A Convex Optimization Approach to Generalized Moment Problems

Christopher I. Byrnes and Anders Lindquist

ABSTRACT In this paper we present a universal solution to the generalized moment problem, with a nonclassical complexity constraint. We show that this solution can be obtained by minimizing a strictly convex nonlinear functional. This optimization problem is derived in two different ways. We first derive this intrinsically, in a geometric way, by path integration of a one-form which defines the generalized moment problem. It is observed that this one-form is closed and defined on a convex set, and thus exact with, perhaps surprisingly, a strictly convex primitive function. We also derive this convex functional as the dual problem of a problem to maximize a cross entropy functional. In particular, these approaches give a constructive parameterization of all solutions to the Nevanlinna-Pick interpolation problem, with possible higher-order interpolation at certain points in the complex plane, with a degree constraint as well as all soutions to the rational covariance extension problem - two areas which have been advanced by the work of Hidenori Kimura. Illustrations of these results in system identifiaction and probablity are also mentioned.

Key words. Moment problems, convex optimization, Nevanlinna-Pick interpolation, covariance extension, systems identification, Kullback-Leibler distance.

1 Introduction

Hidenori Kimura has been at the forefront of many major developments in the theory of systems and control. Among the many things for which he is widely recognized, two of his contributions have influenced the research results we present in this paper,[1] which we dedicate to him on the occasion of his 60th birthday. The first is his pioneering work on the rational covariance extension problem, and the second is his profound work on robust stabilization and control.

In the rational covariance extension problem, one seeks to parameterize and compute all shaping filters, of a bounded degree, which shape white

[1]This research was supported in part by grants from AFOSR, the Swedish Research Council, the Göran Gustafsson Foundation, and Southwestern Bell.

noise into a process with a given window of covariance lags. In [31], Kimura was able to give a neat parameterization of rational filters of a bounded degree and with a given window of Laurent coefficients, in terms of Szegö polynomials of the first and second kind. This parameterization was independently discovered by Georgiou [23]. From the Kimura-Georgiou parameterization one can see that the space of shaping filters with a fixed window of covariance lags is a smooth manifold [3]. Georgiou [23] used degree theory for functions on manifolds to give a very basic existence result and conjectured a refinement of his result that would give a complete parameterization of all solutions, of no more than a given degree, to the rational covariance extension problem. In [6] Georgiou's conjecture was proved by first noting that the Kimura-Georgiou parameterization defined the leaves of a foliation of a space of positive real rational functions having a bounded degree. A second observation used in [6] was that the fast filtering algorithm of [40, 41], viewed as a dynamical system on the same space of positive real functions, also defines a foliation of this space, with its leaves being the stable manifolds through various equilibria [4, 5]. The proof of Georgiou's conjecture was obtained as a corollary of a theorem about the geometry of these two foliations, including the fact that leaves of one intersect leaves of the other transversely.

The paper [32] on robust stabilization of plants with a fixed number of unstable poles and Nyquist plot in a neighborhood of the Nyquist plot of a nominal plant was one of the key contributions which ushered in the era of H^∞ control. In this paper, Kimura gave necessary and sufficient conditions for the solution of this problem in terms of the classical Nevanlinna-Pick interpolation problem, a methodology he would continue to use and develop for robust control [33, 34, 35, 36, 37]. We also refer the reader to the books [21, 53] and the references therein. Nevanlinna-Pick interpolation for bounded-real rational functions is now one of the tools commonly used in robust control. Indeed, it is widely known that the sensitivity function of a controlled plant must take on certain prescribed values at the unstable poles and zeroes of the plant to be internally stabilized and that the bounded-real condition gives a bound on the H^∞ norm of the sensitivity function. Again it is desirable that the interpolant, being the sensitivity function, has an a priori bounded degree. This leads naturally to the Nevanlinna-Pick problem with degree constraint. A complete parameterization of the class of such interpolants was conjectured by Georgiou in [24] and recently settled in [25] for interpolation at distinct points in the complex plane. This can again be enhanced using the geometry of foliations on a space of rational, positive real functions, with the foliation defined by covariance windows being replaced by a foliation whose leaves are defined by fixing the interpolation values [8].

The question of actually finding, or computing solutions to either problem can be solved in the context of nonlinear convex optimization. In [7], we presented a convex optimization approach for determining an arbitrary solution to the rational covariance extension problem with degree constraint for interpolation at distinct points in the complex plane. In this way, one obtains both an algorithm for solving the covariance extension problem and a constructive proof of Georgiou's conjecture. Similarly, in [9] a generalized entropy criterion is developed for solving the rational Nevanlinna-Pick problem with degree constraints. In both problems, the primal problem of maximizing this entropy gain has a very well-behaved dual problem in a finite-dimensional space and gives algorithms for solving both problems with the degree constraints.

At this point, one should ask whether there is a unified point of view from which one can also see why there must be a strictly convex functional whose minimization solves a given problem. While it is true that the general rational Nevanlinna-Pick interpolation problem, allowing for higher-order interpolation at points in the extended complex plane, does include the rational covariance extension problem, there is a more compelling generalization of these problems. It is classical mathematical fact that both problems can be formulated as special cases of the generalized moment problem, dating back to Chebychev and Markov and formulated in terms of convex functional analysis by Caratheodory, Toeplitz, and others. We refer to the classic book [38] for a neat exposition of these topics. In fact, we shall adopt their notation and refer to positive real functions as Carathéodory functions.

In this paper we present a universal solution to the generalized moment problem, with a nonclassical complexity constraint, obtained by minimizing a strictly convex nonlinear functional. This optimization problem is derived in two different ways. We first answer the question of why, intrinsically, there should always be an equivalent convex optimization problem. We settle this question in a geometric way by path integration of a one-form which defines the generalized moment problem. This exposition follows the original calculation in [16] where it is observed that this one-form is closed and defined on a convex set, and thus exact. Since its integral is therefore path-independent, it is intrinsic and is, perhaps surprisingly, a strictly convex functional. In Section 5 we give a new derivation of this convex functional as the dual problem of a problem to maximize a cross entropy functional. In particular, these approaches give a constructive parameterization of all solutions to the Nevanlinna-Pick interpolation problem, with possible higher-order interpolation at certain points in the complex plane, with a degree constraint.

2 The rational covariance extension problem as a trigonometric moment problem with a complexity constraint

A function of a complex variable that is analytic in the open unit disc and maps points there into the open right half-plane will called a *Carathéodory function* in this paper. The *covariance extension problem* is to find an infinite extension $c_{n+1}, c_{n+2}, \dots$ of a given positive sequence $n+1$ real numbers $c_0, c_1, \cdots, c_n$ such that

$$f(z) = \frac{1}{2}c_0 + c_1 z + c_2 z^2 + c_3 z^3 + \dots$$

is a rational Carathéodory function of degree at most n. Here, by positivity of a sequence we mean positive definiteness of the Toeplitz matrix

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & \cdots & c_0 \end{bmatrix}.$$

A function f that is analytic in the disc is a Carathéodory function if and only if

$$\Phi(\theta) := \mathrm{Re}\{f(e^{i\theta})\}$$

is positive for all $\theta \in [-\pi, \pi]$. Note that

$$\Phi(\theta) = c_0 + 2\sum_{k=0}^{\infty} c_k \cos k\theta.$$

In particular,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(\theta) \cos k\theta \, d\theta = c_k \tag{2.1}$$

for $k = 0, 1, 2, \dots$.

As pointed out in [20], the rational covariance extension problem is a trigonometric moment problem: Given a positive sequence $c_0, c_1, \cdots, c_n$, find all positive Φ such that (2.1) holds for $k = 0, 1, \dots, n$. However, the rational covariance extension problem involves additional constraints. Indeed, for f to be rational of degree at most n it is necessary and sufficient that

$$\Phi(\theta) = \frac{P(\theta)}{Q(\theta)}, \tag{2.2}$$

where P and Q are positive trigonometric polynomials in the span of $\{\cos k\theta\}_{k=0}^{n}$. This is the complexity constraint.

3 The generalized moment problem with complexity constraint

There is a vast literature on the generalized moment problem (see, e.g., [1, 2, 27, 38, 49]), in part because so many problems and theorems in pure and applied mathematics, physics and engineering can be formulated as moment problems. The classical problem can be formulated in the following way.

Consider a sequence of continuous functions

$$\alpha_0, \alpha_1, \cdots, \alpha_n,$$

defined on some interval $\mathcal{I}$ of the real line. Suppose that $\alpha_0, \ldots, \alpha_{r-1}$ are real functions and $\alpha_r, \ldots, \alpha_n$ are complex-valued functions whose real and imaginary parts, taken together with $\alpha_0, \ldots, \alpha_{r-1}$ are linearly independent over $\mathbb{R}$. This is no restriction since a purely imaginary moment condition can always be reduced to a real one. We also assume that $\alpha_0 = 1$. Let $\mathfrak{P}$ be the real vector space that is the sum of the real span of $\alpha_0, \ldots, \alpha_{r-1}$ and the complex span of $\alpha_r, \ldots, \alpha_n$. Hence, the real dimension of $\mathfrak{P}$ is $2n-r+2$. Moreover, let $\mathfrak{P}_+$ be the subset of all functions in $\mathfrak{P}$ that have a positive real part on $\mathcal{I}$, and let $\overline{\mathfrak{P}}_+$ be its closure. Classically, one can replace the assumption that $\alpha_0 = 1$ by simply assuming that $\mathfrak{P}_+$ is nonempty, in which case $\mathfrak{P}_+$ is an open convex subset of $\mathfrak{P}$.

The moment problem is then to find all monotone, nondecreasing functions μ of bounded variation such that

$$\int_{\mathcal{I}} \alpha_k(t) d\mu(t) = c_k, \quad k = 0, 1, \cdots, n, \tag{3.3}$$

where $c_0, c_1, \cdots, c_n$ is a given sequence of complex numbers. This problem has a solution if and only if the sequence $c_0, c_1, \cdots, c_n$ is *positive* in the sense that

$$\langle c, q \rangle := \operatorname{Re} \sum_{k=0}^{n} q_k c_k > 0 \tag{3.4}$$

for all $(q_0, q_1, \cdots, q_n) \in \mathbb{R}^r \times \mathbb{C}^{n-r+1}$ such that

$$q := \sum_{k=0}^{n} q_k \alpha_k \in \overline{\mathfrak{P}}_+ \setminus \{0\}. \tag{3.5}$$

Clearly, this is a necessary condition since

$$\langle c, q \rangle = \int_{\mathcal{I}} \left[\operatorname{Re} \sum_{k=0}^{n} q_k c_k \right] d\mu > 0,$$

whenever (3.5) holds, but it can be shown [38] that it is also sufficient. Denote by $\mathfrak{C}_+$ the space of positive sequences. Since $\mathfrak{P}_+$ is nonempty, so is $\mathfrak{C}_+$.

Next, we introduce the nonclassical *complexity constraint*

$$\frac{d\mu}{dt} = \frac{\Psi(t)}{Q(t)}, \quad Q = \mathrm{Re}\{q\}, q \in \mathfrak{P}_+, \tag{3.6}$$

where Ψ is a given positive function in $L^1(\mathcal{I})$ such that

$$\left| \int_{\mathcal{I}} \alpha_k \frac{\Psi}{Q}\, dt \right| < \infty, \quad k = 0, 1, \dots, n, \tag{3.7}$$

for all $Q = \mathrm{Re}\{q\}, q \in \mathfrak{P}_+$. Let us call the class of such functions $\mathcal{G}_+$. If $\mathcal{I}$ is a finite interval, (3.7) holds for all Ψ in the space $L^1_+(\mathcal{I})$ of positive functions in $L^1(\mathcal{I})$ so then $\mathcal{G}_+ = L^1_+(\mathcal{I})$.

To motivate this problem formulation, let us give a couple of examples, the first of which we have already encountered, for real data, in Section 2.

Example 3.1. A function of a complex variable that is analytic in the open unit disc and maps points there into the open left half-plane is called a *Carathéodory function.* Given $n+1$ complex numbers $c_0, c_1, \cdots, c_n$, find an infinite extension $c_{n+1}, c_{n+2}, \dots$ such that

$$f(z) = \frac{1}{2}c_0 + c_1 z + c_2 z^2 + c_3 z^3 + \dots$$

is a rational Carathéodory function of degree at most n. This *covariance extension problem* is a trigonometric moment problem with degree constraint and can be formulated as above by taking $\Phi = 2\mathrm{Re}\{f\}$, $\mathcal{I} = [-\pi, \pi]$ and

$$\alpha_k(\theta) = e^{ik\theta}, \quad k = 0, 1, \dots, n.$$

Moreover, the degree constraint is equivalent to (3.6) if we take

$$\Psi = \mathrm{Re}\{\psi\}, \quad \psi \in \mathfrak{P}_+. \tag{3.8}$$

Finally, $c_0, c_1, \dots, c_n$ is positive if and only if the Toeplitz matrix

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ \bar{c}_1 & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{c}_n & \bar{c}_{n-1} & \cdots & c_0 \end{bmatrix}$$

is positive definite.

Example 3.2. Given $n+1$ distinct points $z_0, z_1, \dots, z_n$ in the open unit disc, consider the problem to determine the rational Carathéodory functions f of degree at most n satisfying the interpolation condition

$$f(z_k) = c_k, \quad k = 0, 1, \dots, n, \tag{3.9}$$

where $c_0, c_1, \dots, c_n$ are prescribed values in the open right half of the complex plane with c_0 real. If the points $z_0, z_1, \dots, z_n$ are not distinct, the interpolation conditions are modified in the following way. If $z_k = z_{k+1} = \cdots = z_{k+m-1}$, the corresponding interpolation conditions are replaced by

$$f(z_k) = c_k, \quad f'(z_k) = c_{k+1}, \quad \dots, \quad \frac{1}{(m-1)!} f^{(m-1)}(z_k) = c_{k+m-1} \tag{3.10}$$

This Nevanlinna-Pick interpolation problem differs from the classical one in that a degree constraint on the interpolant f has been introduced, a restriction motivated by applications [9, 10]. In fact, many problems in systems and control can be reduced to Nevanlinna-Pick interpolation [19, 21, 28, 32, 53, 50], and, as the interpolant generally can be interpreted as a transfer function, the bound on the degree is a natural complexity constraint.

To reformulate this interpolation problem as a generalized moment problem of the type described above, we note that, by the Herglotz Theorem, any rational Carathéodory function of degree at most n such that $f(0)$ is real can be represented as

$$f(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \Phi(e^{i\theta}) d\theta \tag{3.11}$$

in terms of a rational spectral density Φ of degree at most $2n$, which is given by

$$\Phi(e^{i\theta}) = \mathrm{Re}\{f(e^{i\theta})\}. \tag{3.12}$$

Differentiating (3.11), we obtain

$$\frac{1}{\nu!} f^{(\nu)}(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{2e^{i\theta}}{(e^{i\theta} - z)^{\nu+1}} \Phi(e^{i\theta}) d\theta, \quad \nu = 1, 2, \dots.$$

Consequently, the interpolation conditions can be written

$$\int_{-\pi}^{\pi} \alpha_k(\theta) \Phi(e^{i\theta}) d\theta = c_k, \quad k = 0, 1, \dots, n, \tag{3.13}$$

where

$$\alpha_k(\theta) = \frac{1}{2\pi} \frac{e^{i\theta} + z_k}{e^{i\theta} - z_k}$$

for single interpolation points and

$$\begin{aligned}
\alpha_k(\theta) &= \frac{1}{2\pi} \frac{e^{i\theta} + z_k}{e^{i\theta} - z_k} \\
\alpha_{k+1}(\theta) &= \frac{1}{2\pi} \frac{2e^{i\theta}}{(e^{i\theta} - z_k)^2} \\
&\vdots \\
\alpha_{k+m-1}(\theta) &= \frac{1}{2\pi} \frac{2e^{i\theta}}{(e^{i\theta} - z_k)^m}
\end{aligned}$$

whenever $z_k = z_{k+1} = \cdots = z_{k+m-1}$. Taking $\mathcal{I} = [-\pi, \pi]$ and $d\mu = \Phi d\theta$, (3.13) are precisely of the moment conditions (3.3). Moreover the degree constraint is equivalent to (3.6) if we take Ψ in the class (3.8).

If the interpolation points $\{z_k\}_0^n$ are distinct, the sequence $c_0, c_1, \dots, c_n$ is positive if and only if the Pick matrix

$$\left[\frac{c_k + \bar{c}_\ell}{1 - z_k \bar{z}_\ell}\right]_{k,\ell=0}^n$$

is positive definite. If $z_k = 0$, $k = 0, 1, \dots, n$, we have the trigonometric moment problem with degree constraint, described in Example 3.1.

Example 3.3. A popular method in systems identification amounts to estimating the first $n + 1$ coefficients in an orthogonal basis function expansion

$$G(z) = \frac{1}{2} c_0 f_0(z) + \sum_{k=1}^{\infty} c_k f_k(z)$$

of a transfer function $G(z)$ [51, 52]. The functions $f_0, f_1, f_2, \dots$ are orthonormal on the unit circle, i.e.,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_j(e^{i\theta})^* f_k(e^{i\theta}) d\theta = \delta_{jk}.$$

A general class of such functions is given by

$$f_k(z) = \frac{\sqrt{1 - |\xi_k|^2}}{z - \xi_k} \prod_{j=0}^{k-1} \frac{1 - \xi_j^* z}{z - \xi_j},$$

where $\xi_0, \xi_1, \xi_2, \dots$ are poles to be selected by the user [29].

Given the estimated coefficients $c_0, c_1, \cdots, c_n$, the usual problem considered in the literature [47] is to find a rational function G of smallest degree which match these coefficients. Here, however, we consider the corresponding problem where G is a Carathéodory function of degree at most n, a problem that has remained open. It can be reformulated as a generalized moment problem with complexity constraint by observing that

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_k(e^{i\theta}) \Phi(e^{i\theta})\, d\theta, \quad k = 0, 1, \dots, n,$$

where $\Phi(e^{i\theta}) = 2\mathrm{Re}\{G(e^{i\theta})\}$. The degree constraint is enforced by choosing Ψ as in (3.8).

There is one and only one solution to the generalized moment problem with complexity constraint formulated above, and it is determined by the unique minimum of a strictly convex functional. In fact, the following theorem is a version of a result proved in [16], generalizing similar results in analytic interpolation theory [7, 9, 10, 11, 15].

Theorem 3.4. *Let $\mathfrak{P}$ be spanned by C^2 functions $\alpha_0, \alpha_1, \dots, \alpha_n$, whose nonzero real and imaginary parts form a linearly independent set over $\mathbb{R}$. Moreover, let $c \in \mathfrak{C}_+$. Then, for any $\Psi \in \mathcal{G}_+$, there is one and only one $q \in \mathfrak{P}_+$ solving the generalized moment problem*

$$\int_{\mathcal{I}} \alpha_k(t) \frac{\Psi(t)}{Q(t)}\, dt = c_k, \quad k = 0, 1, \dots, n, \tag{3.14}$$

where $Q = \operatorname{Re}\{q\}$. In fact, c and Ψ determine Q as the unique minimum in $\mathfrak{P}_+$ of the strictly convex functional

$$\mathbb{J}_\Psi(q) = \langle c, q \rangle - \int_{\mathcal{I}} \Psi \log Q dt. \tag{3.15}$$

Note that $\mathbb{J}_\Psi(q)$ is finite for all $q \in \mathfrak{P}_+$. In fact, by Jensen's inequality,

$$-\log \int_{\mathcal{I}} \frac{\Psi}{Q} dt \leq \int_{\mathcal{I}} \Psi \log Q dt \leq \log \int_{\mathcal{I}} Q\Psi dt,$$

where both bounds are finite by (3.7). (To see this, for the lower bound take $k = 0$; for the upper bound first take $Q = 1$ in (3.7) and the form the appropriate linear combination.)

In Section 4 we shall outline a geometric proof of Theorem 3.4, in which the functional (3.15) is derived by integrating a 1-form along a path. It is the proof presented in [16]. Another proof is based on duality theory and follows the same pattern as in [7, 9, 10, 11]. In fact, the problem to maximize $\mathbb{J}_\Psi$ over $\mathfrak{P}_+$ is the dual problem in the sense of mathematical programming of the constrained optimization problem in the following theorem.

Theorem 3.5. *Let $\mathfrak{P}$ be spanned by C^2 functions $\alpha_0, \alpha_1, \dots, \alpha_n$, whose nonzero real and imaginary parts form a linearly independent set over $\mathbb{R}$. For any choice of $\Psi \in \mathcal{G}_+$, the constrained optimization problem to minimize the functional*

$$\mathbb{I}_\Psi(\Phi) = \int_{\mathcal{I}} \Psi(t) \log \frac{\Psi(t)}{\Phi(t)} dt \tag{3.16}$$

over $L^1_+(\mathcal{I})$ subject to the constraints

$$\int_{\mathcal{I}} \alpha_k(t)\Phi(t) dt = c_k, \quad k = 0, 1, \dots, n, \tag{3.17}$$

has unique solution, and it has the form

$$\Phi = \frac{\Psi}{Q}, \quad Q = \operatorname{Re}\{q\},$$

where $q \in \mathfrak{P}_+$ is the unique minimum of (3.15).

The proof of this theorem can be found in Section 5.

Example 3.6. Consider the problem of determining a probability density Φ on $\mathcal{I} := (-\infty, \infty)$ of the form

$$\Phi(t) = \frac{\Psi(t)}{Q(t)}, \quad Q(t) = q_n t^{2n} + q_{n-1} t^{2n-1} + \cdots + q_0 \tag{3.18}$$

and satisfying the moment conditions

$$\int_{-\infty}^{\infty} t^k \Phi(t) dt = c_k, \quad k = 0, 1, \ldots, 2n, \tag{3.19}$$

where, of course, $c_0 = 1$. Here Ψ could be an a priori selected probability distribution. This is a moment problem of the type described above with

$$\alpha_k(t) = t^k, \quad k = 0, 1, \ldots, 2n.$$

This is a Hamburger moment problem with complexity constaint. The sequence $c_0, c_1, \ldots, c_{2n}$ is positive if and only if the Hankel matrix

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n} \end{bmatrix}$$

is positive definite. Remarkably, the function

$$\mathbb{I}(\Phi, \Psi) = \mathbb{I}_\Psi(\Phi)$$

is the cross-entropy [26], gain of information [46], directed divergency [39], or the Kullback-Leibler distance between Ψ and Φ. Then the optimization problem of Theorem 3.5 is equivalent to minimizing $\mathbb{I}(\Phi, \Psi)$ subject to the moment conditions (3.3). This gives an interesting interpretation, further pursued in [12], to the present problem: Given an a priori probability density Ψ, we want to find another probability density Φ that has prescribed moments up to order $2n$ and that minimizes the Kullback-Leibler distance to Ψ. Similar optimization problems have been considered in the statistical literature, where, however, minimization is generally with respect to Ψ; see, e.g., [17, 18].

4 A geometric derivation of the convex optimization problem

Following [16], we prove Theorem 3.4 by first constructing the dual functional from the moment equations (3.14). To this end, for any $\Psi \in \mathcal{G}_+$, we define the map

$$F^\Psi : \mathfrak{P}_+ \to \mathfrak{C}_+$$

componentwise via

$$F_k^\Psi(q) = \int_{\mathcal{I}} \alpha_k \frac{\Psi}{Q}\, dt.$$

We want to prove that the map F^Ψ is a homeomorphism. For this we use the following global inverse function theorem proven in [16].

Lemma 4.1. *Suppose M and N are n-dimensional, topological manifolds and that N is connected. Consider a continuous map $f : M \to N$. Then, f is a homeomorphism if and only if f is injective and proper. In this case, M is connected.*

Recall that f is said to be *proper* if the inverse image $f^{-1}(K)$ is compact for all compact $K \in M$. That F^Ψ is proper follows from the following lemma, also proven in [16].

Lemma 4.2. *Suppose $\mathfrak{P}$ is a vector space consisting of C^2-smooth functions. Then $F^\Psi : \mathfrak{P}_+ \to \mathfrak{C}_+$ is proper.*

The domain $\mathfrak{P}_+$ is an open convex subset of $\mathbb{R}^{2n-r+2}$, and clearly the same holds for $\mathfrak{C}_+$. Therefore, it will follow from Lemma 4.1 that F^Ψ is a homeomorphism if we can show that F^Ψ is injective. We proceed to proving precisely this by constructing the functional (3.15).

Since $\mathfrak{P}_+$ is an open, convex subset of $\mathbb{R}^{2n-r+2}$, it is diffeomorphic to $\mathbb{R}^{2n-r+2}$ and hence Euclidean. Parameterizing $q \in \mathfrak{P}_+$ via $q = \sum_{k=0}^n q_k \alpha_k$, we construct the 1-form

$$\omega = \operatorname{Re}\left\{ \sum_{k=0}^n \left[c_k - F_k(q)\right] dq_k \right\},$$

on $\mathfrak{P}_+$, which can be written as

$$\begin{aligned}
\omega &= \operatorname{Re}\left\{ \sum_{k=0}^n c_k dq_k - \int_{\mathcal{I}} \sum_{k=0}^n \alpha_k \frac{\Psi}{Q} dq_k dt \right\} \\
&= \operatorname{Re} \sum_{k=0}^n c_k dq_k - \int_{\mathcal{I}} \frac{\Psi}{Q}\, dQ\, dt.
\end{aligned}$$

Taking the exterior derivative on $\mathfrak{P}_+$, we obtain

$$d\omega = \int_{\mathcal{I}} \frac{\Psi}{Q^2}\, dQ \wedge dQ\, dt = 0,$$

establishing that the 1-form ω is closed.

Therefore, by the Poincaré Lemma, there exist a smooth function $\mathbb{J}_\Psi$ such that $\omega = d\mathbb{J}_\Psi$, and hence

$$\mathbb{J}_\Psi = \int \omega = \int \left(\operatorname{Re} \sum_{k=0}^n c_k dq_k - \int_{\mathcal{I}} \frac{\Psi}{Q}\, dQ\, dt \right),$$

with the integral being independent of the path between two endpoints. Computing the path integral

$$\int_{q_0}^{q_1} \omega = \left[\langle c, q \rangle - \int_{\mathcal{I}} \Psi \log Q \; dt \right]_{q_0}^{q_1},$$

we obtain, modulo a constant of integration,

$$\mathbb{J}_\Psi(q) = \langle c, q \rangle - \int_{\mathcal{I}} \Psi \log Q \; dt. \tag{4.20}$$

Since

$$\frac{\partial \mathbb{J}_\Psi}{\partial q_k} = c_k - \int_{\mathcal{I}} \alpha_k \frac{\Psi}{Q} \, dt, \quad k = 0, 1, \ldots, n, \tag{4.21}$$

the functional (4.20) has a critical point precisely at the solution of the generalized moment problem. To see this on the second factor of $\mathbb{R}^r \times \mathbb{C}^{n-r+1}$, we decompose the exterior differential as the sum $d = \partial + \bar{\partial}$, where $\bar{\partial}$ is the Cauchy-Riemann differential. Since $\mathbb{J}_\Psi$ is real, to say that $d\mathbb{J}_\Psi = 0$ is to say that $\partial \mathbb{J}_\Psi = 0$ or, equivalently, that $\bar{\partial}\mathbb{J}_\Psi = 0$. Finally, by inspection we see that $\partial \mathbb{J}_\Psi = 0$ is the set of defining equations of the generalized moment problem.

Now, it is easy to see that $\mathbb{J}_\Psi$ is strictly convex. Hence there is at most one critical point, which must be a minimum. Consequently, F^Ψ is injective, as claimed.

5 Duality theory

Next we turn to the constraint optimization problem of Theorem 3.5, reformulated as a maximization problem: Maximize the cross entropy

$$\mathbb{I}_\Psi(\Phi) = \int_{\mathcal{I}} \Psi \log \frac{\Phi}{\Psi} \, dt \tag{5.22}$$

over $L_+^1(\mathcal{I})$ subject to the constraints

$$\int_{\mathcal{I}} \alpha_k(t)\Phi(t)dt = c_k, \quad k = 0, 1, \ldots, n. \tag{5.23}$$

Note that $\mathbb{I}_\Psi(\Phi)$, here redefined to be the negative of that in Theorem 3.5, is bounded from above. To see this, we may without restriction assume that $\int_{\mathcal{I}} \Psi dt = 1$. Then, Jensen's inequality [48, p. 61] yields $\mathbb{I}_\Psi(\Phi) \leq \log c_0$.

To solve this optimization problem, we form the Lagrangian

$$L(\Phi, q) = \mathbb{I}_\Psi(\Phi) + \mathrm{Re} \sum_{k=0}^{n} q_k \left[c_k - \int_{\mathcal{I}} \alpha_k \Phi dt \right],$$

where $(q_0, q_1, \ldots, q_n) \in \mathbb{R}^r \times \mathbb{C}^{n-r+1}$ are Lagrange multipliers, complex for the complex moment conditions and real for the real ones. Setting

$$Q = \mathrm{Re} \sum_{k=0}^{n} q_k \alpha_k,$$

we obtain

$$L(\Phi, q) = \int_{\mathcal{I}} \Psi \log \frac{\Phi}{\Psi}\, dt + \langle c, q\rangle - \int_{\mathcal{I}} Q\Phi dt.$$

Clearly, the dual function

$$\psi(q) = \sup_{\Phi \in L^1_+(\mathcal{I})} L(\Phi, q)$$

takes finite values only if $q \in \mathfrak{P}_+$. For such q, form the directional derivative

$$\delta L(\Phi, q; h) = \lim_{\epsilon \to 0} \frac{L(\Phi + \epsilon h, q) - L(\Phi, q)}{\epsilon}$$

in any direction $h \in L^1(\mathcal{I})$ to obtain

$$\delta L(\Phi, q; h) = \int_{\mathcal{I}} \left[\frac{\Psi}{\Phi} - Q\right] h dt,$$

provided the limit exists. Consequently, for each such Φ,

$$\delta L(\Phi, q; h) = 0, \quad \text{for all } h \in L^1(\mathcal{I})$$

if and only if

$$\Phi = \frac{\Psi}{Q},$$

which inserted into the dual functional yields

$$\psi(q) = \langle c, q\rangle - \int_{\mathcal{I}} \Psi \log Q\, dt - \int_{\mathcal{I}} \Psi\, dt.$$

Since the last term is constant, the dual problem to minimize $\psi(q)$ over $\mathfrak{P}_+$ is equivalent to the optimization problem

$$\min_{q \in \mathfrak{P}_+} \mathbb{J}_\Psi(q)$$

of Theorem 3.4.

In Section 4 we proved that this optimization problem has a unique solution at some $\hat{q} \in \mathfrak{P}_+$ satisfying the moment conditions (5.23). Then, setting $\hat{Q} = \mathrm{Re}\{\hat{q}\}$,

$$\hat{\Phi} := \frac{\Psi}{\hat{Q}} \in L^1_+(\mathcal{I}). \tag{5.24}$$

Since the function $\Phi \to L(\Phi, \hat{q})$ is strictly concave and

$$\delta L(\hat{\Phi}, \hat{q}; h) = \int_{\mathcal{I}} \left[\frac{\Psi}{\hat{\Phi}} - \hat{Q} \right] h \, dt = 0, \quad \text{for all } h \in L^1(\mathcal{I}),$$

we have

$$L(\Phi, \hat{q}) \leq L(\hat{\Phi}, \hat{q}) \quad \text{for all } \Phi \in L^1_+(\mathcal{I}), \tag{5.25}$$

with equality if and only if $\Phi = \hat{\Phi}$.

However, $L(\Phi, \hat{q}) = \mathbb{I}_\Psi(\Phi)$ for all Φ satisfying the moment conditions (5.23). In particular, since (3.14) holds for $Q = \hat{Q}$, $L(\hat{\Phi}, \hat{q}) = \mathbb{I}_\Psi(\hat{\Phi})$. Consequently, (5.25) implies that

$$\mathbb{I}_\Psi(\Phi) \leq \mathbb{I}_\Psi(\hat{\Phi})$$

for all $\Phi \in L^1_+(\mathcal{I})$ satisfying the moment conditions (5.23) with equality if and only if $\Phi = \hat{\Phi}$. Hence, $\mathbb{I}_\Psi$ has a unique maximum on the space of all $\Phi \in L^1_+(\mathcal{I})$ satisfying the constraints (5.23), and it is given by (5.24). This proves Theorem 3.5.

6 Conclusions

It is well known that many important problems in control theory and statistical signal processing can be formulated as moment problems. However, modern engineering applications require complexity constraints, thus significantly altering the mathematical problem so that classical theory does not apply. These applications also require complete classes of solutions, rather than just one single solution, and complete smooth parameterizations of these classes.

In this paper we presented a universal solution to the generalized moment problem, with a nonclassical complexity constraint, obtained by minimizing a strictly convex nonlinear functional. This optimization problem was derived in two different ways, first geometrically by path integration of a one-form, and second via duality theory of mathematical programming.

We have thus provided a unified framework that generalizes previous work on interpolation of the Carathéodory and of the Nevanlinna-Pick type [7, 9, 10, 15]. We refer to [9, 13, 14] for applications to signal processing and to [42, 43, 44, 45] for applications to robust control. Algorithms using homotopy continuation methods, based on our convex optimization approach, have been developed for Carathéodory extension in [22] and for Nevanlinna-Pick interpolation in [43, 45].

The results presented in this paper have interesting interpretations also in probability and statistics, where moment problems are prevalent. Indeed, the complexity constraint may, for example, represent a priori information about a probability distribution to be estimated from moments.

References

[1] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Hafner Publishing, New York, 1965.

[2] N.I. Ahiezer and M. Krein, *Some Questions in the Theory of Moments*, American Mathematical Society, Providence, Rhode Island, 1962.

[3] C. I. Byrnes and A. Lindquist, *On the geometry of the Kimura-Georgiou parameterization of modelling filter*, Inter. J. of Control **50** (1989), 99–105.

[4] C. I. Byrnes, A. Lindquist, and T. McGregor, *Predictability and unpredictability in Kalman filtering*, IEEE Transactions Auto. Control **36** (1991), 563–579.

[5] C. I. Byrnes, A. Lindquist, and Y. Zhou, *On the nonlinear dynamics of fast filtering algorithms*, SIAM J. Control and Optimization, **32** (1994), 744–789.

[6] C. I. Byrnes, A. Lindquist, S.V. Gusev and A. S. Matveev, *A complete parametrization of all positive rational extensions of a covariance sequence*, IEEE Trans. Automatic Control **AC-40** (1995), 1841-1857.

[7] C. I. Byrnes, S.V. Gusev, and A. Lindquist, *A convex optimization approach to the rational covariance extension problem,* SIAM J. Control and Optimization **37** (1999), 211–229.

[8] C. I. Byrnes and A. Lindquist, *On the duality between filtering and Nevanlinna-Pick interpolation*, SIAM J. Control and Optimization **39** (2000), 757–775.

[9] C. I. Byrnes, T. T. Georgiou and A. Lindquist, *A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint*, IEEE Trans. Automatic Control **AC-46** (2001), 822–839.

[10] C.I. Byrnes, T.T. Georgiou, and A. Lindquist, *A new approach to spectral estimation: A tunable high-resolution spectral estimator*, IEEE Trans. on Signal Processing **SP-49** (Nov. 2000), 3189–3205.

[11] C.I. Byrnes, T.T. Georgiou, and A. Lindquist, *Generalized interpolation in H^∞: Solutions of bounded complexity*, preprint.

[12] C.I. Byrnes, T.T. Georgiou, and A. Lindquist, forthcoming paper.

[13] C. I. Byrnes, P. Enqvist, and A. Lindquist, *Cepstral coefficients, covariance lags and pole-zero models for finite data strings*, IEEE Trans. on Signal Processing **SP-50** (2001), 677–693.

[14] C. I. Byrnes, P. Enqvist, and A. Lindquist, *Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach*, SIAM J. Control and Optimization, **41** (2002), 23–59.

[15] C. I. Byrnes, S.V. Gusev, and A. Lindquist, *From finite covariance windows to modeling filters: A convex optimization approach*, SIAM Review **43** (Dec. 2001), 645–675.

[16] C. I. Byrnes and A. Lindquist, *Interior point solutions of variational problems and global inverse function theorems*, submitted for publication.

[17] I. Csiszár, *I-divergence geometry of probability distributions and minimization problems*, Ann. Probab. **3** (1975), 146–158.

[18] I. Csiszár and F. Matúš, *Information projections revisited*, prepint, 2001.

[19] Ph. Delsarte, Y. Genin and Y. Kamp, *On the role of the Nevanlinna-Pick problem in circuits and system theory*, Circuit Theory and Applications **9** (1981), 177–187.

[20] Ph. Delsarte, Y. Genin, Y. Kamp and P. van Dooren, *Speech modelling and the trigonometric moment problem*, Philips J. Res. **37** (1982), 277–292.

[21] J. C. Doyle, B. A. Frances and A. R. Tannenbaum, *Feedback Control Theory*, Macmillan Publ. Co., New York, 1992.

[22] P. Enqvist, *Spectral estimation by Geometric, Topological and Optimization Methods*, PhD thesis, Optimization and Systems Theory, KTH, Stockholm, Sweden, 2001.

[23] T. T. Georgiou, *Realization of power spectra from partial covariance sequences*, IEEE Transactions on Acoustics, Speech, and Signal Processing **ASSP-35** (1987), 438–449.

[24] T. T. Georgiou, *A topological approach to Nevanlinna-Pick interpolation*, SIAM J. Math. Analysis **18** (1987), 1248–1260.

[25] T. T. Georgiou, *The interpolation problem with a degree constraint*, IEEE Transactions on Automatic Control **44(3)**, 631–635, 1999.

[26] I. J. Good, *Maximum entropy for hypothesis formulation, especially for multidimentional contingency tables*, Annals Math. Stat. **34** (1963), 911–934.

[27] U. Grenander and G. Szegö, *Toeplitz forms and their applications*, Univ. California Press, 1958.

[28] J. W. Helton, *Non-Euclidean analysis and electronics*, Bull. Amer. Math. Soc. **7** (1982), 1–64.

[29] P. S. C. Heuberger, T. J. de Hoog, P. M. J. Van den Hof, Z. Szabó and J. Bokar, *Minimal partial realization from orthonormal basis function expansions*, Proc. 40th IEEE Conf. Decision and Control, Orlando, Florida, USA, December 2001, 3673–3678.

[30] R. E. Kalman, *Realization of covariance sequences*, Proc. Toeplitz Memorial Conference (1981), Tel Aviv, Israel, 1981.

[31] H. Kimura, *Positive partial realization of covariance sequences*, Modelling, Identification and Robust Control (C. I. Byrnes and A. Lindquist, eds.), North-Holland, 1987, pp. 499–513.

[32] H. Kimura, *Robust stabilizability for a class of transfer functions*, IEEE Trans. Automat. Control, **AC-29** (1984), pp. 788–793.

[33] H. Kimura, *State space approach to the classical interpolation problem and its applications*, Three decades of mathematical system theory, 243–275, Lecture Notes in Control and Inform. Sci., 135, Springer, Berlin, 1989.

[34] H. Kimura and H. Iwase, *On directional interpolation in H^∞*. Linear circuits, systems and signal processing: theory and application (Phoenix, AZ, 1987), 551–560, North-Holland, Amsterdam, 1988.

[35] H. Kimura, *Conjugation, interpolation and model-matching in H^∞*, Internat. J. Control **49** (1989), 269–307.

[36] H. Kimura, *Directional interpolation in the state space*, Systems Control Lett. **10** (1988), 317–324.

[37] H. Kimura, *Directional interpolation approach to H^∞-optimization and robust stabilization*, IEEE Trans. Automat. Control **AC-32** (1987), 1085–1093.

[38] M.G. Krein and A.A. Nudelman, *The Markov Moment Problem and Extremal Problems*, American Mathematical Society, Providence, Rhode Island, 1977.

[39] S. Kullback, *Information Theory and Statistics*, John Wiley, New York, 1959.

[40] A. Lindquist, *A new algorithm for optimal filtering of discrete-time stationary processes*, SIAM J. Control **12** (1974), 736–746.

[41] A. Lindquist, *Some reduced-order non-Riccati equations for linear least-squares estimation: the stationary, single-output case*, Int. J. Control **24** (1976), 821–842.

[42] R. Nagamune and A. Lindquist, *Sensitivity shaping in feedback control and analytic interpolation theory*, Optimal Control and Partial Differential Equations, J.L. Medaldi et al (editors), IOS Press, Amsterdam, 2001, pp. 404–413.

[43] R. Nagamune, *A Robust Solver Using a Continuation Method for Nevanlinna-Pick Interpolation with Degree Constraint*, submitted to IEEE Transactions on Automatic Control.

[44] R. Nagamune, *Closed-loop Shaping Based on Nevanlinna-Pick Interpolation with a Degree Bound*, submitted to IEEE Transactions on Automatic Control.

[45] R. Nagamune, *Robust Control with Complexity Constraint*, PhD thesis, Optimization and Systems Theory, Royal Institute of Technology, Stockholm, Sweden, 2002.

[46] A. Renyi, *Probability Theory*, 1970.

[47] Z. Szabó, P. Heuberger, J. Bokar and P. An den Hof, *Extended Ho-Kalman algorithm for systems represented in generalized orthonormal bases*, Automatica **36** (2000), 1809–1818.

[48] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1966.

[49] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, Mathematical Surveys I, American Mathematical Society, New York, 1943.

[50] A. Tannenbaum, *Feedback stabilization of linear dynamical plants with uncertainty in the gain factor*, Int. J. Control **32** (1980), 1–16.

[51] B. Wahlberg, *Systems identification using Laguerre models*, IEEE Trans. Automatic Control **AC-36** (1991), 551–562.

[52] B. Wahlberg, *System identification using Kautz models*, IEEE Trans. Automatic Control **AC-39** (1994), 1276–1282.

[53] K. Zhou, *Essentials of Robust Control*, Prentice-Hall, 1998.

Department of Systems Science and Mathematics
Washington University
St. Louis, MO 63130, USA
E-mail: chrisbyrnes@seas.wustl.edu

Department of Mathematics
Division of Optimization and Systems Theory
Royal Institute of Technology
100 44 Stockholm, Sweden
E-mail: alq@math.kth.se