_1) \otimes_K \Lambda_{\diamond_2}(w_2). \quad (3.4)$$

Let us define  $\otimes : \mathbb{R}^U \times \mathbb{S}^m \rightarrow \mathbb{R}^{U \times m \times m}$  such that  $(u \otimes V)_{i,j,k} = u_i V_{j,k}$  and let us represent the coefficients of a polynomial matrix by a tensor  $S \in \mathbb{R}^{U \times m \times m}$ . Then we may write  $S = \sum_{i \in [1..m], j \in [1..i]} S_{i,j} \otimes E_{i,j}$ , where  $E_{i,j} \in \mathbb{R}^{m \times m}$  is a matrix of zeros and ones with  $E_{i,j} = E_{j,i} = 1$  and  $S_{i,j} \in \mathbb{R}^U$  are the coefficients of the polynomial in row  $i$  and column  $j$ . Applying Lemma 3.1, the lifting operator for  $K_{\text{SOSPSD}}$ ,  $\Lambda_{\text{SOSPSD}} : \mathbb{R}^{U \times m \times m} \rightarrow \mathbb{S}^{Lm}$  is:

$$\Lambda_{\text{SOSPSD}}(S) = \Lambda_{\lambda \otimes \bar{\diamond}}(S) = \Lambda_{\lambda \otimes \bar{\diamond}}\left(\sum_{i \in [1..m], j \in [1..i]} S_{i,j} \otimes E_{i,j}\right) \quad (3.5a)$$

$$= \sum_{i \in [1..m], j \in [1..i]} \Lambda_{\text{SOS}}(S_{i,j}) \otimes_K E_{i,j}. \quad (3.5b)$$

The output is a block matrix, where each  $L \times L$  submatrix in the  $i$ th group of rows and  $j$ th group of columns is  $\Lambda_{\text{SOSPSD}}(S)_{i,j} = \Lambda_{\text{SOS}}(S_{i,j})$  for all  $i, j \in [1..m]$ . The adjoint operator  $\Lambda_{\text{SOSPSD}}^* : \mathbb{S}^{Lm} \rightarrow \mathbb{R}^{U \times m \times m}$  may also be defined blockwise,  $\Lambda_{\text{SOSPSD}}^*(S)_{i,j} = \Lambda_{\text{SOS}}^*(S_{i,j})$  for all  $i, j \in [1..m]$  where  $S_{i,j} \in \mathbb{R}^{L \times L}$  is the  $(i, j)$ th submatrix in  $S$ . Note that the Hessian of  $s \mapsto -\log \det(\Lambda_{\text{SOSPSD}}(s))$  can be evaluated in  $\mathcal{O}(LU^2m^3)$  time, which is implemented in the Hypatia solver [4].

Likewise, we use a tensor  $s \in \mathbb{R}^{U \times m}$  to describe the coefficients of a polynomial vector, and write  $s_i \in \mathbb{R}^U$  to denote the vector of coefficients of the polynomial in component  $i$ . Applying Lemma 3.1 again, we obtain the (blockwise) definition of the lifting operator for  $K_{\text{SOS } \ell_2}$ ,  $\Lambda_{\text{SOS } \ell_2} : \mathbb{R}^{U \times m} \rightarrow \mathbb{S}^{Lm}$ :

$$\Lambda_{\text{SOS } \ell_2}(s)_{i,j} = \begin{cases} \Lambda_{\text{SOS}}(s_1) & i = j \\ \Lambda_{\text{SOS}}(s_j) & i = 1, j \neq 1 \\ \Lambda_{\text{SOS}}(s_i) & i \neq 1, j = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall i, j \in [1..m], \quad (3.6)$$



where  $\Lambda_{\text{SOS } \ell_2}(\mathbf{s})_{i,j} \in \mathbb{S}^L$  is the  $(i, j)$ th submatrix of  $\Lambda_{\text{SOS } \ell_2}(\mathbf{s})$ . Thus  $\Lambda_{\text{SOS } \ell_2}(\mathbf{s})$  has a block arrowhead structure. The output of the adjoint operator  $\Lambda_{\text{SOS } \ell_2}^* : \mathbb{S}^{Lm} \rightarrow \mathbb{R}^{U \times m}$  may be defined as:

$$\Lambda_{\text{SOS } \ell_2}^*(\mathbf{S})_i = \begin{cases} \sum_{j \in \llbracket 1..m \rrbracket} \Lambda_{\text{SOS}}^*(\mathbf{S}_{j,j}) & i = 1 \\ \Lambda_{\text{SOS}}^*(\mathbf{S}_{1,i}) + \Lambda_{\text{SOS}}^*(\mathbf{S}_{i,1}) & i \neq 1 \end{cases} \quad \forall i \in \llbracket 1..m \rrbracket, \quad (3.7)$$

where  $\Lambda_{\text{SOS } \ell_2}^*(\mathbf{S})_i \in \mathbb{R}^U$  is the  $i$ th slice of  $\Lambda_{\text{SOS } \ell_2}^*(\mathbf{S})$  and  $\mathbf{S}_{i,j} \in \mathbb{R}^{L \times L}$  is the  $(i, j)$ th block in  $\mathbf{S}$  for all  $i, j \in \llbracket 1..m \rrbracket$ .

## 4 Efficient barriers for SOS-L2 and SOS-L1

As for  $K_{\text{SOSPSD}}^*$  and  $K_{\text{SOS } \ell_2}^*$ , we show that a barrier for  $K_{\text{SOS } \ell_1}^*$  can be obtained by composing a linear lifting operator with the logdet barrier. This is sufficient to optimize over  $K_{\text{SOSPSD}}$ ,  $K_{\text{SOS } \ell_2}$  and  $K_{\text{SOS } \ell_1}$  without high dimensional SDP formulations. However, for  $K_{\text{SOS } \ell_2}^*$  and  $K_{\text{SOS } \ell_1}^*$  we can derive improved barriers by composing non-linear functions with the logdet barrier instead. We show that these compositions are indeed LHSCBs.

### 4.1 SOS-L2

Recall Eq. (3.3) suggests that checking membership in  $K_{\text{SOS } \ell_2}^*$  amounts to checking positive definiteness of  $\Lambda_{\text{SOS } \ell_2}(\mathbf{s})$  with side dimension  $Lm$ . This membership check corresponds to a *straightforward* LHSCB with parameter  $Lm$  given by  $\mathbf{s} \mapsto -\log\det(\Lambda_{\text{SOS } \ell_2}(\mathbf{s}))$ . We now show that by working with a Schur complement of  $\Lambda_{\text{SOS } \ell_2}(\mathbf{s})$ , we obtain a membership check for  $K_{\text{SOS } \ell_2}^*$  that requires factorizations of only two matrices with side dimension  $L$  and implies an LHSCB with parameter  $2L$ .

Let  $\Pi : \mathbb{R}^{U \times m} \rightarrow \mathbb{S}^L$  return the Schur complement:

$$\Pi(\mathbf{s}) = \Lambda_{\text{SOS}}(\mathbf{s}_1) - \sum_{i \in \llbracket 2..m \rrbracket} \Lambda_{\text{SOS}}(\mathbf{s}_i) \Lambda_{\text{SOS}}(\mathbf{s}_1)^{-1} \Lambda_{\text{SOS}}(\mathbf{s}_i). \quad (4.1)$$

By Eqs. (3.3) and (4.1):

$$\begin{aligned} K_{\text{SOS } \ell_2}^* &= \{\mathbf{s} \in \mathbb{R}^{U \times m} : \Lambda_{\text{SOS } \ell_2}(\mathbf{s}) \succeq 0\} = \text{cl}\{\mathbf{s} \in \mathbb{R}^{U \times m} : \Lambda_{\text{SOS } \ell_2}(\mathbf{s}) \succ 0\} \\ &= \text{cl}\{\mathbf{s} \in \mathbb{R}^{U \times m} : \Lambda_{\text{SOS}}(\mathbf{s}_1) \succ 0, \Pi(\mathbf{s}) \succ 0\}. \end{aligned} \quad (4.2)$$

Equation (4.2) describes a simple membership check. Furthermore, the function  $F : \mathbb{R}^{U \times m} \rightarrow \mathbb{R}$  defined by:

$$F(\mathbf{s}) = -\log\det(\Pi(\mathbf{s})) - \log\det(\Lambda_{\text{SOS}}(\mathbf{s}_1)) \quad (4.3a)$$

$$= -\log\det(\Lambda_{\text{SOS } \ell_2}(\mathbf{s})) + (m-2)\log\det(\Lambda_{\text{SOS}}(\mathbf{s}_1)), \quad (4.3b)$$

is a  $2L$ -LHSCB barrier for  $K_{\text{SOS } \ell_2}$ .

**Theorem 4.1** *The function  $F$  defined by Equation (4.3) is a  $2L$ -LHSCB for  $K_{\text{SOS } \ell_2}^*$ .*

**Proof** It is easy to verify that  $F$  is a logarithmically homogeneous barrier, so we show it is a  $2L$ -self-concordant barrier for  $K_{\text{SOS } \ell_2}^*$ . We first show that  $\hat{F} : \mathbb{S}_{++}^L \times (\mathbb{R}^{L \times L})^{m-1} \rightarrow \mathbb{R}$  defined as  $\hat{F}(\mathbf{X}_1, \dots, \mathbf{X}_m) = -\log \det(\mathbf{X}_1 - \sum_{i \in \llbracket 2..m \rrbracket} \mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{X}_i^\top) - \log \det(\mathbf{X}_1)$ , is a  $2L$ -self-concordant barrier for the cone:

$$K_{\ell_2}^m = \text{cl} \left\{ (\mathbf{X}_1, \dots, \mathbf{X}_m) \in \mathbb{S}_{++}^L \times (\mathbb{R}^{L \times L})^{m-1} : \mathbf{X}_1 - \sum_{i \in \llbracket 2..m \rrbracket} \mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{X}_i^\top \succ 0 \right\}. \quad (4.4)$$

We then argue that  $F$  is a composition of  $\hat{F}$  with the linear map  $(\mathbf{s}_1, \dots, \mathbf{s}_m) \mapsto (\Lambda_{\text{SOS}}(\mathbf{s}_1), \dots, \Lambda_{\text{SOS}}(\mathbf{s}_m))$  and  $K_{\text{SOS } \ell_2}^*$  is an inverse image of  $K_{\ell_2}^m$  under the same map. Then by Nesterov and Nemirovskii [14, Proposition 5.1.1]  $F$  is self-concordant.

Let  $\Gamma = \mathbb{S}_{++}^L \times (\mathbb{R}^{L \times L})^{m-1}$  and  $\mathbf{G} : \text{int}(\Gamma) \rightarrow \mathbb{S}^L$  be defined as:

$$\mathbf{G}(\mathbf{X}_1, \dots, \mathbf{X}_m) = \mathbf{X}_1 - \sum_{i \in \llbracket 2..m \rrbracket} \mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{X}_i^\top. \quad (4.5)$$

Let us check that  $\mathbf{G}$  is  $(\mathbb{S}_{++}^L, 1)$ -compatible with the domain  $\Gamma$  in the sense of [14, Definition 5.1.1]. This requires that  $\mathbf{G}$  is  $C^3$ -smooth on  $\text{int}(\Gamma)$ ,  $\mathbf{G}$  is concave with respect to  $\mathbb{S}_{++}^L$ , and at each point  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_m) \in \text{int}(\Gamma)$  and any direction  $\mathbf{V} = (\mathbf{V}_1, \dots, \mathbf{V}_m) \in \mathbb{S}^L \times (\mathbb{R}^{L \times L})^{m-1}$  such that  $-\mathbf{X}_1 \preceq \mathbf{V}_1 \preceq \mathbf{X}_1$ , the directional derivatives of  $\mathbf{G}$  satisfy:

$$\frac{d^3 \mathbf{G}}{d\mathbf{X}^3}[\mathbf{V}, \mathbf{V}, \mathbf{V}] \leq -3 \frac{d^2 \mathbf{G}}{d\mathbf{X}^2}[\mathbf{V}, \mathbf{V}]. \quad (4.6)$$

Let  $\mathbf{V} \in \mathbb{S}^L \times (\mathbb{R}^{L \times L})^{m-1}$ . It can be checked that  $\frac{d^3 \mathbf{G}}{d\mathbf{X}^3}$  is continuous on the domain of  $\mathbf{G}$  and we have the directional derivatives:

$$\frac{d^2 \mathbf{G}}{d\mathbf{X}^2}[\mathbf{V}, \mathbf{V}] = -2 \sum_{i \in \llbracket 2..m \rrbracket} (\mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{V}_1 - \mathbf{V}_i) \mathbf{X}_1^{-1} (\mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{V}_1 - \mathbf{V}_i)^\top, \quad (4.7)$$

$$\frac{d^3 \mathbf{G}}{d\mathbf{X}^3}[\mathbf{V}, \mathbf{V}, \mathbf{V}] = 6 \sum_{i \in \llbracket 2..m \rrbracket} (\mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{V}_1 - \mathbf{V}_i) \mathbf{X}_1^{-1} \mathbf{V}_1 \mathbf{X}_1^{-1} (\mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{V}_1 - \mathbf{V}_i)^\top. \quad (4.8)$$

Since  $\mathbf{X}_1 \succ 0$  in  $\text{int}(\Gamma)$ ,  $-\frac{d^2 \mathbf{G}}{d\mathbf{X}^2}[\mathbf{V}, \mathbf{V}] \succeq 0$  and so by Nesterov and Nemirovskii [14, Lemma 5.1.2],  $\mathbf{G}$  is concave with respect to  $\mathbb{S}_{++}^L$ . It remains to show that (4.6) is satisfied. Since the directional derivatives decouple by each index  $i$  in the sum, it is sufficient to show that the inequality is satisfied for each  $i \in \llbracket 2..m \rrbracket$ . For this, it is sufficient that  $6\mathbf{X}_1^{-1} \mathbf{V}_1 \mathbf{X}_1^{-1} \preceq -3 \times -2\mathbf{X}_1^{-1}$  for all  $-\mathbf{X}_1 \preceq \mathbf{V}_1 \preceq \mathbf{X}_1$ , which follows since  $\mathbf{X}_1$  is positive definite on  $\text{int}(\Gamma)$ . Now by [14, proposition 5.1.7],  $\hat{F}$  is a  $2L$ -LHSCB. The same is true for  $F$  by composing  $\hat{F}$  with a linear map.  $\square$

The Hessian of  $F$  can be evaluated in  $\mathcal{O}(LU^2 m^2)$  time, which is implemented in Hypatia.

## 4.2 SOS-L1

By combining Eqs. (1.2) and (2.10), the  $K_{\text{SOS } \ell_1}$  cone admits the semidefinite representation:

$$K_{\text{SOS } \ell_1} = \left\{ \mathbf{s} \in \mathbb{R}^{U \times m} : \exists \mathbf{S}_1, \mathbf{S}_{2,+}, \mathbf{S}_{2,-}, \dots, \mathbf{S}_{m,+}, \mathbf{S}_{m,-} \in \mathbb{S}_{+}^L, \right. \\ \left. \begin{aligned} \mathbf{s}_1 &= \Lambda_{\text{SOS}}^*(\mathbf{S}_1) + \sum_{i \in \llbracket 2..m \rrbracket} \Lambda_{\text{SOS}}^*(\mathbf{S}_{i,+} + \mathbf{S}_{i,-}), \\ \mathbf{s}_i &= \Lambda_{\text{SOS}}^*(\mathbf{S}_{i,+}) - \Lambda_{\text{SOS}}^*(\mathbf{S}_{i,-}) \quad \forall i \in \llbracket 2..m \rrbracket \end{aligned} \right\}. \quad (4.9)$$

Its dual cone is:

$$K_{\text{SOS } \ell_1}^* = \left\{ \mathbf{s} \in \mathbb{R}^{U \times m} : \Lambda_{\text{SOS}}(\mathbf{s}_1 + \mathbf{s}_i) \succeq 0, \Lambda_{\text{SOS}}(\mathbf{s}_1 - \mathbf{s}_i) \succeq 0 \quad \forall i \in \llbracket 2..m \rrbracket \right\}. \quad (4.10)$$

Equation (4.10) suggests that checking membership in  $K_{\text{SOS } \ell_1}^*$  amounts to checking positive definiteness of  $2(m-1)$  matrices of side dimension  $L$ . This membership check corresponds to a *straightforward* LHSCB with parameter  $2L(m-1)$  that is given by  $\mathbf{s} \mapsto -\sum_{i \in \llbracket 2..m \rrbracket} \log \det(\Lambda_{\text{SOS}}(\mathbf{s}_1 + \mathbf{s}_i) \Lambda_{\text{SOS}}(\mathbf{s}_1 - \mathbf{s}_i))$ . We now describe a membership check for  $K_{\text{SOS } \ell_1}^*$  that requires factorizations of only  $m$  matrices, and corresponds to an LHSCB with parameter  $Lm$ .

**Lemma 4.2** *The set  $\{\mathbf{X} \in \mathbb{S}_{++}^L, \mathbf{Y} \in \mathbb{S}^L : -\mathbf{X} \preceq \mathbf{Y} \preceq \mathbf{X}\}$  is equal to  $K_{\ell_2}^2 = \text{cl}\{\mathbf{X} \in \mathbb{S}_{++}^L, \mathbf{Y} \in \mathbb{S}^L : \mathbf{X} - \mathbf{Y}\mathbf{X}^{-1}\mathbf{Y} \succ 0\}$ .*

**Proof** For inclusion in one direction:

$$\text{cl}\{\mathbf{X} \in \mathbb{S}_{++}^L, \mathbf{Y} \in \mathbb{S}^L : \mathbf{X} - \mathbf{Y}\mathbf{X}^{-1}\mathbf{Y} \succ 0\} \quad (4.11a)$$

$$= \left\{ \mathbf{X} \in \mathbb{S}_{++}^L, \mathbf{Y} \in \mathbb{S}^L : \frac{\mathbf{X}}{\mathbf{Y}} \frac{\mathbf{Y}}{\mathbf{X}} \succeq 0, \frac{\mathbf{X}}{-\mathbf{Y}} \frac{-\mathbf{Y}}{\mathbf{X}} \succeq 0 \right\} \quad (4.11b)$$

$$\subseteq \left\{ \mathbf{X} \in \mathbb{S}_{++}^L, \mathbf{Y} \in \mathbb{S}^L : 2\mathbf{v}^\top \mathbf{X} \mathbf{v} \pm 2\mathbf{v}^\top \mathbf{Y} \mathbf{v} \geq 0, \quad \forall \mathbf{v} \in \mathbb{R}^L \right\} \quad (4.11c)$$

$$= \{\mathbf{X} \in \mathbb{S}_{++}^L, \mathbf{Y} \in \mathbb{S}^L : \mathbf{X} + \mathbf{Y} \succeq 0, \mathbf{X} - \mathbf{Y} \succeq 0\}. \quad (4.11d)$$

For the other direction, suppose  $-\mathbf{X} \prec \mathbf{Y} \prec \mathbf{X}$ . Then  $\mathbf{X} \succ 0, \mathbf{Y} + \mathbf{X} \succ 0, \mathbf{X} - \mathbf{Y} \succ 0$ . Note that  $(\mathbf{Y} + \mathbf{X})\mathbf{X}^{-1}(\mathbf{X} - \mathbf{Y}) = \mathbf{X} - \mathbf{Y}\mathbf{X}^{-1}\mathbf{Y}$  is symmetric. Due to Subramanian and Bhagwat [18, Corollary 1], this product of three matrices also has nonnegative eigenvalues. We conclude that  $-\mathbf{X} \prec \mathbf{Y} \prec \mathbf{X}$  implies  $\mathbf{X} \succ 0$  and  $\mathbf{X} - \mathbf{Y}\mathbf{X}^{-1}\mathbf{Y} \succeq 0$ . Since  $-\mathbf{X} \preceq \mathbf{Y} \preceq \mathbf{X} = \text{cl}\{-\mathbf{X} \prec \mathbf{Y} \prec \mathbf{X}\}$ , taking closures gives the result.  $\square$

By Lemma 4.2 we can write the dual cone as:

$$K_{\text{SOS } \ell_1}^* = \text{cl} \left\{ \begin{aligned} &\mathbf{s} \in \mathbb{R}^{U \times m} : \Lambda_{\text{SOS}}(\mathbf{s}_1) \succ 0, \\ &\Lambda_{\text{SOS}}(\mathbf{s}_1) - \Lambda_{\text{SOS}}(\mathbf{s}_i) \Lambda_{\text{SOS}}(\mathbf{s}_1)^{-1} \Lambda_{\text{SOS}}(\mathbf{s}_i) \succ 0, \quad \forall i \in \llbracket 2..m \rrbracket \end{aligned} \right\}. \quad (4.12)$$

**Theorem 4.3** The function  $F : \mathbb{R}^{U \times m} \rightarrow \mathbb{R}$  given by:

$$F(\mathbf{s}) = -\sum_{i \in \llbracket 2..m \rrbracket} \log \det(\Lambda_{\text{SOS}}(\mathbf{s}_1) - \Lambda_{\text{SOS}}(\mathbf{s}_i) \Lambda_{\text{SOS}}(\mathbf{s}_1)^{-1} \Lambda_{\text{SOS}}(\mathbf{s}_i)) - \log \det(\Lambda_{\text{SOS}}(\mathbf{s}_1)) \quad (4.13)$$

is an  $Lm$ -LHSCB for  $K_{\text{SOS } \ell_1}^*$ .

**Proof** It is easy to verify that  $F$  is a logarithmically homogeneous barrier, and we show it is an  $Lm$ -self-concordant barrier. As in Theorem 4.1, we define an auxiliary cone:

$$K_{\ell_\infty}^m = \{(\mathbf{X}_1, \dots, \mathbf{X}_m) \in \mathbb{S}_+^L \times (\mathbb{R}^{L \times L})^{m-1} : (\mathbf{X}_1, \mathbf{X}_i) \in K_{\ell_2}^2 \forall i \in \llbracket 2..m \rrbracket\}. \quad (4.14)$$

Let  $\hat{F} : \mathbb{S}_{++}^L \times (\mathbb{R}^{L \times L})^{m-1} \rightarrow \mathbb{R}$  be defined as  $\hat{F}(\mathbf{X}_1, \dots, \mathbf{X}_m) = -\sum_{i \in \llbracket 2..m \rrbracket} \log \det(\mathbf{X}_1 - \mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{X}_i^\top) - \log \det(\mathbf{X}_1)$ . We argue that  $\hat{F}$  is an  $Lm$ -self-concordant barrier for  $K_{\ell_\infty}^m$ .  $F$  is a composition of  $\hat{F}$  with the same linear map used in Theorem 4.1 and self-concordance of  $F$  then follows by the same reasoning.

Let  $\Gamma = \mathbb{S}_+^L \times (\mathbb{R}^{L \times L})^{m-1}$  and  $\mathbf{H} : \text{int}(\Gamma) \rightarrow (\mathbb{S}_+^L)^{m-1}$  be defined by:

$$\mathbf{H}(\mathbf{X}_1, \dots, \mathbf{X}_m) = (\mathbf{X}_1 - \mathbf{X}_2 \mathbf{X}_1^{-1} \mathbf{X}_2^\top, \dots, \mathbf{X}_1 - \mathbf{X}_m \mathbf{X}_1^{-1} \mathbf{X}_m^\top). \quad (4.15)$$

We claim that  $\mathbf{H}$  is  $((\mathbb{S}_+^L)^{m-1}, 1)$ -compatible with the domain  $\Gamma$ . This amounts to showing that for all  $i \in \llbracket 2..m \rrbracket$ , the mapping  $\mathbf{H}_i : \mathbb{S}_{++}^L \times \mathbb{R}^{L \times L} \rightarrow \mathbb{S}^L$ ,  $\mathbf{H}_i(\mathbf{X}) = \mathbf{X}_1 - \mathbf{X}_i \mathbf{X}_1^{-1} \mathbf{X}_i^\top$  is  $(\mathbb{S}_+^L, 1)$ -compatible with the domain  $\mathbb{S}_+^L \times \mathbb{R}^{L \times L}$  (the requirements for compatibility decouple for each  $i$ ). The latter holds since  $\mathbf{H}_i$  is equivalent to the function  $\mathbf{G}$  from Theorem 4.1 with  $m = 2$ . Then by Nesterov and Nemirovskii [14, Lemma 5.1.7],  $\hat{F}$  is an  $Lm$ -self-concordant barrier.  $\square$

The Hessian of  $F$  can be evaluated in  $\mathcal{O}(LU^2m)$  time, which is implemented in Hypatia. Note that we rely on an analogy of a representation for the  $\ell_\infty$ -norm cone (see [4, Section 5.1]) in Eq. (4.12). From this we derive an LHSCB that is analogous to the  $\ell_\infty$ -norm cone LHSCB. On the other hand, we are not aware of an efficient LHSCB for its dual, the  $\ell_1$ -norm cone, so we cannot use the same technique to derive an LHSCB for the dual of a polynomial analogy to the  $\ell_\infty$ -norm cone.

**Funding** Open Access funding provided by the MIT Libraries. This work has been partially funded by the National Science Foundation under Grant OAC-1835443 and the Office of Naval Research under Grant N00014-18-1-2079.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. *Math. Program.* **95**(1), 3–51 (2003)
2. Aylward, E.M., Parrilo, P.A., Slotine, J.J.E.: Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming. *Automatica* **44**(8), 2163–2170 (2008)
3. Blekherman, G., Parrilo, P.A., Thomas, R.R.: Semidefinite optimization and convex algebraic geometry. SIAM (2012)
4. Coey, C., Kapelevich, L., Vielma, J.P.: Solving natural conic formulations with Hypatia.jl. *INFORMS J. Comput.* (2021). <https://doi.org/10.1287/ijoc.2022.1202>
5. Faybusovich, L.: Self-concordant barriers for cones generated by chebyshev systems. *SIAM J. Optim.* **12**(3), 770–781 (2002)
6. Genin, Y., Hachez, Y., Nesterov, Y., Van Dooren, P.: Optimization problems over positive pseudopolynomial matrices. *SIAM J. Matrix Anal. Appl.* **25**(1), 57–79 (2003)
7. Hall, G.: Engineering and business applications of sum of squares polynomials. arXiv preprint [arXiv:1906.07961](https://arxiv.org/abs/1906.07961) (2019)
8. Henrion, D., Lasserre, J.B.: Convergent relaxations of polynomial matrix inequalities and static output feedback. *IEEE Trans. Autom. Control* **51**(2), 192–202 (2006)
9. Hoffman, M.E., Withers, W.D.: Generalized Chebyshev polynomials associated with affine Weyl groups. *Trans. Am. Math. Soc.* **308**(1), 91–104 (1988)
10. Hol, C.W., Scherer, C.W.: Sum of squares relaxations for polynomial semidefinite programming. In: *Proc. Symp. on Mathematical Theory of Networks and Systems (MTNS)*, Leuven, Belgium, Citeseer (2004)
11. Kojima, M.: Sums of squares relaxations of polynomial semidefinite programs. *Inst. of Technology* (2003)
12. Kojima, M., Muramatsu, M.: An extension of sums of squares relaxations to polynomial optimization problems over symmetric cones. *Math. Program.* **110**(2), 315–336 (2007)
13. Nesterov, Y.: Squared functional systems and optimization problems. In: *High performance optimization*, Springer, pp 405–440 (2000)
14. Nesterov, Y., Nemirovskii, A.: Interior-point polynomial algorithms in convex programming. *Stud. Appl. Math., SIAM* (1994)
15. Papp, D., Alizadeh, F.: Semidefinite characterization of sum-of-squares cones in algebras. *SIAM J. Optim.* **23**(3), 1398–1423 (2013)
16. Papp, D., Yildiz, S.: Sum-of-squares optimization without semidefinite programming. *SIAM J. Optim.* **29**(1), 822–851 (2019)
17. Skajaa, A., Ye, Y.: A homogeneous interior-point algorithm for nonsymmetric convex conic optimization. *Math. Program.* **150**(2), 391–422 (2015)
18. Subramanian, R., Bhagwat, K.: On a theorem of Wigner on products of positive matrices. *Proceedings of the Indian Academy of Sciences-Section A Part 3, Mathematical Sciences* **88**(1):31–34 (1979)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.