CORE DISCUSSION PAPER
2003/34

OPTIMIZATION PROBLEMS OVER NON-NEGATIVE POLYNOMIALS WITH INTERPOLATION CONSTRAINTS

Yvan HACHEZ[1] and Yurii NESTEROV[1]

April 2003

Abstract

Optimization problems over several cones of non-negative polynomials are described; we focus on linear constraints on the coefficients that represent interpolation constraints. For these problems, the complexity of solving the dual formulation is shown to be almost independent of the number of constraints, provided that an appropriate preprocessing has been performed. These results are also extended to non-negative matrix polynomials and to interpolation constraints on the derivatives.

Keywords: convex optimization, non-negative polynomials, interpolation constraints.

[1]CORE and INMA, Université catholique de Louvain, Belgium. E-mail: {hachez,nesterov}@core.ucl.ac.be
A research fellowship from the Belgian National Fund for Scientific Research is gratefully acknowledged by the first author.
This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

Non-negative polynomials are natural objects to model various engineering problems. Among the most representative applications are the filter design problems [1, 2, 5]. Recently, self-concordant barriers for several cones of non-negative polynomials have been proposed in the literature [11]. They are usually based on results dating back to the beginning of the *20*th century [9]. Indeed, these cones and their properties were extensively studied by several well-known mathematicians (Fejér, Riesz, Toeplitz, Markov, . . .), as testified by their correspondences and papers.

Nowadays, convex optimization techniques allow us to efficiently treat these cones, which are parametrized by semidefinite matrices [12, 15]. Although general semidefinite programming solvers could be used to solve the associated problems, the inherent structure of these polynomial problems must be exploited to derive much more efficient algorithms [1, 6]. They are based on the matrix structure that shows up in the dual problem. In particular, solving the standard conic formulation on cones of non-negative polynomials using the dual matrix structure has been studied in [6].

In this article, we consider particular conic formulations, of which the linear constraints are interpolations constraints. Indeed, natural linear constraints on the coefficients of a polynomial are obtained as interpolation conditions on the polynomial or its derivatives; each of them has an unambiguous interpretation. We show that the associated optimization problems can be solved very efficiently in a number of flops almost independent of the polynomial degree. Moreover, these formulations have some interesting properties that are worth pointing out.

In Section 2, we remind the reader of the characterization of non-negative scalar polynomials using the cone of positive semidefinite matrices. This step is of paramount importance before formulating the conic convex problems of interest, i.e., minimizing a linear function of the coefficients of a non-negative polynomial subject to interpolation constraints, see Section 3. Under mild assumptions, these problems can be solved very efficiently as described in Section 4. Although non-negative matrix polynomials can also be characterized using positive semidefinite matrices, closed formulas and nice interpretations are more difficult to obtain. Section 5 shows how to extend our previous results to this new setting. In Section 6, interpolation constraints on the derivatives are considered. Although we could have started this paper with the general setting (non-negative matrix polynomials with general interpolation constraints), this general approach and the associated notation would have shadowed most of the basic ideas underlying our results.

Notation The optimization problems considered in this article are assumed to be stated in terms of appropriate scalar products defined over the space of complex matrices. For any couple of matrices X and Y, let us set their Frobenius scalar product as follows

$$\langle X, Y\rangle_F \doteq \mathrm{Re}(\mathrm{Trace}\, XY^*) \equiv \mathrm{Re}\sum_{i=1}^{m}\sum_{j=1}^{n} x_{i,j}\overline{y}_{i,j}, \tag{1}$$

where $\{x_{i,j}\}_{i,j}$ and $\{y_{i,j}\}_{i,j}$ are the scalar entries of the matrices X and Y, respectively. Both matrices must have the same dimension $m \times n$, but they are not necessarily square. The above definition can thus be applied to vectors. Since this scalar product induces the Frobenius norm, i.e. $\|X\|_F^2 = \langle X, X\rangle_F$, it is called the Frobenius scalar product in what follows. It also follows from the definition that

$$\langle X, Y\rangle_F = \langle \mathrm{Re}(X), \mathrm{Re}(Y)\rangle + \langle \mathrm{Im}(X), \mathrm{Im}(Y)\rangle$$

where $\langle\cdot,\cdot\rangle$ stands for the standard scalar product of matrices, i.e. $\langle X, Y\rangle = \mathrm{Trace}\, XY^*$. Positive semidefiniteness of a matrix Y is denoted by $Y \succeq 0$. The sets of positive semidefinite real symmetric and complex Hermitian matrices (of order n) are denoted by $\mathcal{S}_n^+$ and $\mathcal{H}_n^+$, respectively. The column vector

$$\pi_n(s) = \begin{bmatrix}1 & s & \cdots & s^n\end{bmatrix}^T,$$

with $s = x$ or z, is often used to represent a polynomial by its coefficients. The (block) diagonal matrix defined by the (block) vector y is denoted by $\mathrm{diag}(y)$. The complex unit is written as $\jmath$, i.e. $\jmath^2 = -1$. The elements of the canonical basis are written as $\{e_k\}_k$, i.e. $I_n = \begin{bmatrix}e_0 & \dots & e_{n-1}\end{bmatrix}$ is the identity matrix of size n.

2 Non-negative polynomials

Let us summarize a few facts about non-negative polynomials. First of all, the characterization of such polynomials depends on the curve of the complex plane on which they are defined. These curves are typically the real axis $\mathbb{R}$, the unit circle $e^{\jmath\mathbb{R}}$ or the imaginary axis $\jmath\mathbb{R}$. Optimization problems on the latter curve are not considered in what follows; they can be reduced in a straightforward manner to optimization problems on the real line. The set of non-negative polynomials on any of these three curves is clearly a convex cone $\mathcal{K}$, i.e.

$$\mathcal{K} + \mathcal{K} \subseteq \mathcal{K}, \quad \alpha\mathcal{K} \subseteq \mathcal{K}, \quad \forall \alpha \geq 0$$

In this article, this special structure is used to formulate various optimization problems in conic form, based on interpolation constraints. Let us now examine the cones of interest and their duals.

2.1 Real line

Denote the cone of polynomials (of degree $2n$) non-negative on the whole real line $\mathbb{R}$ by

$$\mathcal{K}_{\mathbb{R}} = \{p \in \mathbb{R}^{2n+1} : p(x) = \sum_{k=0}^{2n} p_k x^k \geq 0, \forall x \in \mathbb{R}\}$$

and define the inner product between two real vectors $p = \begin{bmatrix} p_0, & \dots, & p_{2n} \end{bmatrix}^T$ and $q = \begin{bmatrix} q_0, & \dots, & q_{2n} \end{bmatrix}^T$ by $\langle p, q \rangle_{\mathbb{R}} \doteq \sum_{k=0}^{2n} p_k q_k = \langle p, q \rangle$. As a direct consequence of Markov-Lukács Theorem, the cone $\mathcal{K}_{\mathbb{R}}$ can be characterized as follows [11].

Theorem 2.1. *A polynomial $p(x) = \sum_{k=0}^{2n} p_k x^k$ is non-negative on the real axis if and only if there exists a positive semidefinite symmetric matrix $Y = \{y_{ij}\}_{i,j=0}^{n}$ such that ($y_{ij} = 0$ for i or j outside their definition range)*

$$p_k = \sum_{i+j=k} y_{ij}, \quad \textit{for } k = 0, \dots, 2n. \tag{2}$$

Remark. Identities (2) can be rewritten as $p(x) = \langle Y\pi_n(x), \pi_n(x) \rangle$.

Remember that the dual cone $\mathcal{K}_{\mathbb{R}}^{\star}$ is defined by

$$\mathcal{K}_{\mathbb{R}}^{\star} = \{s \in \mathbb{R}^{2n+1} : \langle p, s \rangle_{\mathbb{R}} \geq 0, \quad \forall p \in \mathcal{K}_{\mathbb{R}}\}.$$

Let $H(s)$ be the Hankel matrix defined by the vector $s \in \mathbb{R}^{2n+1}$, i.e.,

$$H(s) = \begin{bmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & s_{2n-1} \\ s_n & \cdots & s_{2n-1} & s_{2n} \end{bmatrix}. \tag{3}$$

Then the cone dual to $\mathcal{K}_{\mathbb{R}}$ is characterized by $H(s) \succeq 0$, i.e.,

$$\mathcal{K}_{\mathbb{R}}^{\star} = \{s \in \mathbb{R}^{2n+1} : H(s) \succeq 0\}.$$

The operator dual to $H(\cdot)$ allows us to write (2) as $p = H^*(Y)$, which means that

$$p_k = \langle Y, H(e_k) \rangle, \quad k = 0, \ldots, 2n.$$

Let us now interpret the primal and dual objects. Given $p \in \operatorname{int} \mathcal{K}_{\mathbb{R}}$, there exists a positive definite matrix Y such that $p = H^*(Y)$ and such that its inverse is a Hankel matrix, say $Y = H(s)^{-1}$ [11]. Remember that any positive definite Hankel matrix $H(s)$ admits a Vandermonde factorization [3]

$$Y^{-1} = H(s) = V \operatorname{diag}(\{p_i\}_{i=0}^{n})^{-1} V^T \tag{4}$$

where $p_i > 0, \forall i$ and V is a nonsingular Vandermonde matrix defined by the nodes $\{x_i\}_{i=0}^{n}$, i.e.

$$V = \begin{bmatrix} 1 & \ldots & 1 \\ x_0 & \ldots & x_n \\ \vdots & & \vdots \\ x_0^n & \ldots & x_n^n \end{bmatrix}. \tag{5}$$

Let L^T be the inverse of V, i.e. $L^T V = I_{n+1}$. It is well known that the rows of L^T are the coefficients of the Lagrange polynomials $\{l_i(x)\}_{i=0}^{n}$ associated with the distinct points $\{x_i\}_{i=0}^{n}$:

$$l_i(x_j) = \langle Le_i, \pi_n(x_j) \rangle = \delta_{ij}, \quad 0 \le i, j \le n,$$

where δ_{ij} is the Kronecker delta. Therefore, we obtain an explicit expression of the positive definite matrix Y in terms of the Lagrange polynomials

$$Y = L \operatorname{diag}(\{p_i\}_{i=0}^{n}) L^T. \tag{6}$$

This representation is equivalent to the decomposition of $p(x)$ as a weighted sum of $n+1$ squared Lagrange polynomials

$$p(x) = \sum_{i=0}^{n} p(x_i)(l_i(x))^2.$$

To see this, we plug the expression (6) of Y into the identity $p(x) =$

$\langle Y\pi_n(x), \pi_n(x)\rangle$, which can subsequently be rewritten as

$$p(x) = \langle L\,\mathrm{diag}(\{p_i\}_{i=1}^{n+1})L^T\pi_n(x), \pi_n(x)\rangle = \sum_{i=0}^{n} p_i \langle Le_ie_i^T L^T, \pi_n(x)\pi_n(x)^T\rangle$$

$$= \sum_{i=0}^{n} p_i(\langle Le_i, \pi_n(x)\rangle)^2 = \sum_{i=0}^{n} p_i(l_i(x))^2$$

As $\{l_i(x)\}_{i=0}^n$ are the Lagrange polynomials associated with the points $\{x_i\}_{i=0}^n$, it is straightforward to check that $p_i = p(x_i), \forall i$.

2.2 Unit circle

On the unit circle, the non-negative polynomials of interest are the trigonometric polynomials. Remember that a trigonometric polynomial of degree n has the form

$$p(\theta) = \sum_{k=0}^{n}[a_k \cos(k\theta) + b_k \sin(k\theta)], \quad \theta \in [0, 2\pi]. \tag{7}$$

where $\{a_k\}_{k=0}^n$ and $\{b_k\}_{k=0}^n$ are two sets of real coefficients. Without loss of generality, we can assume that $b_0 = 0$.

If we define the complex coefficients $\{p_k\}_{k=0}^n$ as

$$p_k = a_k + \jmath b_k, \quad k = 0, \ldots, n,$$

the pseudo-polynomial

$$p(z) = \langle p, \pi_n(z)\rangle_F = \mathrm{Re}(\sum_{k=0}^{n} p_k z^{-k}), \quad |z| = 1, \tag{8}$$

evaluated on the unit circle is equivalent to trigonometric polynomial (7). Therefore, we can use either (7) or (8) to represent the *same* mathematical object.

Denote the cone of trigonometric polynomials (of degree n) non-negative on the unit circle by

$$\mathcal{K}_{\mathbb{C}} = \{p \in \mathbb{R} \times \mathbb{C}^n : \langle p, \pi_n(z)\rangle_F \geq 0, z = e^{j\theta}, \theta \in [0, 2\pi)\}.$$

and define the inner product between two vectors $p = \begin{bmatrix} p_0, & \ldots, & p_n \end{bmatrix}^T \in \mathbb{R} \times \mathbb{C}^n$ and $q = \begin{bmatrix} q_0, & \ldots, & q_n \end{bmatrix}^T \in \mathbb{R} \times \mathbb{C}^n$ by $\langle p, q\rangle_{\mathbb{C}} \doteq \langle p, q\rangle_F$. As a direct consequence of Fejér-Riesz Theorem, see e.g., [14, Part 6, Problems 40 and 41], this cone can be characterized as follows [11].

Theorem 2.2. *A trigonometric polynomial $p(z) = \langle p, \pi_n(z)\rangle_F$ is non-negative on the unit circle if and only if there exists a positive semidefinite Hermitian matrix $Y = \{y_{i,j}\}_{i,j=0}^n$ such that ($y_{i,j} = 0$ for i or j outside their definition range)*

$$p_k = \begin{cases} \sum_{i-j=0} y_{i,j}, & k = 0 \\ 2\sum_{i-j=k} y_{i,j}, & k = 1, \dots, n \end{cases} \tag{9}$$

Remark. As before, identities (9) can be rewritten using the vector $\pi_n(z)$, i.e. $p(z) = \langle Y\pi_n(z), \pi_n(z)\rangle$.

By definition, the cone dual to $\mathcal{K}_{\mathbb{C}}$ is the set of vectors $s \in \mathbb{R} \times \mathbb{C}^n$ satisfying the inequalities

$$\langle p, s\rangle_{\mathbb{C}} \geq 0, \quad \forall p \in \mathcal{K}_{\mathbb{C}}.$$

Let $T(s)$ be the Hermitian Toeplitz matrix defined by the vector $s \in \mathbb{R}\times\mathbb{C}^n$, i.e.,

$$T(s) = \begin{bmatrix} s_0 & \bar{s}_1 & \cdots & \bar{s}_n \\ s_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bar{s}_1 \\ s_n & \ddots & s_1 & s_0 \end{bmatrix}. \tag{10}$$

Then the cone dual to $\mathcal{K}_{\mathbb{C}}$ is characterized by $T(s) \succeq 0$, i.e.

$$\mathcal{K}_{\mathbb{C}}^{\star} = \{s \in \mathbb{R} \times \mathbb{C}^n : T(s) \succeq 0\}.$$

The operator dual to $T(\cdot)$ allows us to write (9) as $p = T^*(Y)$, which means that

$$p_k = \langle Y, T_k\rangle, \quad k = 0, \dots, n$$

where the matrices $\{T_k\}_{k=0}^n$ are defined by the identity $T(s) = \frac{1}{2}\sum_{k=0}^n (T_k s_k + T_k^T \overline{s}_k), \forall s \in \mathbb{R} \times \mathbb{C}^n$.

As before, a better understanding of the primal and dual objects is obtained by considering the decomposition of $p(z) \in \operatorname{int} \mathcal{K}_{\mathbb{C}}$ as a weighted sum of squared Lagrange polynomials.

3 The optimization problem

The problem of optimizing over the cone of non-negative polynomials, subject to linear constraints on the coefficients of these polynomials, has already been studied by the authors in a wider framework [6]. Remember that this class of problems is exactly the standard *conic formulation* introduced in [12]. In this section, we now focus on the particular case of scalar polynomials constrained by interpolation constraints. The consequent structures of the primal and dual problems lead to efficient algorithms for solving such problems.

3.1 Real line

Several important optimization problems on the real line can be formulated as the following *primal* problem

$$\begin{array}{ll} \min & \langle c, p\rangle \\ \text{s.t.} & \langle a_i, p\rangle = b_i, \quad i = 0, \dots, k-1, \\ & p \in \mathcal{K}_{\mathbb{R}}, \end{array} \tag{11}$$

where the matrix of constraints $A = \{a_i\}_{i=0}^{k-1} \in \mathbb{R}^{k\times(2n+1)}$ is a full row rank matrix. Clearly, the constraints $Ap = b$ are linear constraints on the coefficients of the polynomial $p(x) = \sum_{i=0}^{2n} p_i x^i$ whereas the constraint $p \in \mathcal{K}_{\mathbb{R}}$ is semi-infinite. Note that the number k of linear constraints must satisfy $1 \leq k \leq 2n+1$. Moreover, if $k = 2n+1$, (11) is clearly not an optimization problem.

From a computational point of view, the problem dual to (11) has a considerable advantage over its primal counterpart. It reads as follows

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & s + \sum_{i=0}^{k-1} a_i y_i = c, \\ & s \in \mathcal{K}_{\mathbb{R}}^{\star}. \end{array} \tag{12}$$

Since its constraints are equivalent to $H(c - A^T y) \succeq 0$, the Hankel structure allows us to solve this dual problem efficiently [6].

Using Theorem 2.1, the primal optimization problem (11) can also be recast as a semidefinite programming problem :

$$\begin{array}{ll} \min & \langle H(c), Y\rangle \\ \text{s.t.} & \langle H(a_i), Y\rangle = b_i, \quad i = 0, \dots, k-1, \\ & Y \in \mathcal{S}_+^{n+1}. \end{array}$$

Let us now focus on interpolation constraints. Clearly, an interpolation constraint on a polynomial p is a linear constraint :

$$p(x_i) = \langle p, \pi_{2n}(x_i)\rangle = b_i.$$

Assume that all linear constraints of (11) are interpolation constraints, i.e.

$$\langle a_i, p \rangle \doteq \langle \pi_{2n}(x_i), p \rangle = b_i, \quad i = 0, \dots, k-1. \tag{13}$$

Then the dual problem (12) is equivalent to

$$\begin{array}{ll} \max & \langle b, y \rangle \\ \text{s. t.} & H(c) - \sum_{i=0}^{k-1} y_i H(\pi_{2n}(x_i)) \succeq 0. \end{array}$$

As the Hankel structure satisfies

$$H(\pi_{2n}(x)) = \pi_n(x)\pi_n(x)^T, \quad \forall x \in \mathbb{R},$$

we finally obtain the following formulation

$$\begin{array}{ll} \max & \langle b, y \rangle \\ \text{s. t.} & H(c) - V \operatorname{diag}(y) V^T \succeq 0, \end{array} \tag{14}$$

where the *Vandermonde matrix* V is defined by the nodes $\{x_0, \dots, x_{k-1}\}$, i.e.

$$V = \begin{bmatrix} 1 & \dots & 1 \\ x_0 & \dots & x_{k-1} \\ \vdots & & \vdots \\ x_0^n & \dots & x_{k-1}^n \end{bmatrix}.$$

Assumption 3.1. The components of the vector b are strictly positive.

Remark. Since we work with non-negative polynomials, this assumption is not restrictive. If there exists an integer i such that $b_i = 0$, one can factorize $p(x)$ as $p(x) = \tilde{p}(x)(x - x_i)^2$ and rewrite the optimization problem using the polynomial $\tilde{p}(x)$.

3.2 Unit circle

Several important optimization problems on the unit circle can be formulated as the following *primal* problem

$$\begin{array}{ll} \min & \langle c, p \rangle_{\mathbb{C}} \\ \text{s. t.} & \langle a_i, p \rangle_{\mathbb{C}} = b_i, \quad i = 0, \dots, k-1, \\ & p \in \mathcal{K}_{\mathbb{C}}, \end{array} \tag{15}$$

with linearly independent constraints. From a computational point of view, the problem dual to (15) has again a considerable advantage over its primal counterpart. This dual problem reads as follows

$$\begin{array}{ll} \max & \langle b, y \rangle \\ \text{s.t.} & s + \sum_{i=0}^{k-1} y_i a_i = c, \\ & s \in \mathcal{K}_{\mathbb{C}}^{\star}. \end{array} \tag{16}$$

As in the real line setting, one can use the Toeplitz structure of its constraints to get fast algorithms. Using Theorem 2.2, the primal optimization problem (15) can be reformulated as the semidefinite programming problem

$$\begin{array}{ll} \min & \langle T(c), Y \rangle \\ \text{s.t.} & \langle T(a_i), Y \rangle = b_i, \quad i = 0, \ldots, k-1, \ . \\ & Y \in \mathcal{H}_{+}^{n+1}. \end{array}$$

An interpolation constraint on the trigonometric polynomial p corresponds to

$$p(\theta_i) = \sum_{k=0}^{n} [a_k \cos(k\theta_i) + b_k \sin(k\theta_i)] = b_i \geq 0, \quad \theta_i \in [0, 2\pi],$$

and is equivalent to the linear constraint

$$\langle a_i, p \rangle \doteq p(z_i) = \langle p, \pi_n(z_i) \rangle = b_i, \quad z_i = e^{\jmath \theta_i}. \tag{17}$$

Note that the identity

$$T(\pi_{2n}(z)) = \pi_n(z) \pi_n(z)^*, \quad \forall z \in \mathbb{T},$$

holds for the Toeplitz structure. If all linear constraints of (15) are interpolation constraints, the dual can therefore be written as

$$\begin{array}{ll} \max & \langle b, y \rangle \\ \text{s.t.} & T(c) - V \operatorname{diag}(y) V^* \succeq 0, \end{array} \tag{18}$$

where the *Vandermonde matrix* V is defined by the points $\{z_0, \ldots, z_{k-1}\}$, i.e.

$$V = \begin{bmatrix} 1 & \ldots & 1 \\ z_0 & \ldots & z_{k-1} \\ \vdots & & \vdots \\ z_0^n & \ldots & z_{k-1}^n \end{bmatrix}.$$

As before, we make the next non-restrictive assumption.

Assumption 3.2. The components of the vector b are strictly positive.

Table 1: Interiors of primal and dual cones

	$\mathcal{K} = \mathcal{K}_{\mathbb{R}}$	$\mathcal{K} = \mathcal{K}_{\mathbb{C}}$
$p \in \operatorname{int} \mathcal{K}$	$p(t) > 0, \forall t \in \mathbb{R}$ and $p_{2n} > 0$	$p(z) > 0, \forall \lvert z \rvert = 1$
$s \in \operatorname{int} \mathcal{K}^{\star}$	$H(s)$ is positive definite	$T(s)$ is positive definite

4 Solving the optimization problem

In this section, we focus on problems with interpolation conditions, see (13) and (17). We discuss the interpretation of the so-called "strict feasibility" assumption in our context of polynomials. Then we give the explicit solution of three particular optimization problems (one interpolation constraint, two interpolation constraints, property of the objective function). In the general setting, we show that solving the dual problem can be very efficiently done, provided that strict feasibility holds.

4.1 Strict feasibility

The standard assumption on the primal and dual problems is the so-called "strict feasibility" assumption. This assumption is necessary in order to properly define the primal and dual central-paths and thus to solve our pair of primal and dual problems [10]. Moreover, it ensures that the optimal values of both problems coincide, which is an important property to solve our class of problem efficiently.

Assumption 4.1 (Strict feasibility). There exist points $\tilde{p} \in \operatorname{int} \mathcal{K}$, $\tilde{s} \in \operatorname{int} \mathcal{K}^{\star}$ and $\tilde{y} \in \mathbb{R}^k$ that satisfy the following linear systems

$$\begin{aligned} \langle a_i, \tilde{p} \rangle &= b_i, \quad 0 \leq i \leq k-1, \\ \hat{s} + \sum_{i=0}^{k-1} a_i \tilde{y}_i &= c. \end{aligned}$$

As mentioned in Table 1, the interiors of the primal and dual cones are characterized in terms of polynomials and structured matrices, respectively. However, our particular problem classes allow us to further discuss the interpretation of the previous assumption. More specifically, we shall see that some information about strict feasibility of our problems is known in advance.

Real line

First, we analyze the strict feasibility of the primal constraints. If the number of interpolation points is less or equal to $n+1$, i.e. $k \leq n+1$, it is clear that there exists a strictly positive polynomial $\tilde{p}$ such that $A\tilde{p} = b$. Assume that $k = n+1$ and let $\{l_i(x)\}_{i=0}^{n}$ be the set of Lagrange polynomials of degree n associated with the interpolation points. By definition, these polynomials satisfy the identities

$$l_i(x_j) = \delta_{ij}, \quad 0 \leq i, j \leq n,$$

where δ_{ij} is the well-known Kronecker delta. The polynomial

$$\tilde{p}(x) = \sum_{i=0}^{n} b_i (l_i(x))^2$$

clearly satisfies all our interpolation constraints and belongs to $\operatorname{int} \mathcal{K}_{\mathbb{R}}$. For the case $k < n+1$, we can add $n+1-k$ "extra" interpolation constraints and check that the (original) primal problem is always strictly feasible. If the number of interpolation points is strictly greater that $n+1$, we cannot say anything in advance about primal strict feasibility.

Let us now analyze the strict feasibility of the dual constraints. Because of the structure of our interpolation constraints, the interior of the dual space is the set of vectors $s = c - A^T y$ such that

$$H(s) = H(c - A^T y) = H(c) - \sum_{i=0}^{k-1} y_i \pi_n(x_i) \pi_n(x_i)^T \succ 0.$$

If $k \geq n+1$, we conclude from this inequality that there always exists $s = c - A^T y \in \operatorname{int} \mathcal{K}_{\mathbb{R}}^{\star}$. Another simple situation arises when $c \in \operatorname{int} \mathcal{K}_{\mathbb{R}}^{\star}$, i.e $H(c) \succ 0$. Then the dual problem is always strictly feasible. For instance, this situation occurs when minimizing the integral of the polynomial $p(x)$ on a finite interval $I \subseteq (-\infty, +\infty)$:

$$\langle c, p \rangle = \int_I p(x) dx = \sum_{i=0}^{2n} p_i \left(\int_I x^i dx \right),$$

subject to interpolation constraints. This situation is frequent in practice and one easily checks that $c \in \operatorname{int} \mathcal{K}_{\mathbb{R}}^{\star}$ in that case. Indeed, the inner product $\langle c, p \rangle$ is positive for all non-zero $p \in \mathcal{K}_{\mathbb{R}}$.

Remark. If the dual problem is strictly feasible, one can always reformulate the problem in order to ensure that $c \in \operatorname{int} \mathcal{K}_{\mathbb{R}}^{\star}$.

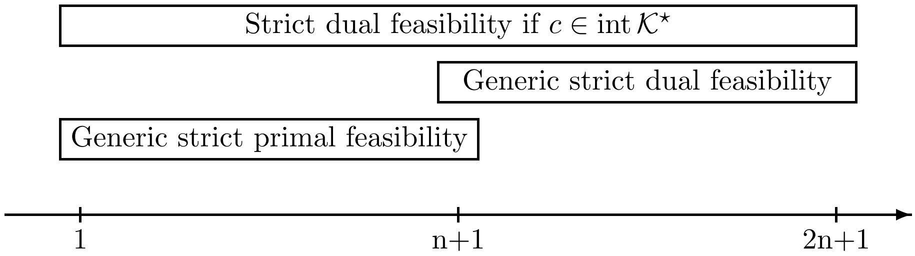

Figure 1: Generic strict feasibility as a function of the number of interpolation constraints

We have summarized our discussion on Figure 1. Let us point out a remarkable property of our class of problems, which is clearly exhibited on this figure. If the number of constraints is equal to $n+1$, both primal and dual problems are strictly feasible and this property is *independent* of the data. Except for that particular case, there usually exists a trade-off between strict primal and dual feasibility.

Unit circle

Using exactly the same argument, one can show that the primal problem is always strictly feasible if the number of interpolation constraints is less or equal to $n+1$. As in the real line, there exists a trade-off between strict feasibility of the primal and dual constraints unless $k = n+1$. If $c \in \operatorname{int} \mathcal{K}^\star_{\mathbb{C}}$, i.e $T(c) \succ 0$, the dual problem is always strictly feasible.

Therefore, the largest class of interpolation problems on non-negative polynomials (degree $2n$ or n, in the real line or unit circle setting, respectively) for which strict feasibility holds and does not depend on the interpolation points, satisfies the following assumption.

Assumption 4.2. The number k of interpolation constraints is less or equal to $n+1$ and the objective vector c satisfies $H(c) \succ 0$ (real line setting) or $T(c) \succ 0$ (unit circle setting), i.e. $c \in \operatorname{int} \mathcal{K}^\star$.

From now on, we focus on problems that fulfil this assumption. First, we consider several problems for which explicit solutions are easily computed from the data.

4.2 One interpolation constraint

Real line

Suppose that one wants to solve the primal problem

$$\min\{\langle c, p\rangle : p(\bar{x}) = b, p \in \mathcal{K}_{\mathbb{R}}\}.$$

The dual problem reads as follows

$$\begin{array}{ll} \max & by \\ \text{s. t.} & H(c) \succeq y\pi_n(\bar{x})[\pi_n(\bar{x})]^T. \end{array}$$

Without loss of generality, the scalar b is assumed to be equal to 1. Therefore, the optimal value of this problem

$$\frac{1}{\langle H(c)^{-1}\pi_n(\bar{x}), \pi_n(\bar{x})\rangle},$$

is equal to the optimal value of y. Using Assumption 4.2, the optimal vector p is thus given by

$$p = H^*(qq^T), \quad q = \frac{H(c)^{-1}\pi_n(\bar{x})}{\langle H(c)^{-1}\pi_n(\bar{x}), \pi_n(\bar{x})\rangle}.$$

One can check that

$$\begin{aligned} p(\bar{x}) = \langle \pi_{2n}(\bar{x}), p\rangle &= \langle \pi_n(\bar{x})\pi_n(\bar{x}), qq^T\rangle = (\langle \pi_n(\bar{x}), q\rangle)^2 = 1, \\ \langle c, p\rangle = \langle H(c), qq^T\rangle &= \frac{1}{\langle H(c)^{-1}\pi_n(\bar{x}), \pi_n(\bar{x})\rangle}. \end{aligned}$$

As p is feasible and the corresponding objective value $\langle c, p\rangle$ is equal to the dual optimal one, the polynomial $p(x) = \langle p, \pi_{2n}(x)\rangle$ is optimal.

Unit circle Let us now solve the primal problem

$$\min\{\langle c, p\rangle : p(\bar{z}) = \langle p, \pi_n(\bar{z})\rangle_{\mathbb{R}} = b, p \in \mathcal{K}_{\mathbb{C}}\}. \tag{19}$$

As in the real line setting, both primal and dual optimal solutions are computed explicitly by making use of Assumption 4.2. They are equal to :

$$\begin{aligned} y &= \frac{1}{\langle T(c)^{-1}\pi_n(\bar{z}), \pi_n(\bar{z})\rangle}, \\ p = T^*(qq^*), \quad q &= \frac{T(c)^{-1}\pi_n(\bar{z})}{\langle T(c)^{-1}\pi_n(\bar{z}), \pi_n(\bar{z})\rangle}. \end{aligned}$$

Example 4.1 (Moving average system, [13]). Let $h[n]$ be a discrete time signal and $\mathcal{H}(e^{\jmath\omega})$ be its Fourier transform. The function $|\mathcal{H}(e^{\jmath\omega})|^2$ is known as the *energy density spectrum* because it determines how the energy is distributed in frequency. Let us compute the signal that has the minimum energy

$$2\pi E = \int_{-\pi}^{\pi} |\mathcal{H}(e^{\jmath\omega})|^2 d\omega$$

and satisfies $|\mathcal{H}(e^{\jmath 0})| = 1$.

This is exactly an example of the problem class (19). Since $p(e^{\jmath\omega}) = |\mathcal{H}(e^{\jmath\omega})|^2$ is a trigonometric polynomial, $\int_{-\pi}^{\pi} p(e^{\jmath\omega})d\omega = p_0$. The vector c that defines the objective function is thus equal to $c = [1, 0, \dots, 0]^T = e_0$. The interpolation constraint is obviously defined by $\bar{z} = \pi_n(e^{\jmath 0}) = e$ and $b = 1$.

Therefore, the optimal primal solution is given by

$$p = T^*(qq^*), \quad q = \frac{[1, \dots, 1]^T}{n+1}.$$

and the corresponding Fourier transform $\mathcal{H}(e^{\jmath\omega})$ can be set to

$$\mathcal{H}(e^{\jmath\omega}) = \sum_{i=0}^{n} \frac{1}{n+1} e^{-\jmath\omega}.$$

Note that $|\mathcal{H}(e^{\jmath\omega})|^2$ is an approximation of a low-pass filter, see Figure 2. The corresponding signal is exactly the impulse response of the *moving average system* :

$$h[k] = \begin{cases} \frac{1}{n+1}, & 0 \le k \le n+1, \\ 0, & \text{otherwise.} \end{cases}$$

Since convolution of a discrete signal $x[n]$ with $h[n]$ returns a signal $y[n]$ such that

$$y[k] = \frac{1}{n+1} \sum_{l=0}^{n} x[k-l],$$

$y[n]$ is the "moving average" of $x[n]$.

4.3 Two interpolation constraints

Before investigating problems with two interpolation constraints, we need to solve explicitly a 2-dimensional semidefinite programming problem.

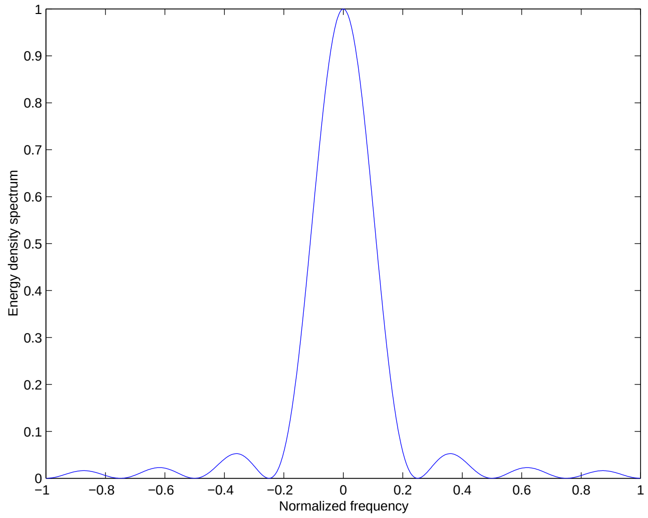

Figure 2: Energy density spectrum ($|\mathcal{H}(e^{\jmath\omega})|^2$ – n=7)

Proposition 4.1. *Let $b_0, b_1 \in \operatorname{int}\mathbb{R}_+$ and $\alpha, \gamma \in \mathbb{R}$ and $\beta \in \mathbb{C}$. The optimal value of the optimization problem*

$$\begin{array}{ll} \max & b_0 y_0 + b_1 y_1 \\ \text{s.t.} & \begin{bmatrix} \alpha & \overline{\beta} \\ \beta & \gamma \end{bmatrix} \succeq \begin{bmatrix} y_0 & 0 \\ 0 & y_1 \end{bmatrix} \end{array} \tag{20}$$

is reached at the optimal point

$$y_0 = \alpha - |\beta|\sqrt{\frac{b_1}{b_0}}, \qquad y_1 = \gamma - |\beta|\sqrt{\frac{b_0}{b_1}}$$

and it is equal to $b_0\alpha + b_1\gamma - 2|\beta|\sqrt{b_0 b_1}$.

Proof. The constraints are equivalent to

$$\alpha - y_0 \geq 0, \qquad \gamma - y_1 - \frac{|\beta|^2}{\alpha - y_0} \geq 0.$$

Maximizing the linear function $b_0 y_0 + b_1 y_1$ on this 2-dimensional convex region is straightforward (see Figure 3). Clearly, the system of equations

$$\frac{|\beta|^2}{(\alpha - y_0)^2} = \frac{b_0}{b_1}, \quad y_1 = \gamma - \frac{|\beta|^2}{\alpha - y_0},$$

provides us with the optimal point (y_0, y_1). □

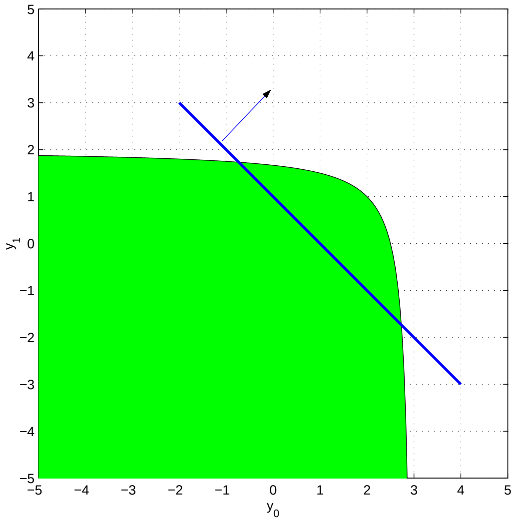

Figure 3: Feasibility region of (20) with $\alpha = 3, \beta = (1 + \jmath)/\sqrt{2}$ and $\gamma = 2$

Real line

If the number of interpolation constraints is equal to 2, the dual problem (12) is given by

$$\begin{array}{ll} \max & \langle b, y \rangle \\ \text{s.t.} & H(c) \succeq y_0 \pi_n(x_0)[\pi_n(x_0)]^T + y_1 \pi_n(x_1)[\pi_n(x_1)]^T \end{array}.$$

Equivalently, the dual constraint can be recast as

$$H(c) - \begin{bmatrix} \pi_n(x_0) & \pi_n(x_1) \end{bmatrix} \begin{bmatrix} y_0 & 0 \\ 0 & y_1 \end{bmatrix} \begin{bmatrix} \pi_n(x_0) & \pi_n(x_1) \end{bmatrix}^T \succeq 0.$$

Let us define the matrix $M_H(c; x_0, x_1)$ by

$$M_H(c; x_0, x_1) = \begin{bmatrix} \langle H(c)^{-1} \pi_n(x_0), \pi_n(x_0) \rangle & \langle H(c)^{-1} \pi_n(x_1), \pi_n(x_0) \rangle \\ \langle H(c)^{-1} \pi_n(x_0), \pi_n(x_1) \rangle & \langle H(c)^{-1} \pi_n(x_1), \pi_n(x_1) \rangle \end{bmatrix}.$$

If $\operatorname{diag}(y)$ is positive definite at the optimum, then the previous linear matrix inequality can be recast as

$$M_H(c; x_0, x_1)^{-1} \succeq \operatorname{diag}(y).$$

Indeed, this reformulation follows from the Schur complement formula. Otherwise, our hypothesis on the objective function, $c \in \operatorname{int} \mathcal{K}_{\mathbb{R}}^{\star}$, can be used so as to obtain the same reformulation. We delay the proof of that fact to the general setting, see Proposition 4.3.

Consequently, Proposition 4.1 allows us to solve our dual problem explicitly :

$$y_0 = \frac{1}{\det(M_H)} \left[\langle H(c)^{-1}\pi_n(x_1), \pi_n(x_1)\rangle - |\langle H(c)^{-1}\pi_n(x_0), \pi_n(x_1)\rangle| \sqrt{\frac{b_1}{b_0}} \right],$$

$$y_1 = \frac{1}{\det(M_H)} \left[\langle H(c)^{-1}\pi_n(x_0), \pi_n(x_0)\rangle - |\langle H(c)^{-1}\pi_n(x_0), \pi_n(x_1)\rangle| \sqrt{\frac{b_0}{b_1}} \right],$$

with $\det(M_H) = \det(M_H(c; x_0, x_1))$.

Our primal optimization problem can also be solved explicitly. To see this, define the vector $v = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T$ as the solution of the linear system

$$\begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} M_H(c; x_0, x_1) \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix} = \begin{bmatrix} \sqrt{b_0} \\ \sqrt{b_1} \end{bmatrix}$$

where $\sigma \in \{-1, +1\}$ is the sign of $\langle H(c)^{-1}\pi_n(x_0), \pi_n(x_1)\rangle$. Then the vector

$$p = H^*(qq^*), \quad q = H(c)^{-1} \begin{bmatrix} \pi_n(x_0) & \pi_n(x_1) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}$$

defines a non-negative polynomial $p(x) = (\langle q, \pi_n(x)\rangle)^2$ that satisfies $p(x_0) = b_0$ and $p(x_1) = b_1$. Indeed, we have

$$\begin{bmatrix} q(x_0) \\ q(x_1) \end{bmatrix} = \begin{bmatrix} \pi_n(x_0) & \pi_n(x_1) \end{bmatrix}^T q = \begin{bmatrix} \sqrt{b_0} \\ \sigma\sqrt{b_1} \end{bmatrix}$$

Moreover, the inner product $\langle c, p\rangle$ is equal to the optimal dual value : the vector p is thus optimal.

Unit circle

As in the real line setting, the dual problem can be rewritten as

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & M_T(c; z_0, z_1)^{-1} \succeq \operatorname{diag}(y) \end{array}$$

where

$$M_T(c; z_0, z_1) = \begin{bmatrix} \langle T(c)^{-1}\pi_n(z_0), \pi_n(z_0)\rangle & \langle T(c)^{-1}\pi_n(z_1), \pi_n(z_0)\rangle \\ \langle T(c)^{-1}\pi_n(z_0), \pi_n(z_1)\rangle & \langle T(c)^{-1}\pi_n(z_1), \pi_n(z_1)\rangle \end{bmatrix}.$$

The optimal dual solution is now equal to

$$y_0 = \frac{1}{\det(M_T)}\left[\langle T(c)^{-1}\pi_n(z_1), \pi_n(z_1)\rangle - |\langle T(c)^{-1}\pi_n(z_0), \pi_n(z_1)\rangle|\sqrt{\frac{b_1}{b_0}}\right],$$
$$y_1 = \frac{1}{\det(M_T)}\left[\langle T(c)^{-1}\pi_n(z_0), \pi_n(z_0)\rangle - |\langle T(c)^{-1}\pi_n(z_0), \pi_n(z_1)\rangle|\sqrt{\frac{b_0}{b_1}}\right],$$

with $\det(M_T) = \det(M_T(c; z_0, z_1))$. Let us define the vector $\begin{bmatrix} v_0 & v_1 \end{bmatrix}^T$ as the solution of the linear system

$$\begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix}^* M_T(c; z_0, z_1) \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix} = \begin{bmatrix} \sqrt{b_0} \\ \sqrt{b_1} \end{bmatrix}$$

where σ is equal to $e^{-\jmath \arg\langle T(c)^{-1}\pi_n(z_1), \pi_n(z_0)\rangle}$. The vector

$$p = T^*(qq^*), \quad q = T(c)^{-1}\begin{bmatrix} \pi_n(z_0) & \pi_n(z_1) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}$$

corresponds to a trigonometric polynomial $p(z) = |q(z)|^2$ that satisfies our interpolation constraints and such that $\langle c, p\rangle = \langle b, y\rangle$. This vector p is thus the (primal) optimal one.

4.4 More interpolation constraints ($k \le n+1$)

If Assumption 4.2 holds and $k \le n+1$, the previous analysis can always be carried out. We first focus on the unit circle setting and show the connection with spectral factorization of trigonometric polynomials. The real line problem is then solved using a similar methodology. Let us start with two preliminary results.

Preliminary results

Proposition 4.2. *Let $C \in \operatorname{int}\mathcal{H}_+^n$ be a positive definite matrix and $V = \begin{bmatrix} V_0 & V_1 \end{bmatrix} \in \mathbb{C}^{n\times n}$ be a nonsingular matrix. If the matrix $W = \begin{bmatrix} W_0 \\ W_1 \end{bmatrix}$ is the (left) inverse of V with compatible partitions, i.e. $\begin{bmatrix} W_0V_0 & W_0V_1 \\ W_1V_0 & W_1V_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$, then we have*

$$(V_1^* C^{-1} V_1)^{-1} = W_1 C W_1^* - W_1 C W_0^* (W_0 C W_0^*)^{-1} W_0 C W_1.$$

Proof. Let us apply the well-known *Schur complement identity*

$$\begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} I & 0 \\ GE^{-1} & I \end{bmatrix} \begin{bmatrix} E & 0 \\ 0 & H - GE^{-1}F \end{bmatrix} \begin{bmatrix} I & E^{-1}F \\ 0 & I \end{bmatrix}$$

to the matrix product

$$WCW^* = \begin{bmatrix} W_0CW_0^* & W_0CW_1^* \\ W_1CW_0^* & W_1CW_1^* \end{bmatrix}.$$

Clearly, we obtain that

$$WCW^* = \begin{bmatrix} I & 0 \\ (W_1CW_0^*)(W_0CW_0^*)^{-1} & I \end{bmatrix} \begin{bmatrix} W_0CW_0^* & 0 \\ 0 & W_1/W_0 \end{bmatrix} \begin{bmatrix} I & (W_0CW_0^*)^{-1}(W_0CW_1^*) \\ 0 & I \end{bmatrix}$$

with $W_1/W_0 = W_1CW_1^* - (W_1CW_0^*)(W_0CW_0^*)^{-1}(W_0CW_1^*)$. Because the matrix WCW^* is nonsingular (by assumption), we have

$$(WCW^*)^{-1} = \begin{bmatrix} I & -(W_0CW_0^*)^{-1}(W_0CW_1^*) \\ 0 & I \end{bmatrix} \begin{bmatrix} W_0CW_0^* & 0 \\ 0 & W_1/W_0 \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ -(W_1CW_0^*)(W_0CW_0^*)^{-1} & I \end{bmatrix}.$$

Hence, the lower right block of the identity

$$(WCW^*)^{-1} = V^*C^{-1}V = \begin{bmatrix} V_0^*C^{-1}V_0 & V_0^*C^{-1}V_1 \\ V_1^*C^{-1}V_0 & V_1^*C^{-1}V_1 \end{bmatrix}$$

is exactly equivalent to

$$V_1^*C^{-1}V_1^* = (W_1CW_1^* - (W_1CW_0^*)(W_0CW_0^*)^{-1}(W_0CW_1^*))^{-1}.$$

□

Proposition 4.3. *Let $C \in \operatorname{int} \mathcal{H}_+^n$ be a positive definite matrix and $V_1 \in \mathbb{C}^{n\times k}$ be a matrix with full column rank ($k \le n$). Then the linear matrix inequality*

$$C \succeq V_1 \operatorname{diag}(y) V_1^* \tag{21}$$

is equivalent to

$$(V_1^*C^{-1}V_1)^{-1} \succeq \operatorname{diag}(y). \tag{22}$$

Proof. If $k = n$, the proof is trivial. Indeed, both inequalities (21) and (22) are congruent. This congruence is defined by the nonsingular matrix V_1^{-1}. If $k < n$, Proposition 4.2 must be used. Let $V_0 \in \mathbb{C}^{n\times(n-k)}$ be a matrix such that $V = \begin{bmatrix} V_0 & V_1 \end{bmatrix} \in \mathbb{C}^{n\times n}$ is nonsingular. The (left) inverse of V is denoted by $W = \begin{bmatrix} W_0 \\ W_1 \end{bmatrix}$. If the rows of W are partitioned according to the partition of V, we have

$$WV = \begin{bmatrix} W_0V_0 & W_0V_1 \\ W_1V_0 & W_1V_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

The linear matrix inequality (21), which can be rewritten as

$$C - \begin{bmatrix} V_0 & V_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{diag}(y) \end{bmatrix} \begin{bmatrix} V_0^* \\ V_1^* \end{bmatrix} \succeq 0,$$

is thus equivalent to

$$\begin{bmatrix} W_0 \\ W_1 \end{bmatrix} C \begin{bmatrix} W_0^* & W_1^* \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{diag}(y) \end{bmatrix} \succeq 0 \tag{23}$$

by congruence. Because $W_0CW_0^*$ is positive definite by assumption, the previous inequality is equivalent to positive semidefiniteness of its Schur complement in (23),

$$W_1CW_1^* - (W_1CW_0^*)(W_0CW_0^*)^{-1}(W_0CW_1^*) \succeq \mathrm{diag}(y).$$

We complete the proof by making use of Proposition 4.2. □

Unit circle

Remember that the optimization problem of interest is

$$\begin{array}{ll} \min & \langle c, p\rangle_{\mathbb{R}} \\ \text{s.t.} & \langle p, \pi_n(z_i)\rangle_{\mathbb{R}} = b_i, \quad i = 0, \dots, k-1, \\ & p \in \mathcal{K}_{\mathbb{C}}. \end{array} \tag{24}$$

If the non-negative trigonometric polynomial $p(z)$ is written as a square by making use of an arbitrary spectral factor $q(z)$, i.e. $p(z) = |q(z)|^2$ or $p = T^*(qq^*)$, the primal optimization problem can be rewritten as

$$\begin{array}{ll} \min & \langle T(c)q, q\rangle \\ \text{s.t.} & \langle q, \pi_n(z_i)\rangle = \sqrt{b_i}e^{\jmath\theta_i}, \quad i = 0, \dots, k-1, \end{array} \tag{25}$$

where $\{\theta_i\}_{i=0}^{k-1}$ is a set of phases.

Define the vector σ by $\sigma_i = \sqrt{b_i} e^{\jmath\theta_i}, i = 0, \dots, k-1$ and the matrix M_T by

$$M_T(c; z_0, \dots, z_{k-1}) = \begin{bmatrix} \pi_n(z_0) & \cdots & \pi_n(z_{k-1}) \end{bmatrix}^* \\ T(c)^{-1} \begin{bmatrix} \pi_n(z_0) & \cdots & \pi_n(z_{k-1}) \end{bmatrix}.$$

As a function of σ, the optimal solution of (25) is equal to

$$q = T(c)^{-1} \begin{bmatrix} \pi_n(z_0) & \cdots & \pi_n(z_{k-1}) \end{bmatrix} M_T(c; z_0, \dots, z_{k-1})^{-1} \sigma \tag{26}$$

and the corresponding optimal value is

$$\langle T(c)q, q \rangle = \langle M_T(c; z_0, \dots, z_{k-1})^{-1} \sigma, \sigma \rangle.$$

Remark. A direct consequence of (26) is that the spectral factor $q(z)$ is decomposed as a sum of "Lagrange-like" polynomials :

$$q(z) = \langle q, \pi_n(z) \rangle_{\mathbb{C}} = \sum_{i=0}^{k-1} e^{\jmath\theta_i} \sigma_i l_i(z)$$

where $l_i(z_j) = \delta_{ij}, \forall i, j$.

Finally, the optimal solution of problem (25) is obtained by minimizing over the vector σ,

$$\begin{array}{ll} \min & \langle M_T(c; z_0, \dots, z_{k-1})^{-1} \sigma, \sigma \rangle \\ \text{s.t.} & |\sigma_i|^2 = b_i, \quad i = 0, \dots, k-1. \end{array} \tag{27}$$

If $m > 2$, an explicit solution is difficult to obtain easily from this new formulation. However, we can solve the semidefinite relaxation of problem (27) :

$$\begin{array}{ll} \min & \langle M_T^{-1}(z_0, \dots, z_{k-1}), X \rangle, \\ \text{s.t.} & \operatorname{diag}(X) = b, \\ & X \in \mathcal{H}_+^k, \end{array} \tag{28}$$

where $\operatorname{diag}(X)$ is the vector defined by the diagonal elements of X. In general, a quadratic problem of the form (27) is NP-hard to solve, see the Appendix. Nevertheless, the particular structure of the quadratic objective function yields an extremely interesting result.

Theorem 4.4. *If Assumption 4.2 holds, relaxation* (28) *of quadratically constrained quadratic problem* (27) *is exact.*

Proof. Using standard convex duality theory, the dual of problem (28) is

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & M_T^{-1}(z_0,\dots,z_{k-1}) \succeq \operatorname{diag}(y), \end{array} \tag{29}$$

which is exactly the dual of the original problem (24) :

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & T(c) \succeq \begin{bmatrix}\pi_n(z_0) & \dots & \pi_n(z_{k-1})\end{bmatrix} \operatorname{diag}(y) \begin{bmatrix}\pi_n(z_0) & \dots & \pi_n(z_{k-1})\end{bmatrix}^* . \end{array} \tag{30}$$

To see this, we define the matrix V_1 as $V_1 = \begin{bmatrix}\pi_n(z_0) & \dots & \pi_n(z_{k-1})\end{bmatrix}$ and we apply Proposition 4.3 with $C = T(c)$. Because the (dual) constraints of (29) and (30) are equivalent, both problems are identical.

By assumption the original problem (24) has no duality gap. Since both problems (24) and (28) have the same dual, the relaxation has also a zero duality gap. This last observation completes our proof. □

The optimal coefficients p can be obtained from the solution X of (28) via the identity

$$p = T^*(T(c)^{-1}V_1 M_T^{-1} X M_T^{-1} V_1^* T(c)^{-1})$$

where $V_1 = \begin{bmatrix}\pi_n(z_0) & \dots & \pi_n(z_{k-1})\end{bmatrix}$ and $M_T = M_T(c; z_0,\dots,z_{k-1})$. To see this, note that

$$\begin{aligned} \langle c, p\rangle &= \langle T(c), T(c)^{-1}V_1 M_T^{-1} X M_T^{-1} V_1^* T(c)^{-1}\rangle \\ &= \langle T(c)^{-1}V_1 M_T^{-1} X M_T^{-1} V_1^*, I\rangle \\ &= \langle V_1^* T(c)^{-1}V_1 M_T^{-1} X M_T^{-1}, I\rangle \\ &= \langle X, M_T^{-1}\rangle \end{aligned}$$

and that, for all i,

$$\begin{aligned} \langle p, \pi_n(z_i)\rangle =& \langle T(\pi_n(z_i)), T(c)^{-1}V_1 M_T^{-1} X M_T^{-1} V_1^* T(c)^{-1}\rangle \\ =& \langle \pi_n(z_i)\pi_n(z_i)^*, T(c)^{-1}V_1 M_T^{-1} X M_T^{-1} V_1^* T(c)^{-1}\rangle \\ =& \langle (\pi_n(z_i)^* T(c)^{-1}V_1 M_T^{-1}) X (M_T^{-1} V_1^* T(c)^{-1}\pi_n(z_i)), I\rangle \\ =& \langle e_i e_i^*, X\rangle = b_i. \end{aligned}$$

Real line

Remember that the optimization problem of interest is

$$\begin{array}{ll} \min & \langle c, p\rangle \\ \text{s.t.} & \langle p, \pi_{2n}(x_i)\rangle = b_i, \quad i = 0,\dots,k-1, \; . \\ & p \in \mathcal{K}_{\mathbb{R}}. \end{array} \tag{31}$$

If we use any complex spectral factor $q(x)$ of our unknown polynomial $p(x) = |q(x)|^2$ as a variable, the previous analysis can be carried out in the real line setting. It leads exactly to the same formulae *provided that* the following substitutions are performed :

1. $T(c)$ is replaced by $H(c)$;
2. the interpolation points $\{z_i\}_{i=0}^{k-1}$ are replaced by $\{x_i\}_{i=0}^{k-1}$;
3. the matrix $M_T(c; z_0, \ldots, z_{k-1})$ is replaced by its "Hankel counterpart"

$$[M_H(c; x_0, \ldots, x_{k-1})]_{ij} = \pi_n(x_i)^* H(c)^{-1} \pi_n(x_j).$$

Let us summarize the most important steps. First, the primal optimization problem (31) is reformulated as

$$\begin{array}{ll} \min & \langle H(c)q, q\rangle \\ \text{s.t.} & \langle q, \pi_n(x_i)\rangle = \sqrt{b_i} e^{\jmath\theta_i}, \quad i = 0, \ldots, k-1, \end{array} \tag{32}$$

which is equivalent to

$$\begin{array}{ll} \min & \langle M_H(c; x_0, \ldots, x_{k-1})^{-1}\sigma, \sigma\rangle \\ \text{s.t.} & |\sigma_i|^2 = b_i, \quad i = 0, \ldots, k-1. \end{array} \tag{33}$$

In practice, this last optimization problem is solved using the following relaxation

$$\begin{array}{ll} \min & \langle M_H^{-1}(x_0, \ldots, x_{k-1}), X\rangle, \\ \text{s.t.} & \operatorname{diag}(X) = b, \\ & X \in \mathcal{H}_+^k. \end{array} \tag{34}$$

As before, the structure of quadratic problem (33) leads to an exact semidefinite relaxation.

Theorem 4.5. *If Assumption 4.2 holds, relaxation* (34) *of quadratically constrained quadratic problem* (33) *is exact.*

Proof. Using standard convex duality theory, the dual of problem (28) is

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & M_H^{-1}(x_0, \ldots, x_{k-1}) \succeq \operatorname{diag}(y), \end{array} \tag{35}$$

which is exactly the dual of the original problem (31) :

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & H(c) \succeq \begin{bmatrix}\pi_n(x_0) & \ldots & \pi_n(x_{k-1})\end{bmatrix} \operatorname{diag}(y) \begin{bmatrix}\pi_n(x_0) & \ldots & \pi_n(x_{k-1})\end{bmatrix}^*. \end{array} \tag{36}$$

To see this, we define the matrix V_1 as $V_1 = \begin{bmatrix} \pi_n(z_0) & \dots & \pi_n(z_{k-1}) \end{bmatrix}$ and we apply Proposition 4.3 with $C = T(c)$ and V_1. Because the (dual) constraints of (35) and (36) are equivalent, both problems are identical.

By assumption the original problem (31) has no duality gap. Since both problems (31) and (34) have the same dual, the relaxation has also a zero duality gap. This last observation completes our proof. □

Complexity

The complexity of solving relaxation (34) or (28) is only a function of the desired accuracy ϵ and the number of interpolation constraints k. If Assumption 4.2 holds and if the original problem has been pre-processed, it can be solved in a number of iterations that does *not* depend on the degree n. Indeed, solving the dual problem (35) or (29) using a *standard* path-following scheme requires $\mathcal{O}(\sqrt{k}\log\frac{1}{\epsilon})$ Newton steps. At each iteration, computing the gradient and the Hessian of a barrier function of the type

$$f(y) = -\log\det(M^{-1} - \operatorname{diag}(y))$$

requires $\mathcal{O}(k^3)$ flops. Note that the pre-processing can be done via fast Hankel or Toeplitz solvers, see [8].

4.5 Still more interpolation constraints ($m > n + 1$)

If the number of interpolation constraints is strictly greater than $n + 1$, strict feasibility of the primal problem depends on the data. Therefore, a *general* procedure that solves efficiently the primal problem and uses the structure of the interpolation constraints is not likely to exist. Indeed, the primal problem might be infeasible ! Let us illustrate this fact by a simple example.

Example 4.2. Consider the set of polynomials of degree $2n = 4$, nonnegative on the real line, and four interpolation points $x = [-2, -1, 1, 2]$. The vector $b = [1, 1, 1, 1]$ gives a strictly feasible primal problem. Indeed, the polynomial $p(x) = \frac{1}{3}(x^4 - 5x^2 + 7)$ satisfies our interpolation constraints and belongs to $\operatorname{int}\mathcal{K}_{\mathbb{R}}$. If the vector b is equal to $[1, 10, 1, 1]$, the polynomial family that satisfies our interpolation constraints is $p(x;\lambda) = \frac{1}{4}((\lambda - 7)x^4 + 6x^3 + (29 - 5\lambda)x^2 - 24x + 4\lambda)$, $\lambda \in \mathbb{R}$. If $p(x;\lambda)$ belonged to $\operatorname{int}\mathcal{K}_{\mathbb{R}}$, λ would be greater than 7. As $p(\frac{5}{4};\lambda) = \frac{1}{1024}(-371\lambda - 3255), \forall\lambda > 0$, these data correspond to an infeasible primal set...

Of course, the dual structure can still be exploited to try reducing the computational cost. For instance, consider a problem on the unit circle with $m > n+1$ interpolation constraints. Clearly, the corresponding Vandermonde matrix V can be divided into a nonsingular square Vandermonde matrix V_0 and a rectangular one V_1

$$V = \begin{bmatrix} V_0 & V_1 \end{bmatrix}, \quad \det V_0 \neq 0.$$

If the dual vector is divided accordingly, the dual constraint can be recast as $T(c) \succeq V_0 \operatorname{diag}(y_0) V_0^* + V_1 \operatorname{diag}(y_1) V_1^*$. Since V_0 is nonsingular, it is equivalent to

$$V_0^{-1} T(c) V_0^{-*} - V_0^{-1} V_1 \operatorname{diag}(y_1) V_1^*) V_0^{-*} \succeq \operatorname{diag}(y_0).$$

Therefore, an appropriate preprocessing leads to the following dual constraint

$$\hat{C} - \hat{V} \operatorname{diag}(y_1) \hat{V}^* \succeq \operatorname{diag}(y_0).$$

Since the Toeplitz structure of the dual is lost, the resulting algorithm cannot use the underlying displacement operator nor a divide-and-conquer strategy to evaluate the gradient and the Hessian of the self-concordant barrier function. This strategy will thus be slower than the one designed in [6]

4.6 Property of the objective function

If $H(c)$ or $T(c)$ is not positive definite, the corresponding dual problem can sometimes be solved explicitly.

Real line If the vector c is such that $H(c)$ can be factorized as

$$H(c) = \begin{bmatrix} V & W \end{bmatrix} \begin{bmatrix} \operatorname{diag}(\lambda_v) & 0 \\ 0 & \operatorname{diag}(\lambda_w) \end{bmatrix} \begin{bmatrix} V^T \\ W^T \end{bmatrix} \tag{37}$$

where $V \in \mathbb{R}^{k \times (n+1)}$ is the Vandermonde matrix defined by the interpolation constraints and $W \in \mathbb{R}^{(n+1-k)\times(n+1)}$ is such that $\begin{bmatrix} V & W \end{bmatrix}$ is full rank, one can easily compute an explicit solution of the optimization. From a theoretical point of view, there exist vectors c such that the proposed factorization does not exist. From a computational point of view, it may also be difficult to compute accurately.

The dual constraints now reads as follows

$$\begin{bmatrix} \operatorname{diag}(\lambda_v - y) & 0 \\ 0 & \operatorname{diag}(\lambda_w) \end{bmatrix} \succeq 0.$$

If $\mathrm{diag}(\lambda_w)$ is not positive semidefinite, the dual optimization problem is infeasible and the primal problem is unbounded. Otherwise, the solution is obtained by setting the dual variables y_i to their upper bounds, i.e. $y = \lambda_v$. This provides us with either a lower bound or the exact value of the optimization problem, depending on whether the problem has a duality gap.

Unit circle The same factorization technique can be applied to $T(c)$, i.e.

$$T(c) = \begin{bmatrix} V & W \end{bmatrix} \begin{bmatrix} \mathrm{diag}(\lambda_v) & 0 \\ 0 & \mathrm{diag}(\lambda_w) \end{bmatrix} \begin{bmatrix} V^* \\ W^* \end{bmatrix}, \tag{38}$$

and leads to the same results and drawbacks.

5 Matrix polynomials

In this section, we show that most of the previous results still holds in the context of non-negative matrix polynomials. As before, these non-negative polynomials could be defined on the real line, on the imaginary axis and on the unit circle. To avoid redundancies, we only consider the cone of matrix polynomials non-negative on the real line, which is again denoted by $\mathcal{K}_{\mathbb{R}}$,

$$0 \preceq P(x) = \sum_{k=0}^{2n} P_k x^k, \ \forall x \in \mathbb{R}; \qquad P_k = P_k^* \in \mathbb{R}^{q \times q}, \forall k. \tag{39}$$

Theorem 2.1 can then be extended to the matrix case [6].

Theorem 5.1. *A matrix polynomial $P(x)$ is non-negative on the real axis if and only if there exists a positive semidefinite symmetric block matrix $Y = \{Y_{ij}\}_{i,j=0}^n$ such that ($Y_{ij} = 0$ for i or j outside their definition range)*

$$P_k = \sum_{i+j=k} Y_{ij}, \quad \textit{for } k = 0, \ldots, 2n. \tag{40}$$

As shown in [6], the dual cone is the set of Hermitian matrix coefficients $S = \begin{bmatrix} S_0 & S_1 & \ldots & S_{2n} \end{bmatrix}$ such that the corresponding block Hankel matrix

$$H(S) = \begin{bmatrix} S_0 & S_1 & \cdots & S_n \\ S_1 & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & S_{2n-1} \\ S_n & \cdots & S_{2n-1} & S_{2n} \end{bmatrix}, \tag{41}$$

is positive semidefinite, i.e. $\mathcal{K}_{\mathbb{R}}^{\star} = \{S : H(S) \succeq 0\}$.

5.1 The optimization problem

Using matrix interpolation constraints, our optimization problem (11) could be extended to

$$\begin{array}{ll} \min & \langle C, P\rangle \equiv \sum_{\ell=0}^{2n}\langle C_\ell, P_\ell\rangle \\ \text{s.t.} & P(x_i) = \sum_{\ell=0}^{2n} P_\ell x_i^\ell = B_i, \quad i = 0,\dots,k-1 \\ & P(x) \succeq 0, \ \forall x \in \mathbb{R} \end{array} . \tag{42}$$

where $\{B_i\}_{i=0}^{k-1}$ is a set of positive definite matrices. Its dual is readily seen to be equal to

$$\begin{array}{ll} \max & \langle B, Y\rangle \equiv \sum_{i=0}^{k-1}\langle B_i, Y_i\rangle \\ \text{s.t.} & S_\ell + \sum_{i=0}^{k-1} x_i^\ell Y_i = C_\ell, \qquad \ell = 0,\dots,2n \\ & H(S) \succeq 0, \end{array} . \tag{43}$$

5.2 Strict feasibility

As before, primal strict feasibility holds if the number k of matrix interpolation constraints is less or equal to $n+1$. To see this, consider $n+1$ distinct interpolation points $\{x_i\}_{i=0}^n$ and the associated Lagrange polynomials $\{L_i(x)\}_{i=0}^n$ of degree n. These polynomials are defined by the identities

$$L_j(x_i) = \delta_{ij} I_m, \qquad 0 \le i, j \le n.$$

Then the polynomial

$$P(x) = \sum_{i=0}^{n} L_i(x) P(x_i) L_i^T(x) = \sum_{i=0}^{n} L_i(x) B_i L_i^T(x)$$

can be rewritten as

$$P(x) = \langle L \operatorname{diag}(\{P(x_i)\}_{i=0}^{k-1}) L^T \Pi_n(x), \Pi_n(x)\rangle$$

where L is nonsingular and $\operatorname{diag}(\{P(x_i)\}_{i=0}^{k-1})$ is positive definite. By construction, we see that $P(x) \in \operatorname{int} \mathcal{K}_{\mathbb{R}}$ and $P(x_i) = B_i, \forall i$.

Since the dual constraints (43) are equivalent to

$$H(C) \succeq \sum_{i=0}^{k-1} \Pi_n(x_i) Y_i \Pi_n(x_i)^T,$$

the dual is strictly feasible if $k \ge n+1$.

Let us state the matrix counterpart of Assumption 4.2 for future use.

Assumption 5.1. The number k of interpolation constraints is less or equal to $n+1$ and the objective block vector C satisfies $H(C) \succ 0$.

Hereafter, we focus on problems satisfying this assumption.

5.3 One interpolation constraint

Let us consider a matrix interpolation problem with one constraint :

$$\begin{array}{ll} \min & \langle C, P\rangle \\ \text{s.t.} & P(\bar{x}) = \sum_{\ell=0}^{2n} P_\ell \bar{x}^\ell = B \succ 0, \\ & P(x) \succeq 0, \ \forall x \in \mathbb{R} \end{array} .$$

Without loss of generality, B is assumed to be the identity matrix, i.e. $B = I_m$. Using the dual matrix variable Y, the dual problem reads

$$\begin{array}{ll} \max & \langle I, Y\rangle \\ \text{s.t.} & H(C) \succeq \Pi_n(\bar{x}) Y \Pi_n(\bar{x})^T \end{array} .$$

Because $H(C) \succ 0$, a standard Schur complement approach shows that the optimal dual variable is

$$Y = [\Pi_n(\bar{x})^T H(C)^{-1} \Pi_n(\bar{x})]^{-1}.$$

The spectral factor

$$Q = H(C)^{-1}\Pi_n(\bar{x})[\Pi_n(\bar{x})^T H(C)^{-1} \Pi_n(\bar{x})]^{-1}$$

allows us to compute the optimal primal variable P

$$P(x) = Q(x)Q(x)^* \quad \Longleftrightarrow \quad P = H^*(QQ^*).$$

It is easy to check that this value of P is optimal, i.e.

$$\langle C, P\rangle = \sum_{\ell=0}^{2n} \langle C_\ell, P_\ell\rangle = \langle H(C)Q, Q\rangle = \langle I, [\Pi_n(\bar{x})^T H(C)^{-1} \Pi_n(\bar{x})]^{-1}\rangle = \langle I, Y\rangle$$

and

$$P(\bar{x}) = \Pi_n(\bar{x})^T QQ^* \Pi_n(\bar{x}) = I_m = B.$$

5.4 More interpolation constraints

If the number of matrix interpolation constraints is less or equal to $n+1$, we can again use an arbitrary spectral factor to get an efficient algorithm, the complexity of which mainly depends on k and m.

Indeed, let $Q(x)$ be an arbitrary spectral factor $Q(x)$ of our unknown polynomial $P(x)$, i.e $P(x) = Q(x)Q(x)^*$. Then the optimization problem can be rewritten as

$$\begin{array}{ll} \min & \langle H(C)Q, Q\rangle \\ \text{s.t.} & Q(x_i) = \sum_{\ell=0}^{2n} Q_\ell x_i^\ell = B_i^{1/2} U_i, \quad i = 0, \ldots, k-1 \end{array} \qquad (44)$$

where $\{U_i\}_{i=0}^{k-1}$ is a set of unitary matrices, i.e. $U_i^* U_i = I_m, \forall i$.

If the definition of M_H is adapted to the matrix case,

$$[M_H(C; x_0, \ldots, x_{k-1})]_{ij} = \Pi_n(x_i)^* H(C)^{-1} \Pi_n(x_j),$$

then the optimal solution of (44), written as a function of

$$U = \begin{bmatrix} U_0 \\ \vdots \\ U_{k-1} \end{bmatrix},$$

is equal to

$$Q = H(C)^{-1} \begin{bmatrix} \Pi_n(x_0) & \cdots & \Pi_n(x_{k-1}) \end{bmatrix} \\ M_H(C; x_0, \ldots, x_{k-1})^{-1} \operatorname{diag}(\{B_i^{1/2}\}_{i=0}^{k-1}) U.$$

As in the scalar case, the optimal solution of the original problem is obtained via the quadratic optimization problem

$$\begin{array}{ll} \min & \langle \operatorname{diag}(\{B_i^{1/2}\}_{i=0}^{k-1}) M_H(C; x_0, \ldots, x_{k-1})^{-1} \operatorname{diag}(\{B_i^{1/2}\}_{i=0}^{k-1}) U, U \rangle \\ \text{s. t.} & U_i^* U_i = I_m, \quad i = 0, \ldots, k-1. \end{array} \tag{45}$$

The associated semidefinite relaxation is

$$\begin{array}{lll} \min & \langle M_H(C; x_0, \ldots, x_{k-1})^{-1}, X \rangle & \\ \text{s. t.} & X_{ii} = B_i, & i = 0, \ldots, k-1 \\ & X \in \mathcal{H}_+^{mk} & \end{array} \tag{46}$$

where X_{ii} is the ith $m \times m$ diagonal block of X. Its dual is given by

$$\begin{array}{ll} \max & \langle B, Y \rangle \\ \text{s. t.} & M_H(C; x_0, \ldots, x_{k-1})^{-1} \succeq \operatorname{diag}(\{Y_i\}_{i=0}^{k-1}) \end{array} \tag{47}$$

and is equal to the dual of the original problem. Therefore, we could proceed as before to obtain the following theorem :

Theorem 5.2. *If Assumption 5.1 holds, relaxation* (46) *of quadratically constrained quadratic problem* (45) *is exact.*

Provided that the original problem has been pre-processed, solving the dual problem (47) does *not* depend on the degree $2n$ of $P(x)$. This result is similar to the scalar case. As Assumption 5.1 guarantees that strict feasibility holds, we obtain an efficient algorithm to solve our problem class.

6 Interpolation conditions on the derivatives

In this section, we present the straightforward extension of our previous results to interpolation conditions on the derivatives. We only consider the scalar case to keep our equations as small as possible.

6.1 Real line

In the real line setting, interpolation constraints on the derivatives are formulated as

$$p^{(\ell)}(x_i) = \langle p, \pi_{2n}^{(\ell)}(x_i)\rangle = b_i$$

where $\pi_{2n}^{(\ell)}(\cdot)$ is the component-wise ℓth derivative of $\pi_{2n}(\cdot)$. Such constraints will be called "interpolation-like" constraints.

If all the linear constraints of (11) are interpolation-like constraints, i.e.

$$\langle a_i, p\rangle \doteq \langle \pi_{2n}^{(\ell_i)}(x_i), p\rangle = b_i, \quad i = 0, \dots, k-1,$$

the dual problem (12) reads now as follows

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & H(c) - \sum_{i=0}^{k-1} y_i H(\pi_{2n}^{(\ell_i)}(x_i)) \succeq 0. \end{array} \tag{48}$$

Let us now prove that $H(\pi_{2n}^{(\ell_i)}(x_i))$ has a special structure.

Proposition 6.1. *Let $\ell \geq 0$. Then*

$$H(\pi_{2n}^{(\ell)}(x)) = \sum_{r=0}^{\ell} C_\ell^r \pi_n^{(r)}(x)(\pi_n^{(\ell-r)}(x))^T, \quad \forall x \in \mathbb{R} \tag{49}$$

and the rank of this matrix is $\min\{\ell, 2n-\ell\} + 1$.

Proof. Since $H(\pi_{2n}(x)) = \pi_n(x)\pi_n(x)^T$ and $H(\cdot)$ is a linear operator, equation (49) is a direct consequence of the chain rule. The rank condition originates from the fact that $\pi_n^{(n+1)}(x) = 0$. □

This proposition allows us to improve the formulation (48) of the dual problem. First of all, assume that the interpolation points are distinct and that $\ell_i \leq n, \forall i$. Let us define a block diagonal matrix

$$\Delta(y) = \operatorname{diag}(\{\Delta_0(y), \dots, \Delta_{k-1}(y)\})$$

where $\Delta_i(y)$ is a $(\ell_i+1)\times(\ell_i+1)$ matrix defined by

$$\Delta_i(y) = \begin{bmatrix} 0 & & C_{\ell_i}^{\ell_i} y_i \\ & \iddots & \\ C_{\ell_i}^{0} y_i & & 0 \end{bmatrix}, \quad i = 0,\ldots,k-1.$$

Using the above proposition, the dual problem can be written as

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s. t.} & H(c) - V\Delta(y)V^T \succeq 0, \end{array} \tag{50}$$

where V is the non-square *confluent Vandermonde matrix*

$$V = \left[\pi_n^{(0)}(x_0) \quad \ldots \quad \pi_n^{(\ell_1)}(x_0) \quad | \quad \ldots \quad | \quad \pi_n^{(0)}(x_m) \quad \ldots \quad \pi_n^{(\ell_m)}(x_m)\right].$$

If the interpolation points are not distinct or if there exists at least one index i such that $\ell_i > n$, the matrix V and the block-diagonal matrix $\Delta(y)$ must be redefined in order to get a dual problem similar to (50). Because the appropriate reformulation is evident, but cumbersome, it has been omitted.

If $H(c) \succ 0$ and the numbers of rows of V is greater than its number of columns, the dual constraint (50) is easily recast using Proposition 4.3 :

$$(V^T H(c)^{-1} V)^{-1} \succeq \Delta(y).$$

The complexity of solving the dual problem (50) depends mostly on the dimension of $\Delta(y)$. That is, an appropriate preprocessing tends to eliminate the dependence on the degree $2n$. Because primal strict feasibility cannot be guaranteed from the knowledge of k, we cannot guarantee that the semidefinite relaxation is exact.

6.2 Unit circle

In the unit circle setting, interpolation constraints on the derivatives, $p^{(\ell_i)}(\theta_i) = b_i$, are equivalent to the linear constraints

$$p^{(\ell_i)}(z_i) = \langle (-\jmath N)^{\ell_i} p, \pi_n(z_i)\rangle = \langle p, (\jmath N)^{\ell_i} \pi_n(z_i)\rangle = b_i, \quad z_i = e^{\jmath\theta_i} \tag{51}$$

where $N = \operatorname{diag}(0, 1, \ldots, n)$.

If all linear constraints of (15) are interpolation-like constraints, i.e.

$$\langle a_i, p\rangle \doteq \langle p, (\jmath N)^{\ell_i} \pi_n(z_i)\rangle = b_i, \quad z_i = e^{\jmath\theta_i}, \quad i = 0,\ldots,k-1,$$

the dual problem (16) reads now as follows

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & T(c) - \sum_{i=0}^{k-1} y_i T((\jmath N)^{\ell_i}\pi_n(z_i)) \succeq 0 \end{array} . \tag{52}$$

Note that $T((\jmath N)^m \pi_n(z))$ has a special structure.

Proposition 6.2. *Let $\ell \geq 0$. Then*

$$T((\jmath N)^\ell \pi_n(z)) = \sum_{r=0}^{\ell} C_\ell^r (\jmath N)^r \pi_n(z)[(\jmath N)^{\ell-r}\pi_n(z)]^* \tag{53}$$

and the rank of this matrix is $\min\{\ell, n\} + 1$.

Proof. Since $T(\pi_n(z)) = \pi_n(z)\pi_n(z)^*$, $\frac{\partial}{\partial\theta}(\pi_n(z)|_{z=e^{\jmath\theta}}) = \jmath N(\pi_n(z)|_{z=e^{\jmath\theta}})$ and $T(\cdot)$ is a linear operator, it is straightforward to check equation (53). □

Assume that the interpolation points are distinct and define the block diagonal matrix

$$\Delta(y) = \text{diag}(\{\Delta_0(y), \ldots, \Delta_{k-1}(y)\})$$

as before. Using the above proposition, the dual problem can be written as

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & T(c) - W\Delta(y)W^* \succeq 0. \end{array}$$

where W is the non-square matrix

$$W = \Big[(\jmath N)^0 \pi_n^{(0)}(z_0), \ldots, (\jmath N)^{\ell_1}\pi_n^{(\ell_1)}(z_0)| \ldots | \\ (\jmath N)^0 \pi_n^{(0)}(z_{k-1}), \ldots, (\jmath N)^{\ell_{k-1}}\pi_n^{(\ell_{k-1})}(x_{k-1})\Big].$$

If $\ell_i \leq 1, \forall i$, the matrix W is the product of a confluent Vandermonde matrix V and a diagonal scaling D, i.e. $W = VD$. If $T(c) \succ 0$ and the numbers of rows of V is greater than its number of columns, the complexity of solving the reformulated dual problem

$$\begin{array}{ll} \max & \langle b, y\rangle \\ \text{s.t.} & (W^*T(c)^{-1}W)^{-1} \succeq \Delta(y), \end{array}$$

depends mostly on the dimension of $\Delta(y)$. That is, an appropriate preprocessing tends to eliminate the dependence on the degree n. However, primal strict feasibility cannot be guaranteed from the knowledge of k so that the exact semidefinite relaxation cannot be certified in general.

7 Conclusion

Conic optimization problems on several cones of non-negative polynomials, with linear constraints generated by interpolation-like constraints, are studied in this article. They naturally induce semidefinite programming problems

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A_i, X\rangle = b_i \quad i = 1, \dots, m \\ & X \succeq 0 \end{array} \tag{54}$$

with low-rank matrices $\{A_i\}_{i=1}^m$. Conditions which guarantee that strict feasibility holds are investigated, see Assumption 4.1 and the associated discussion. Using Proposition 4.3, the associated dual problems can be reformulated efficiently; the complexity of solving the reformulated duals is almost independent of the primal space dimension. Finally, new classes of quadratically constrained quadratic programs with exact semidefinite relaxation are described.

A Appendix

The proof presented in Appendix is based on ideas of A. Nemirovskii.

Proposition A.1. *Let $A = A^*$ be a Hermitian matrix of order $2n+1$. Then the quadratic optimization problem*

$$\begin{array}{ll} \min & \langle Az, z\rangle \\ \text{s.t.} & |z_i| = 1, \quad i = 0, \dots, 2n \end{array} \tag{55}$$

is NP*-hard.*

Proof. Let $\{a_i\}_{i=0}^n \subseteq \mathbb{Z}$ be a finite set of integers. Checking whether there exist $\{x_i\}_{i=0}^n \subseteq \{-1, +1\}$ such that the equality

$$\sum_{i=0}^{2n} a_i x_i = 0 \tag{56}$$

holds is related to the subset sum problem [4, SP13] and is thus NP-complete.

Let $\{z_\ell\}_{\ell=0}^{2n} \subseteq \mathbb{C}$ be a finite set of complex numbers of modulus one and define the quadratic functions

$$P_\ell(z) = |z_0 - z_{2\ell-1}|^2 + |z_{2\ell-1} - z_{2\ell}|^2 + |z_0 - z_{2\ell}|^2, \quad \ell = 1, \dots, n.$$

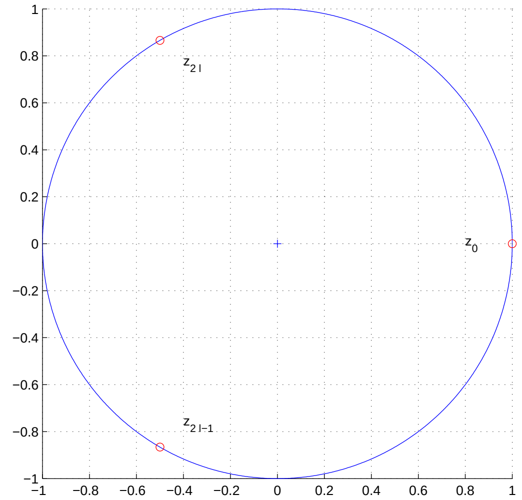

Figure 4: Solution of $\max\{P_\ell(z) : |z_i| = 1, \forall i\}$

Assume that z_0 is equal to 1 without loss of generality. Then the optimization problem

$$\max\{\sum_{\ell=1}^{n} P_\ell(z) : |z_i| = 1, \forall i\}$$

can be solved explicitly, see Figure 4. Note that the inequality

$$\max\{\sum_{\ell=1}^{n} P_\ell(z) - |\sum_{\ell=0}^{n} a_\ell(z_{2\ell+1} - z_{2\ell+2})|^2 : |z_i| = 1, \forall i\} \leq \max\{\sum_{\ell=1}^{n} P_\ell(z) : |z_i| = 1, \forall i\}$$

is tight if and only if Problem (56) is solvable. Since its left hand side is an instance of (55), this quadratic problem is hard to solve. □

References

[1] B. Alkire and L. Vandenberghe, *Convex optimization problems involving finite autocorrelation sequences*, Math. Programming, 93 (2002), pp. 331–359.

[2] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, *Linear matrix inequality formulation of spectral mask constraints with applications to*

FIR filter design, IEEE Trans. Signal Process., 50 (2002), pp. 2702–2715.

[3] S. Feldmann and G. Heinig, *Vandermonde factorization and canonical representations of block Hankel matrices*, Linear Algebra Appl., 241–243 (1996), pp. 247–278.

[4] M. R. Garey and D. S. Johnson, *Computers and intractability: a guide to the theory of NP-completeness*, W. H. Freeman, San Francisco, Calif., 1979.

[5] Y. Genin, Y. Hachez, Y. Nesterov, and P. Van Dooren, *Convex optimization over positive polynomials and filter design*, in Proceedings UKACC International Conference on Control 2000, 2000. CD-ROM Paper SS-41, University of Cambridge, UK, 4-7 September 2000.

[6] ——, *Optimization problems over positive pseudo-polynomial matrices*, SIAM J. Matrix Anal. Appl., (accepted for publication).

[7] Y. Hachez and Y. Nesterov, *Optimization problems over nonnegative polynomials with interpolation constraints*, in Proceedings 15th IFAC World Congress, July 2002. CD-ROM Paper 1650, Barcelona, ES, 21-26 July 2002.

[8] T. Kailath and A. H. Sayed, eds., *Fast reliable algorithms for matrices with structure*, SIAM, Philadelphia, PA, 1999.

[9] S. Karlin and W. J. Studden, *Tchebycheff systems: With applications in analysis and statistics*, vol. 15 of Pure and Applied Mathematics, Interscience Publishers John Wiley & Sons, New York-London-Sydney, 1966.

[10] Y. Nesterov, *Long-step strategies in interior-point primal-dual methods*, Math. Programming, 76 (1997), pp. 47–94.

[11] ——, *Squared functional systems and optimization problems*, in High performance optimization, H.Frenk et al., eds., vol. 33 of Appl. Optim., Kluwer Acad. Publ., Dordrecht, 2000, ch. 17, pp. 405–440.

[12] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, vol. 13 of Studies in Applied Mathematics, SIAM, Philadelphia, PA, 1994.

[13] A. V. Oppenheim and R. W. Schafer, *Discrete-time signal processing*, Prentice Hall Signal Processing Series, Prentice Hall, Englewood Cliffs, NJ, 1989.

[14] G. Pólya and G. Szegő, *Problems and theorems in analysis. Vol. II: Theory of functions, zeros, polynomials, determinants, number theory, geometry*, vol. 216 of Die Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, New York, 1976.

[15] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.