A GENERALIZED ENTROPY CRITERION FOR NEVANLINNA-PICK INTERPOLATION WITH DEGREE CONSTRAINT*

CHRISTOPHER I. BYRNES†, TRYPHON T. GEORGIOU§, AND ANDERS LINDQUIST‡

ABSTRACT. In this paper we present a generalized entropy criterion for solving the rational Nevanlinna-Pick problem for $n + 1$ interpolating conditions and the degree of interpolants bounded by n. The primal problem of maximizing this entropy gain has a very well-behaved dual problem. This dual is a convex optimization problem in a finite-dimensional space and gives rise to an algorithm for finding all interpolants which are positive real and rational of degree at most n. The criterion requires a selection of a monic Schur polynomial of degree n. It follows that this class of monic polynomials completely parameterizes all such rational interpolants, and therefore it provides a set of design parameters for specifying such interpolants. The algorithm is implemented in state space form and applied to several illustrative problems in systems and control, namely sensitivity minimization, maximal power transfer and spectral estimation.

1. Introduction

In this paper we consider the following interpolation problem, which we refer to as the *Nevanlinna-Pick problem with degree constraint.* Given a set of $n + 1$ distinct points

$$\mathcal{Z} := \{z_0, z_1, \ldots, z_n\}$$

in the complement of the unit disc $\mathbb{D}^c := \{z \mid |z| > 1\}$, and a set of $n + 1$ values

$$\mathcal{W} := \{w_0, w_1, \ldots, w_n\}$$

in the open right half of the complex plane, denoted $\mathbb{C}^+$, we seek a parameterization of all functions $f(z)$ which

  (i) satisfy the interpolation conditions

$$f(z_k) = w_k \quad \text{for } k = 0, 1, \ldots, n, \tag{1.1}$$

  (ii) are analytic and have nonnegative real part in $\mathbb{D}^c$, and
 (iii) are rational of (McMillan) degree at most n.

† Department of Systems Science and Mathematics, Washington University, St. Louis, Missouri 63130, USA

§ Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA

‡ Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden

Requiring only condition (i) amounts to standard Lagrange interpolation, the solution of which is elementary. Requiring also condition (ii) yields a classical problem in complex analysis, namely the Nevanlinna-Pick interpolation problem [38]. This problem has a solution if and only if the *Pick matrix*

$$P = \left[\frac{w_k + \bar{w}_\ell}{1 - z_k^{-1} \bar{z}_\ell^{-1}} \right]_{k,\ell=0}^n \tag{1.2}$$

is positive semidefinite [38, 35]. Moreover, the solution is unique if and only if P is singular. Clearly, the case $P > 0$ is what interests us here. If points in $\mathcal{Z}$ are not distinct, the interpolation conditions (i) involve derivatives of $f(z)$, and the Pick matrix is suitably modified [38].

The functions satisfying (ii) are known as *Carathéodory functions* in the mathematical literature. In circuits and systems the same functions are referred to as *positive real*. They play a fundamental role in describing the impedance of RLC circuits, in formalizing questions of stability via energy dissipation in linear and nonlinear systems, and in characterizing the positivity of probability measures in stochastic systems theory. For these reasons, problems involving interpolation by positive real functions play an important role in circuit theory [39, 11, 25], robust stabilization and control [36, 37, 40, 30, 29, 21, 13], signal processing [18, 6, 7, 8, 2], speech synthesis [12], and stochastic systems theory [27, 5, 4].

However, in all these applications, it is important that the interpolating function be rational with a degree which does not exceed some prescribed bound. Degree constraints present some new challenges which need to be incorporated systematically into any useful enhancement of the classical theory. While the Nevanlinna-Schur recursion algorithm and the well-known linear fractional parametrization of all solutions [38] can be used to generate rational solutions, this does not provide any insight into how to parameterize all rational solutions of a given bounded degree. In general, even if the Nevanlinna-Pick problem is solvable, the set of interpolants of degree $k < n$ may be empty, and to determine whether this is the case is often a very hard problem. Hence, at the present time, there is no computationally efficient way to determine *minimum* degree interpolants. However, the set of interpolants of degree at most n is always nonempty, which motivates condition (iii). The surprising fact, to be demonstrated below, is that this set can be parametrized by spectral zeros.

Now, if the rational, positive-real function f is represented as

$$f(z) = \frac{b(z)}{a(z)}, \tag{1.3}$$

where, for the moment, we take $a(z)$ and $b(z)$ to be polynomials of degree n, then

$$\Phi(z) := f(z) + f^*(z) = \frac{\Psi(z)}{a(z)a^*(z)}, \tag{1.4}$$

where $f^*(z) := \overline{f(\bar{z}^{-1})}$ and

$$\Psi(z) := a(z)b^*(z) + a^*(z)b(z). \tag{1.5}$$

(Later, to simplify matters, $a(z), b(z)$ will be taken to be rational functions with fixed poles at the reciprocals of $\mathcal{Z}$.) Since condition (ii) requires that

$$f(z) + f^*(z) \geq 0 \quad \text{on the unit circle,}$$

$\Psi(z)$ is a pseudo-polynomial which is nonnegative on the unit circle. Therefore $\Psi(z)$ has a stable spectral factor $\sigma(z)$ of degree n, i.e., a polynomial solution of

$$\sigma(z)\sigma^*(z) = \Psi(z)$$

having all its zeros in the closed unit disc $\mathbb{D}$, which is unique modulo a factor ± 1. It turns out that the converse is also true. In fact, to each choice of $\sigma(z)$ with n roots in the unit disc, there is one and only one pair $a(z), b(z)$ so that f, defined by (1.3), satisfies (i), (ii) and (iii). Scaling of σ does not affect f, since a and b are scaled by the same factor. Even modulo such scaling, the correspondence $\sigma \mapsto f$ may still fail to be injective, since $a(z)$ and $b(z)$ may have common factors. In fact such common factors do occur when there are solutions of degree less than n.

The Nevanlinna-Pick problem with degree constraint was first considered in [19], where it was shown that, provided the Nevanlinna-Pick problem has a solution, each choice of Ψ corresponds to at least one pair $a(z), b(z)$ such that $f = \frac{b}{a}$ is a solution to the Nevanlinna-Pick problem with degree constraint. It was also conjectured that there is a unique such pair, implying that the solutions (a, b) would be completely parameterized by the choice of zeros of σ. The proof of existence was by means of degree theory and hence nonconstructive. It followed closely the arguments used in [17, 18] to obtain the corresponding existence proof in an important special case, the rational covariance extension problem with degree constraint.

The conjecture was recently established in a stronger form in [6] for the rational covariance extension problem, where it is shown that, under the mild assuption that Ψ is positive on the unit circle, solutions are unique and depend analytically on the problem data. In other words, the rational covariance extension problem is well-posed as an analytic problem. Subsequently, a simpler proof of uniqueness was given in [8] in a form which has been adapted to the rational Nevanlinna-Pick problem in [20], also proving uniqueness for the boundary case when Ψ has zeros on the unit circle.

However, the proofs developed in [18, 19, 6, 8, 20] are all nonconstructive and the question of computing such solutions remained open. This issue was first addressed in [7] for the rational covariance extension problem. In fact, for any positive Ψ, a convex minimization problem was introduced, the solution of which solves the rational co-variance extension problem, thus allowing efficient computation of the corresponding interpolant.

The purpose of the present paper is to develop an analogous computational theory for the rational Nevanlinna-Pick problem. This is done via a generalized entropy functional, akin to that in [7], which incorporates the Nevanlinna-Pick interpolation data and the chosen positive quasipolynomial $\Psi(z)$. The primal problem to maximize this generalized entropy functional requires optimization in infinitely many variables, but the dual problem, which is convex, has finitely many variables, and the minimum corresponds to the required interpolant.

In Section 2 we motivate the Nevanlinna-Pick interpolation problem with degree constraint by examples from systems and control, namely from sensitivity minimization in $\mathcal{H}^\infty$ control, maximal power transfer and spectral estimation. In Section 3 we

review basic facts and set notation. The main results of the paper are then stated in Section 4, in which we define an entropy criterion, which incorporates the data in the rational Nevanlinna-Pick problem. We demonstrate that the infinite-dimensional optimization problem to maximize the entropy criterion has a simple finite-dimensional dual, which in turn is a generalization of the optimization problem in [7]. It is of independent interest that the dual functional contains a barrier-like term, which, in contrast to interior-point methods, does not become infinite on the boundary of the relevant closed convex set but has infinite gradient there. Section 5 contains a proof of the main theorem together with an analysis of the dual problem. In Section 6 we outline a computational procedure for solving the dual problem. In the special case of real interpolants, we develop a state-space procedure, which has the potential to allow extensions to the multivariable case.

2. Motivating examples

To motivate our theory, we now describe a number of applications which lead to Nevanlinna-Pick interpolation problems with degree constraint. We touch upon problems in robust control, in circuit theory and in modeling of stochastic processes. The examples chosen are basic since our aim is only to indicate the range of potential applications of our theory.

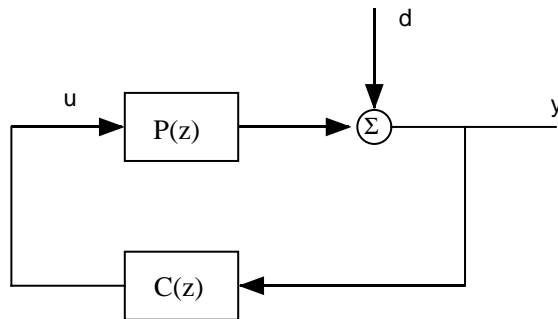**Example 1: Sensitivity minimization.** Consider the following feedback system

Figure 1: Feedback system.

where u denotes the control input to the plant to be controlled, d represents a disturbance, and y is the resulting output, which is also available as an input to a compensator to be designed. Internal stability, and robustness of the output with respect to input disturbances, relies on certain properties of the transfer function from the disturbance to the output, which is given by the sensitivity function $S(z)$ defined via

$$S(z) = (1 - P(z)C(z))^{-1}. \tag{2.1}$$

It is well-known (see, for example, [41, page 100]) that the internal stability of the feedback system is equivalent to the condition that $S(z)$ has all its poles inside the unit disc and satisfies the interpolation conditions

$$S(z_i) = 1, \quad i = 1, 2, \ldots, r \quad \text{and} \quad S(p_j) = 0, \quad j = 1, 2, \ldots, \ell,$$

where $z_1, z_2, \ldots, z_r$ and $p_1, p_2, \ldots, p_\ell$ are the zeros and poles, respectively, of the plant $P(z)$ outside the unit disk. Conversely, if $S(z)$ is any stable, proper rational function which satisfies these interpolation conditions, then $S(z)$ can be represented in the form (2.1) for some rational function $C(z)$.

On the other hand, for disturbance attenuation, S needs to be bounded. The lowest such bound,

$$\alpha_{\mathrm{opt}} = \inf_{S(z_i)=1, S(p_j)=0} \|S\|_\infty, \qquad (2.2)$$

is attained for an S such that $|S(e^{i\theta})| = \alpha_{\mathrm{opt}}$ for all $\theta \in [-\pi, \pi]$. In order to achieve lower sensitivity in selected frequency bands, we must allow higher upper bound $\alpha > \alpha_{\mathrm{opt}}$. Then admissible sensitivity functions S are such that $\frac{1}{\alpha} S(z)$ maps the exterior of the disc into the unit disc. Using the linear fractional transformation $s = \frac{1+z}{1-z}$, which maps the unit disc into the right half plane, the problem then amounts to finding a Carathéodory function

$$f(z) = \frac{\alpha + S(z)}{\alpha - S(z)}$$

which satisfies the interpolation conditions

$$f(p_j) = 1, \quad j = 1, 2, \ldots, \ell \quad \text{and} \quad f(z_i) = \frac{\alpha + 1}{\alpha - 1}, \quad i = 1, 2, \ldots, r.$$

The Macmillan degree of f is the same as the degree of S. The conclusion of our theory is that we can efficiently search over all interpolants of degree at most $n := r + \ell - 1$ to obtain a suitable one. The design parameters which dictate the shape of the sensitivity function are precisely the zeros of

$$\alpha^2 - S(z)S^*(z), \qquad (2.3)$$

which coincide with the zeros of Φ, defined as in (1.4). Hence, they are also zeros of Ψ given by (1.5). The standard approach to shaping the sensitivity function is to formulate a "weighted optimization problem" through a selection of a suitable shaping filter (cf. [15, Chapter 9], [41, Chapter 8]). Typically, a drawback of this approach is an increase in the dimension of the relevant feedback operators by an amount equal to the degree of the shaping filter. Thus the alternative design approach presented here allows for a handle on the degree.

To illustrate our point we consider a simple numerical example which we can work by hand. Let the plant in Figure 1 have the transfer function $P(z) = \frac{1}{z-2}$. This system has one pole and one zero outside the unit disc, namely a pole at 2 and a zero at ∞. Thus the interpolation conditions are $S(\infty) = 1$ and $S(2) = 0$, and, in this simple case, the sensitivity function must be of the form

$$S(z) = \frac{z-2}{z-\beta}, \quad |\beta| < 1.$$

It is easy to see that $\alpha_{\mathrm{opt}} = 2$. We take $\alpha = 2.5$. The one-parameter family of interpolants S such that $\|S\|_\infty \leq 2.5$ is depicted in Figure 2 and parametrized by the zero of (2.3) in $(-1, 1)$, instead of β. Parameterizing the family in terms of such spectral zeros is natural since, as discussed above, it is valid in the general case.

Choosing this spectral zero in the vicinity of $z = -1$, e.g., at -0.9, results in an S with high-pass character. This is

$$S(z) = \frac{z-2}{z-0.2006},$$

with a frequency response shown in Figure 2 with a solid curve. In the same figure, we plot (with dotted curves) the frequency response of S corresponding to a choice of the spectral zero at -0.6, -0.3, 0, 0.3, 0.6, and 0.9.
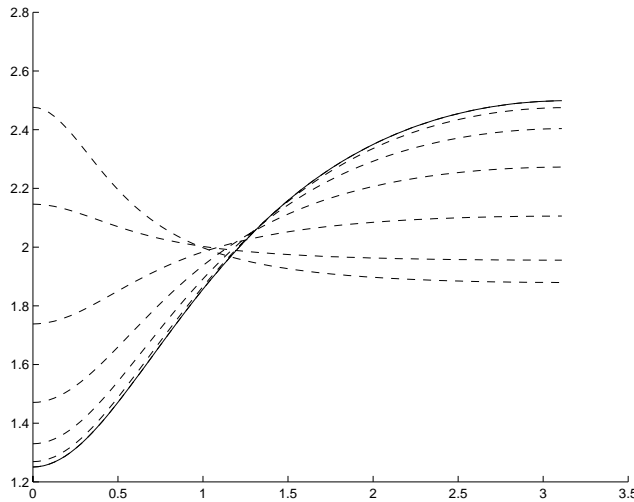

Figure 2: $|S(e^{i\theta})|$ as a function of θ

This simple first-order numerical example was easily worked out by elementary calculations, but higher-order examples require the full power of the theory of this paper.

Example 2: Maximal power transfer. The classical problem of maximal power transfer, first studied by H.W. Bode and reformulated as an interpolation problem by D.C. Youla [39, 10] is illustrated in Figure 3. Here a lossless 2-port coupling is to be designed to achieve a maximal level of power transfer between a generator and a lossy load.
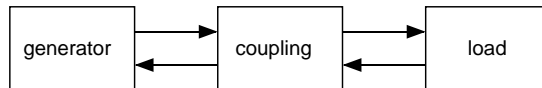

Figure 3: Two port connection.

Let $Z_\ell(s)$ denote the impedance of the passive load and r_g the internal impedance of the generator. The Youla theory rests on the following elements (for details, see [10, Chapter 4]).

(i) $s_1, s_2, \ldots, s_n$ are the right half plane (RHP) *transmission zeros* of $Z_\ell(s)$, i.e., they are the RHP zeros of

$$\Phi_\ell(s) := Z_\ell(s) + Z_\ell(-s),$$

(ii) $Z(s)$ denotes the *driving-point impedance* of the 2-port at the output port when the input port terminates at its reference impedance r_g,

(iii) $B(s)$ is a Blaschke (all-pass) factor with zeros at all open right-half-plane poles of $Z_\ell(-s)$, and

(iv) $\rho(s)$ denotes a *reflection coefficient* at the output port and is given by

$$\rho(s) = B(s)\frac{Z(s) - Z_\ell(-s)}{Z(s) + Z_\ell(s)}.$$

The problem is to maximize the *transducer power gain,*

$$\Phi_{\mathrm{pg}}(s) = 1 - \rho(s)\rho(-s)$$

for $s = i\omega$ at certain preferred range of frequencies ω. This gain is the ratio between average power delivered to the load and the maximum available average power at the source. In order to synthesize a lossless 2-port (e.g., using Darlington synthesis), $Z(s)$ needs to be positive real, which turns out to be the case if and only if $\rho(s)$ is bounded real, i.e., takes values in the unit disc, and satisfies certain interpolation conditions. For simplicity, we assume that the load does not have any transmission zero on the imaginary axis. In this case, the required interpolation conditions are

$$\rho(s_i) = B(s_i) \quad \text{for } i = 1, 2 \ldots, n. \tag{2.4}$$

Thus, the problem of maximizing the transducer power gain amounts to minimizing the H_∞ norm of $\rho(s)$ subject to the constraints (2.4).

Since the transducer power gain is rarely required to be uniform across frequencies, the usual approach to the problem is to specify a desired transducer power gain shape and then to determine whether a solution is feasible. (See [10, Chapter 4]. Also see Helton [26] for an alternative formulation generalizing Youla's theory.) However, in the context of the theory developed in the present paper, we may instead select the zeros of $\Phi_{\mathrm{pg}}(s)$.

As mentioned in the previous example, the theory of the paper applies to any class of functions which is conformally equivalent to positive real functions. Thus we begin by translating the problem to the "discrete-time setting" via the conformal mapping $s = \frac{z-1}{z+1}$, which maps the right-half-plane bijectively onto the complement of the unit disc. We use the notation $g(s) \mapsto \hat{g}(z) := g(\frac{1-z}{1+z})$. In this representation, the transducer power gain becomes

$$\hat{\Phi}_{\mathrm{pg}}(z) = 1 - \hat{\rho}(z)\hat{\rho}(z^{-1}).$$

Next, the conformal mapping

$$f(z) = \frac{1 + \hat{\rho}(z)}{1 - \hat{\rho}(z)}$$

transforms the bounded real function $\hat{\rho}$ to the Carathéodory function f. Then, the roots of $\hat{\Phi}_{\mathrm{pg}}$ are precisely the zeros of

$$\Phi(z) = f(z) + f^*(z),$$

and hence zeros of $\Psi(z)$ in (1.5). The interpolation conditions (2.4) translate directly to interpolation conditions for f via the above transformations. In Section 6 we shall return with a numerical example, which demonstrates the computation theory.

Example 3: Spectral estimation. Consider a scalar zero-mean, stationary Gaussian stochastic process $\{y(t)\}_{\mathbb{Z}}$, and denote by $\Phi(e^{i\theta})$, $\theta \in [-\pi, \pi]$, its power spectral density. Then

$$\Phi(z) = f(z) + f^*(z),$$

where f is a Carathéodory function with the series expansion

$$f(z) = \tfrac{1}{2}c_0 + c_1 z^{-1} + c_2 z^{-2} + \ldots$$

about infinity, where $c_k = \mathrm{E}\{y(t+k)y(t)\}$ for $k = 0, 1, 2, \ldots$. Traditionally, in order to estimate Φ from a realization $y_0, y_1, \ldots, y_N$ of the process, one estimates first a number of covariance samples $c_0, c_1, \ldots, c_n$, where $n << N$, via some *ergodic estimate* such as

$$c_k = \frac{1}{N+1-n} \sum_{t=0}^{N-n} y_{t+k} y_k. \tag{2.5}$$

Knowledge of $c_0, c_1, \ldots, c_n$ imposes certain interpolation conditions on f at infinity. Finding all f satisfying these is the topic which originally motivated the research programs from which the results of the present paper emanated [17, 18, 6, 5, 4, 7, 8]. A complete parameterization of all solutions of degree at most n was provided in [6].

Here we shall take a radically different approach to spectral estimation that is based on nontraditional covariance measurements. The basic idea is to determine covariance estimates after passing the observed time series through a bank of filter with different frequency response and then integrating these statistical measurements in one Markovian model.

Given a number of poles, $p_0, p_1, \ldots, p_n$, of modulus less than one and with $p_0 = 0$, let

$$G_k(z) = \frac{z}{z - p_k} \quad k = 0, 1, \ldots, n \tag{2.6}$$

form a bank of stable filters, driven by y as in Figure 4, and denote by $u_0, u_1, \ldots, u_n$
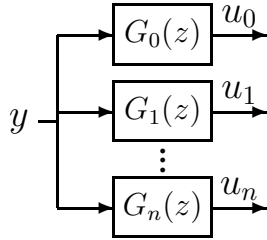

Figure 4: Filter bank.

the corresponding output processes. For simplicity of exposition, we assume that $p_0, p_1, \ldots, p_n$ are distinct and real, hence, for this paper, avoiding the situation with complex pairs of poles. The general case will be presented in [3]. The idea is that the transfer functions G_k are (conjugate) Cauchy kernels in the sense that

$$h(p_k^{-1}) = \int_{-\pi}^{\pi} h(e^{i\theta}) G_k^*(e^{i\theta}) \frac{d\theta}{2\pi} \tag{2.7}$$

for any h which is analytic in $\mathbb{D}^c$ and square-integrable on the unit circle. To see this, note that, if $h(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \ldots$, then, by orthogonality, the integral in (2.7) equals $\sum_{j=0}^{\infty} h_j p_k^j = h(p_k^{-1})$, because $G_k(z) = 1 + p_k z^{-1} + p_k^2 z^{-2} + \ldots$. Therefore, assuming that the filter has come to statistical steady state, the zeroth order covariance lag of the output process u_k is given by

$$c_0(u_k) := \mathrm{E}\{u_k(t)^2\} \;=\; \int_{-\pi}^{\pi} (f(e^{i\theta}) + f^*(e^{i\theta}))|G_k(e^{i\theta})|^2 \frac{d\theta}{2\pi}$$

$$= \; 2\int_{-\pi}^{\pi} f(e^{i\theta})G_k(e^{i\theta})G_k^*(e^{i\theta})\frac{d\theta}{2\pi},$$

and therefore, in view of (2.7), $c_0(u_k) = 2G_k(p_k^{-1})f(p_k^{-1})$. Consequently, the zeroth order covariance data for the outputs of the filter bank supply the interpolation constraints

$$f(p_k^{-1}) = \frac{1}{2}(1 - p_k^2)c_0(u_k) \quad k = 0, 1, \ldots, n, \tag{2.8}$$

where $c_0(u_k), k = 0, 1, \ldots, n$ can be determined via ergodic estimates. An advantage of this approach is that interpolation of the spectrum can be chosen closer to the unit circle in precisely the frequency band where high resolution is desired. We shall return with a numerical example at the end of Section 6.

3. Preliminaries and notation

For simplicity, in this paper we only consider the case where the interpolation points in $\mathbb{Z}$ are distinct. The general case works similarly. Moreover, from now on, *we assume that the Pick matrix* (1.2) *is positive definite*, to avoid the degenerate case where the solution is unique. Also, for convenience, we normalize the problem so that

$$z_0 = \infty \quad \text{and} \quad f(\infty) \text{ is real.}$$

This is done without loss of generality since, first, the transformation

$$z \to \frac{1 - \bar{z}_0 z}{z - z_0}$$

sends an arbitrary z_0 to infinity and is a bianalytic map from $\mathbb{D}^c$ into itself, and, secondly, we can subtract the same imaginary constant from all values w_k without altering the problem.

Denote by $\mathcal{L}_2$ the space of functions which are square-integrable on the unit circle. This is a Hilbert space with inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta})g^*(e^{i\theta})d\theta.$$

Moreover, for an $f \in \mathcal{L}_2$, let

$$f(e^{i\theta}) = \sum_{k=-\infty}^{\infty} f_k e^{-ik\theta}$$

be its Fourier representation. In this notation,

$$\langle f, g \rangle = \sum_{k=-\infty}^{\infty} f_k \bar{g}_k.$$

Next, let $\mathcal{H}_2$ be the standard Hardy space of all functions which are analytic in the *exterior* of the unit disc, $\mathbb{D}^c$, and have square-integrable limits on the boundary

$$\lim_{r \to +1} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(re^{i\theta})|^2 d\theta < \infty.$$

As usual, $\mathcal{H}_2$ is identified with the subspace of $\mathcal{L}_2$ with vanishing negative Fourier coefficients. More precisely, for $f \in \mathcal{H}_2$,

$$f(z) = f_0 + f_1 z^{-1} + f_2 z^{-2} + \dots.$$

The class of all Carathéodory functions in $\mathcal{H}_2$ will be denoted by $\mathcal{C}$. Moreover, we denote by $\mathcal{C}_+$ the subclass of *strictly positive real functions*, whose domain of analyticity includes the unit circle and have positive real part.

Now, consider the data $\mathcal{Z}$ and $\mathcal{W}$ with the standing assumption that $z_0 = \infty$. It is a well-known consequence of Beurling's Theorem [24] that the kernel of the evaluation map $E : \mathcal{H}_2 \to \mathbb{C}^{n+1}$ defined via

$$E(f) = \begin{bmatrix} f(z_0) \\ f(z_1) \\ \vdots \\ f(z_n) \end{bmatrix},$$

is given by

$$\ker(E) = B\mathcal{H}_2,$$

where $B(z)$ is the *Blaschke product*

$$B(z) := z^{-1} \prod_{k=1}^{n} \frac{1 - z_k^{-1} z}{z - \bar{z}_k^{-1}}.$$

Now, let $\mathcal{H}(B)$ be the orthogonal complement of $B\mathcal{H}_2$ in $\mathcal{H}_2$, i.e., the subspace satisfying

$$\mathcal{H}_2 = B\mathcal{H}_2 \oplus \mathcal{H}(B),$$

which will be referred to as the *coinvariant subspace* corresponding to B, since $B\mathcal{H}_2$ is invariant under the shift z^{-1}. Connecting $B\mathcal{H}_2$ to the filter bank in Example 3 in Section 2, we see that, provided $z_k := p_k^{-1}$ for $k = 0, 1, \dots, n$, as suggested by the interpolation problem, the filter-bank transfer functions (2.6) form a basis of $\mathcal{H}(B)$. However, we prefer to work in a basis, $g_0, g_1, \dots, g_n$, for which $g_0 = G_0 = 1$ is orthogonal to the rest of the base elements. Thus we choose

$$g_0(z) = 1, \quad g_k(z) = G_k(z) - 1 = \frac{1}{z\bar{z}_k - 1}, \quad k = 1, 2, \dots, n. \tag{3.1}$$

For future reference, we list the four identities

$$
\begin{aligned}
\langle f, g_0 \rangle &= f(\infty) \\
\langle f, g_k \rangle &= f(z_k) - f(\infty), \quad k = 1, 2, \ldots, n \\
\langle f^*, g_0 \rangle &= \overline{f(\infty)} \\
\langle f^*, g_k \rangle &= 0, \quad k = 1, 2, \ldots, n,
\end{aligned}
\tag{3.2}
$$

which hold for all $f \in \mathcal{H}_2$. In fact, they follow readily from (2.7) and $\langle f^*, G_k \rangle = \bar{f}(\infty)$ with the corresponding conjugated identities. We also remark that there is a natural basis for $\mathcal{H}_2$ obtained by extending $\{g_0, g_1, \ldots, g_n\}$ via

$$
g_k(z) = z^{n+1-k} B(z) \quad \text{for } k = n+1, n+2, \ldots.
\tag{3.3}
$$

The subspace $\mathcal{H}(B)$ consists precisely of all rational functions of the form

$$
p(z) = \frac{\pi(z)}{\tau(z)},
$$

where

$$
\tau(z) = \prod_{k=1}^{n} (z - \bar{z}_k^{-1})
\tag{3.4}
$$

and $\pi(z) = \pi_0 z^n + \pi_1 z^{n-1} + \cdots + \pi_n$ is some polynomial of degree at most n. Therefore, any rational function of degree at most n can be written as

$$
f(z) = \frac{b(z)}{a(z)} \quad \text{where } a, b \in \mathcal{H}(B).
$$

Throughout the paper we shall use such representations for rational functions, and in particular the functions $a(z)$, $b(z)$ and $\sigma(z)$, introduced in Section 1 will belong to $\mathcal{H}(B)$. Hence, $\Psi(z)$, defined by (1.5), will be a symmetric pseudo-polynomial in the basis elements of $\mathcal{H}(B)$ and $\mathcal{H}(B)^*$, where, in particular, $\mathcal{H}(B) \cap \mathcal{H}(B)^*$ is the space of constant functions. In general, the space of pseudo-polynomials in this basis will be denoted by $\mathcal{S}$, and is defined by

$$
\mathcal{S} = \mathcal{H}(B) \vee \mathcal{H}(B)^* = \operatorname{span}\{g_n^*, \ldots, g_1^*, g_0, g_1, \ldots, g_n\}.
\tag{3.5}
$$

In particular, $\Psi \in \mathcal{S}$, and so do ab^* and a^*b. Moreover, we define the subset

$$
\mathcal{S}_+ = \{S \in \mathcal{S} \mid S^* = S \text{ and } S(e^{i\theta}) > 0 \text{ for all } \theta\}
\tag{3.6}
$$

of symmetric and positive functions in $\mathcal{S}$. Any $S \in \mathcal{S}_+$ is a coecive spectral density.

4. A generalized entropy criterion for Nevanlinna-Pick interpolation

Given any function $\Psi(z) \in \mathcal{S}_+$, consider, for each $f \in \mathcal{C}_+$, the generalized entropy gain

$$
\mathbb{I}_\Psi(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[\Phi(e^{i\theta})]\Psi(e^{i\theta})d\theta
\tag{4.1}
$$

where

$$
\Phi(z) := f(z) + f^*(z)
\tag{4.2}
$$

is the corresponding spectral density.

Entropy integrals such as (4.1) have, of course, a long history. For example, see [23, 28] for use of entropy gains in signal processing, and see [33] for use in $\mathcal{H}^\infty$ control. The expression in formula (4.1) reduces to the standard entropy gain in the signal processing literature

$$\mathbb{I}_1(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[f(e^{i\theta}) + f^*(e^{i\theta})]d\theta \tag{4.3}$$

when we set $\Psi = 1$. The unique maximizing function of $\mathbb{I}_1$ subject to the interpolation constraints (1.1) can be obtained by the Nevanlinna-Pick algorithm [38] and is often referred to as the *central* or *maximum entropy* solution.

Since $\Psi(z) \in \mathcal{S}_+$, there is a unique factorization

$$\Psi(z) = \sigma(z)\sigma^*(z) \tag{4.4}$$

such that $\sigma \in \mathcal{H}(B)$ has no zeros in the closure of $\mathbb{D}^c$, i.e., $\sigma(z)$ is a *minimum-phase* spectral factor of $\Psi(z)$. In particular, $\sigma(\infty) \neq 0$. It turns out that there is a unique solution f to the Nevanlinna-Pick problem with degree constraint which maximizes the above entropy functional. Moreover, this solution satisfies

$$f(z) + f^*(z) = \frac{\sigma(z)\sigma^*(z)}{a(z)a^*(z)}, \tag{4.5}$$

where $a \in \mathcal{H}(B)$ is also minimum-phase. Hence the entropy maximization forces a preselected spectral zero structure for the interpolating function, as seen from the following theorem, whose proof will be concluded in the next section, when all necessary lemmas have been established.

Theorem 4.1. *Given a $\Psi \in \mathcal{S}_+$, there exists a unique solution to the constrained optimization problem*

$$\max_{f \in \mathcal{C}_+} \mathbb{I}_\Psi(f) \tag{4.6}$$

subject to the constraints

$$f(z_k) = w_k \quad \text{for } k = 0, 1, \ldots, n. \tag{4.7}$$

Moreover, this solution is of the form

$$f(z) = \frac{b(z)}{a(z)}, \quad a, b \in H(B), \tag{4.8}$$

and hence of degree at most n, and

$$a(z)b^*(z) + b(z)a^*(z) = \Psi(z). \tag{4.9}$$

Conversely, if $f \in \mathcal{C}_+$ satisfies conditions (4.7), (4.8) and (4.9), it is the unique solution of (4.6).

Theorem 4.1 provides a complete parameterization of all pairs (a, b), defining a strictly positive real solution (4.8) to the Nevanlinna-Pick problem with degree constraint, in terms of the zeros of the minimum-phase spectral factor $\sigma(z)$ of the spectral density $\Psi \in \mathcal{S}_+$. These zeros may be chosen arbitrarily in the open unit disc.

Corollary 4.2 (Spectral Zero Assignability Theorem). *For each minimum-phase* $\sigma \in \mathcal{H}(B)$, *normalized so that* $\sigma(\infty) = 1$, *there exists a unique minimum-phase* $a(z) \in \mathcal{H}(B)$ *such that the unique positive real function* $f(z)$ *satisfying* (4.5) *solves the interpolation problem* (4.7). *In other words, there is a bijective correspondence between pairs* (a, b) *solving the Nevanlinna-Pick problem with degree constraint and the set of* n *points in the open unit disc, these being the zeros of* $\sigma(z)$.

The *primal problem* (4.6) is an infinite-dimensional optimization problem. However, since there are only finitely many interpolation constraints, there is a dual problem with finitely many variables. From conditions (4.8) and (4.9), we see that

$$f(z) + f^*(z) = \frac{\Psi(z)}{Q(z)},$$

where $Q(z) = a(z)a^*(z) \in \mathcal{S}_+$. In terms of the basis introduced in Section 3,

$$Q(z) = \bar{q}_n g_n^*(z) + \ldots + \bar{q}_1 g_1^*(z) + q_0 g_0(z) + q_1 g_1(z) + \ldots + q_n g_n(z). \quad (4.10)$$

Since $g_0(z) \equiv 1$, $q_0 = \langle Q, g_0 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q(e^{i\theta}) d\theta$. Therefore, since Q is positive on the circle, q_0 is real and positive. Hence, we may identify Q with the vector $q := (q_0, q_1, \ldots, q_n)$ of coefficients belonging to the set

$$\mathcal{Q}_+ := \{ q \in \mathbb{R} \times \mathbb{C}^n \mid Q(e^{i\theta}) > 0 \quad \text{for all } \theta \}.$$

Clearly, $q \in \mathcal{Q}_+$ if and only if $Q \in \mathcal{S}_+$. As we shall see shortly the q-parameters will essentially be the Lagrange multipliers for the dual problem.

Now, consider the Lagrange function

$$L(f, \lambda) = \mathbb{I}_\Psi(f) + \lambda_0(w_0 - f(z_0)) + 2\text{Re}\left\{ \sum_{k=1}^{n} \bar{\lambda}_k [w_k - f(z_k)] \right\}. \quad (4.11)$$

Since the primal problem (4.6) amounts to maximizing a strictly concave function over a convex region, the Lagrange function has a saddle point [32, p.458] provided there is a stationary point in $\mathcal{C}_+$, and, in this case, the optimal Lagrange vector $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n+1}$ can be determined by solving the *dual problem* to minimize

$$\rho(\lambda) = \max_{f \in \mathcal{C}_+} L(f, \lambda). \quad (4.12)$$

Now, consider the linear map $\boldsymbol{\lambda} : \mathcal{Q}_+ \rightarrow \mathbb{R} \times \mathbb{C}^n$ defined by

$$\lambda_0 = 2(q_0 - \text{Re} \sum_{j=1}^{n} q_j)$$
$$\lambda_k = q_k \quad \text{for } k = 1, 2, \ldots, n. \quad (4.13)$$

The function ρ takes finite values only for a subset of $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_n) \in \mathbb{R} \times \mathbb{C}^n$ and, in particular, on the set

$$\Lambda_+ := \boldsymbol{\lambda}(\mathcal{Q}_+). \quad (4.14)$$

We have the following proposition, the proof of which is deferred to Appendix A.

Proposition 4.3. *For each $\lambda \in \Lambda_+$, the map $f \mapsto L(f, \lambda)$ has a unique maximum in $\mathcal{C}_+$, , and it is given by*

$$f(z) + f^*(z) = \frac{\Psi(z)}{Q(z)}, \tag{4.15}$$

where Q is defined from (4.10) and $q = \boldsymbol{\lambda}^{-1}(\lambda)$.

This proposition defines, for each $\lambda \in \Lambda_+$, a function $f_\lambda \in \mathcal{C}_+$, which, as is easy to check, can be written as

$$f_\lambda(z) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{z + e^{i\theta}}{z - e^{i\theta}} \frac{\Psi(e^{i\theta})}{Q_\lambda(e^{i\theta})} d\theta$$

in terms of the corresponding $Q_\lambda \in \mathcal{S}_+$. We want to show that there is a unique minimizing λ, denoted $\hat{\lambda}$, such that $f_{\hat{\lambda}} \in \mathcal{C}_+$ satisfies the interpolation condition (4.7). In this case, setting $\hat{f} := f_{\hat{\lambda}}$,

$$\rho(\lambda) \geq \rho(\hat{\lambda}) = L(\hat{f}, \hat{\lambda}), \quad \text{for all } \lambda \in \Lambda_+.$$

Now, for any $f \in \mathcal{C}_+$ which satisfies the interpolation constraints (4.7),

$$\mathbb{I}_\Psi(f) = L(f, \hat{\lambda}) \leq L(\hat{f}, \hat{\lambda}).$$

In particular, this holds for $f = \hat{f}$ so that $\mathbb{I}_\Psi(\hat{f}) = L(\hat{f}, \hat{\lambda})$. Hence,

$$\mathbb{I}_\Psi(f) \leq \mathbb{I}_\Psi(\hat{f}) = \rho(\hat{\lambda}) \leq \rho(\lambda), \tag{4.16}$$

if f satisfies the interpolation constraints. Consequently, if we can show that ρ has a minimum $\hat{\lambda} \in \Lambda_+$, then $\mathbb{I}_\Psi$ has a maximum in $\mathcal{C}_+$, and the optimal values of the two problems coincide.

It turns out to be more convenient to use the q's as dual variables.

Proposition 4.4. *The dual functional (4.12) is*

$$\rho(\boldsymbol{\lambda}(q)) = \mathbb{J}_\Psi(q) + c,$$

where

$$\mathbb{J}_\Psi(q) = 2w_0 q_0 + 2Re\left\{ \sum_{k=1}^{n} (w_k - w_0)\bar{q}_k \right\} - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[Q(e^{i\theta})]\Psi(e^{i\theta})d\theta, \tag{4.17}$$

and

$$c := \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\log[\Psi(e^{i\theta})] - 1 \right) \Psi(e^{i\theta})d\theta.$$

We are now in a position to formulate the dual version of Theorem 4.1, the proof of which will be deferred to the next section. For simplicity, we remove the constant term c, which does not affect the optimization.

Theorem 4.5. *Given a $\Psi \in \mathcal{S}_+$, there exists a unique solution to the dual problem*

$$\min_{q \in \mathcal{Q}_+} \mathbb{J}_\Psi(q). \tag{4.18}$$

Moreover, to the minimizing q there corresponds an $f \in \mathcal{C}_+$ such that

$$\frac{\Psi(z)}{Q(z)} = f(z) + f^*(z) \tag{4.19}$$

where Q is given by (4.10). *Moreover, this function f satisfies conditions* (4.7), (4.8) *and* (4.9) *in Theorem* 4.1, *namely*

$$f(z_k) = w_k \quad for \ k = 0, 1, \ldots, n, \tag{4.20}$$

$$f(z) = \frac{b(z)}{a(z)}, \quad a, b \in \mathcal{H}(B), \tag{4.21}$$

$$\Psi(z) = a(z)b^*(z) + b(z)a^*(z). \tag{4.22}$$

Conversely, any $f \in \mathcal{C}_+$ which satisfies these conditions can be constructed from the unique solution of (4.18) *via* (4.19).

We conclude by noting that if the problem data is real or self-conjugate, and Ψ is real, then both the function $f(z)$ constructed above, and the function $\bar{f}(\bar{z})$, satisfy the conditions of Theorems 4.1 and 4.5 so that, by uniqueness, they must coincide.

Corollary 4.6. *Assume that the the sets $\mathcal{Z}$ and $\mathcal{W}$ are self-conjugate and that $w_k = \bar{w}_j$ whenever $z_k = \bar{z}_j$, and that Ψ is real. Then, the optimizing functions f, Q in Theorems* 4.1 *and* 4.5 *have real coefficients. In particular, there is a unique pair of real functions $a(z)$ and $b(z)$ in $\mathcal{H}(B)$, devoid of zeros in closure of $\mathbb{D}^c$, such that*

$$\Psi(z) = a(z)b(z^{-1}) + a(z^{-1})b(z)$$

$$f(z) = \frac{b(z)}{a(z)} \in \mathcal{C}_+$$

$$f(z_k) = w_k \quad for \ k = 0, 1, ..., n.$$

We shall return to the special case covered in Corollary 4.6 in Section 6, and we shall refer to it as the *self-conjugate case*.

5. The convex optimization problem

In this section, we shall analyze the functional $\mathbb{J}_\Psi(q)$, constructed in the previous section. We shall show that it has a unique minimum in $\mathcal{Q}_+$, which is instrumental in proving Theorem 4.1 and Theorem 4.5. To this end, we first extend $\mathbb{J}_\Psi(q)$ to the closure $\mathcal{Q}$ of $\mathcal{Q}_+$, and consider

$$\mathbb{J}_\Psi : \mathcal{Q} \longrightarrow \mathbb{R} \cup \{\infty\}.$$

Proposition 5.1. *The functional $\mathbb{J}_\Psi(q)$ is a C^∞ function on $\mathcal{Q}_+$ and has a continuous extension to the boundary that is finite for all $q \neq 0$. Moreover, $\mathbb{J}_\Psi$ is strictly convex, and $\mathcal{Q}$ is a closed and convex set.*

This proposition, along with Propositions 5.2 and 5.4 below, are analogous to related results in [7], developed for the covariance extension problem. Their proofs are similar, *mutatis mutandis*, to those developed in [7], except for Lemma 5.3 below. The complete proofs are adapted to the present framework and included in the appendix for the convenience of the reader.

In order to ensure that $\mathbb{J}_\Psi$ achieves a minimum on $\mathcal{Q}$, it is important to know whether $\mathbb{J}_\Psi$ is proper, i.e., whether $\mathbb{J}_\Psi^{-1}(K)$ is compact whenever K is compact. In this case, of course, a unique minimum will exist.

Proposition 5.2. *For all $r \in \mathbb{R}$, $\mathbb{J}_\Psi^{-1}(-\infty, r]$ is compact. Thus $\mathbb{J}_\Psi$ is proper (i.e., $\mathbb{J}_\Psi^{-1}(K)$ is compact whenever K is compact) and bounded from below.*

The proof of this proposition, given in the appendix, relies on the analysis of the growth of $\mathbb{J}_\Psi$, which entails a comparison of linear and logarithmic growth. To this end, the following lemma is especially important. We note that its proof is the only point in our construction and argument in which we use the Pick condition in an essential way. Denote the linear part of $\mathbb{J}_\Psi(q)$ by

$$J(q) := 2w_0 q_0 + 2\mathrm{Re}\left\{ \sum_{k=1}^n (w_k - w_0)\bar{q}_k \right\}$$

$$= 2w_0 q_0 + \sum_{k=1}^n (w_k - w_0)\bar{q}_k + \sum_{k=1}^n (\bar{w}_k - w_0)q_k. \tag{5.1}$$

Lemma 5.3. *For each nonzero $q \in \mathcal{Q}$, $J(q) > 0$.*

Proof. Since $P > 0$, there exists a strictly positive real interpolant. Choose an arbitrary such interpolant, and denote it by f. Then, recalling that $z_0 = \infty$, (3.2) yields

$$2w_0 = \langle f + f^*, g_0 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(e^{i\theta}) + f^*(e^{i\theta})]g_0^*(e^{i\theta})d\theta$$

and

$$w_k - w_0 = \langle f + f^*, g_k \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(e^{i\theta}) + f^*(e^{i\theta})]g_k^*(e^{i\theta})d\theta$$

for $k = 1, 2, \ldots, n$. For any q in $\mathcal{Q}$, we compute

$$J(q) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(e^{i\theta}) + f^*(e^{i\theta})]Q(e^{i\theta})d\theta \geq 0,$$

and $J(q) = 0$ if and only if $Q \equiv 0$. $\square$

Finally, we need to exclude the possibility that the minimum occurs on the boundary. This is the content of the following proposition, also proved in the appendix.

Proposition 5.4. *For $\Psi \in \mathcal{S}_+$, the functional $\mathbb{J}_\Psi$ never attains a minimum on the boundary $\partial\mathcal{Q}$.*

Hence we have established that $\mathbb{J}_\Psi(q)$ is strictly convex, has compact sublevel sets and the minimum does not occur on the boundary of $\mathcal{Q}$. Consequently, it has a unique minimum, which occurs in the open set $\mathcal{Q}_+$. Clearly, this minimum point will be a stationary point with vanishing gradient. As the following lemma shows, the gradient

becomes zero precisely when the interpolation conditions are satisfied, and in fact the value of the gradient depends only on the mismatch at the interpolation points.

Before stating the lemma, however, let us, for the convenience of the reader, review a few basic facts from complex function theory. In what follows, it will be convenient to use complex partial differential operators acting on smooth, but not necessarily complex analytic, functions. In particular, if we write the complex vector $q_k = x_k + iy_k$ as a sum of real and imaginary parts, this defines the differential operators

$$\frac{\partial}{\partial q_k} = \frac{1}{2}\left(\frac{\partial}{\partial x_k} - i\frac{\partial}{\partial y_k}\right) \quad \text{and} \quad \frac{\partial}{\partial \bar{q}_k} = \frac{1}{2}\left(\frac{\partial}{\partial x_k} + i\frac{\partial}{\partial y_k}\right)$$

which operate on smooth functions. Indeed, the second operator is the Cauchy-Riemann operator which characterizes the analytic functions F of q_k via

$$\frac{\partial F}{\partial \bar{q}_k} = 0.$$

And, for example, while conjugation, viewed as the function defined by $\bar{q}_k = x_k - iy_k$, is of course not analytic, it is smooth and satisfies

$$\frac{\partial \bar{q}_k}{\partial q_k} = 0 \quad \text{and} \quad \frac{\partial \bar{q}_k}{\partial \bar{q}_k} = 1.$$

Lemma 5.5. *At any point $q \in \mathcal{Q}_+$ the gradient of $\mathbb{J}_\Psi$ is given by*

$$\frac{\partial \mathbb{J}_\Psi}{\partial q_0} = 2[w_0 - f(z_0)], \tag{5.2}$$

$$\frac{\partial \mathbb{J}_\Psi}{\partial \bar{q}_k} = [w_k - f(z_k)] - [w_0 - f(z_0)], \quad for\ k = 1, 2, \ldots, n, \tag{5.3}$$

where f is the $\mathcal{C}_+$ function satisfying

$$f(z) + f^*(z) = \frac{\Psi(z)}{Q(z)} \tag{5.4}$$

with $Q(z) \in \mathcal{S}$ corresponding to q as in (4.10).

Proof. The existence of a function f as claimed in the statement is obvious by virtue of the fact that $\frac{\Psi(z)}{Q(z)}$ is bounded and greater than zero on the unit circle. Recalling that

$$\frac{\partial q_k}{\partial \bar{q}_k} = 0,$$

for $k > 0$ we have

$$\frac{\partial \mathbb{J}_\Psi}{\partial \bar{q}_k} = (w_k - w_0) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{g_k^*(e^{i\theta})}{Q(e^{i\theta})}\Psi(e^{i\theta})d\theta$$
$$= (w_k - w_0) - \langle f + f^*, g_k \rangle.$$

Since $f \in \mathcal{H}_2$, this is the same as (5.3). To see this, use (3.2) and note that $z_0 = \infty$. For the case $k = 0$ we need to take the real derivative:

$$\frac{\partial \mathbb{J}_\Psi}{\partial q_0} = 2w_0 - \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{g_0(e^{i\theta})}{Q(e^{i\theta})}\Psi(e^{i\theta})d\theta$$
$$= 2w_0 - \langle f + f^*, g_0 \rangle,$$

which, again using (3.2), yields (5.2). $\square$

We are now prepared for the proof of our main results.

Proof of Theorem 4.5. Propositions 5.1, 5.2 and 5.4 establish the existence of a unique minimum in $q \in \mathcal{Q}_+$. Then Lemma 5.5 shows that the interpolation conditions are met for the corresponding $\mathcal{C}_+$-function f satisfying (5.4). The construction of such a function proceeds as follows. Since $Q \in \mathcal{S}_+$ and is rational, it admits a rational spectral factorization $Q(z) = a(z)a^*(z)$, where $a(z) = \frac{\alpha(z)}{\tau(z)}$ with $\alpha(z)$ a stable polynomial of degree at most n. Hence, $a \in \mathcal{H}(B)$. Then, we solve the linear equation $a(z)b^*(z) + b(z)a^*(z) = \Psi(z)$ for b. This linear equation has always a unique solution because a has no zeros in $\mathbb{D}^c$; cf. the discussion in [9]. Then $f(z) = \frac{b(z)}{a(z)}$, and all conditions of the theorem are satisfied.

Conversely, given an $f \in \mathcal{C}_+$ satisfying (4.21) and (4.22), a unique $q \in \mathcal{Q}_+$ can be obtained from (5.4). Finally, in view of Lemma 5.5, the interpolation conditions (4.20) imply that the gradient of $\mathbb{J}_\Psi$ for the corresponding q is zero. Thus it is the unique minimizing q. $\square$

Proof of Theorem 4.1. Let us denote by $\hat{q}$ the minimizing q in Theorem 4.5. Then, since $\hat{q} \in \mathcal{Q}_+$, we have $\hat{\lambda} := \lambda(\hat{q}) \in \Lambda_+$ in the notation of Proposition 4.4. Let $\hat{f}$ be the unique corresponding $f \in \mathcal{C}_+$ defined via Proposition 4.3. By Theorem 4.5, $\hat{f}$ satisfies conditions (4.7), (4.8) and (4.9). Then, since thus $\hat{f}$ satisfies the interpolation condition, (4.16) holds, implying that $\hat{f}$ is the maximizing f of Theorem 4.1. Conversely, if $\hat{f}$ satisfies (4.7), (4.8) and (4.9), by Theorem 4.5, the corresponding $\hat{q}$, defined via (4.19), is the unique maximizing solution to the dual problem. Therefore, it follows in the same way as above, that $\hat{f}$ is the unique maximizing solution to the primal problem. $\square$

An interesting, and useful, aspect of the functionals studied using interior point methods is that they contain a barrier term, which is infinite on the boundary of the closed convex set in question. At first glance, the logarithmic integrand in $\mathbb{J}_\Psi(q)$ might seem to be a barrier-like term but, as we have seen in Section 5, by a theorem of Szegö, the logarithmic integrand is in fact integrable for nonzero Q having zeros on the boundary of the unit circle. Hence $\mathbb{J}_\Psi(q)$ does not become infinite on the entire boundary $\partial \mathcal{Q}$ of $\mathcal{Q}$. Nonetheless, $\mathbb{J}_\Psi(q)$ has a very interesting barrier-type property as described in the following proposition and proven in the Appendix.

Proposition 5.6. *The dual functional $\mathbb{J}_\Psi(q)$ has an infinite gradient on the boundary $\partial \mathcal{Q}$.*

As far as computation is concerned, this is a a useful property of the convex optimization problem.

6. Computational procedure

Given $\Psi(z)$, define the class $\mathcal{P}$ of (strictly) positive real functions

$$f(z) = \frac{b(z)}{a(z)}, \quad a, b \in \mathcal{H}(B)$$

having the property that

$$a(z)b^*(z) + b(z)a^*(z) = \Psi(z). \tag{6.1}$$

We want to determine the unique function in $\mathcal{P}$ which also satisfies the interpolation conditions. To this end, we shall construct a sequence of functions,

$$f^{(0)}, f^{(1)}, f^{(2)}, \cdots \in \mathcal{P}$$

which converges to the required interpolant.

As before, we may write (6.1) as

$$f(z) + f^*(z) = \frac{\Psi(z)}{Q(z)}, \tag{6.2}$$

where $Q \in \mathcal{S}_+$ satisfies

$$a(z)a^*(z) = Q(z). \tag{6.3}$$

It is easy to see that this defines a bijection

$$\mathcal{I} : \mathcal{Q}_+ \to \mathcal{P} : \quad Q \mapsto f. \tag{6.4}$$

To see this, note that

$$a(z) = \frac{\alpha(z)}{\tau(z)}, \quad b(z) = \frac{\beta(z)}{\tau(z)} \quad \text{and} \quad \Psi(z) = \frac{d(z, z^{-1})}{\tau(z)\tau^*(z)},$$

where $\alpha(z)$ and $\beta(z)$ are polynomials of at most degree n and $d(z, z^{-1})$ is a pseudo-polynomial, also of at most degree n. Since $f = \beta/\alpha$ is strictly positive real, both $\alpha(z)$ and $\beta(z)$ must be Schur polynomials. Then, determine $\alpha(z)$ via a stable polynomial factorization

$$\alpha(z)\alpha^*(z) = \tau(z)\tau^*(z)Q(z), \tag{6.5}$$

and solve the linear system

$$\alpha(z)\beta^*(z) + \beta(z)\alpha^*(z) = d(z, z^{-1}) \tag{6.6}$$

for β. In fact, (6.6) is a linear (Hankel + Toeplitz) system $S(\alpha)\beta = d$ in the coefficients of the polynomials, which is nonsingular since $\alpha(z)$ is a Schur polynomial; see, e.g., [9]. Then

$$f(z) = \frac{\beta(z)}{\alpha(z)}.$$

Given an $f \in \mathcal{P}$ we can determine the corresponding gradient of $\mathbb{J}_\Psi(q)$ by means of Lemma 5.5. The following lemma gives the equations for the $(n+1) \times (n+1)$ Hessian matrix

$$H(q) = \left[\frac{\partial^2 \mathbb{J}_\Psi}{\partial \bar{q}_k \partial \bar{q}_\ell} \right]_{k,\ell=0}^n. \tag{6.7}$$

Lemma 6.1. *Let $h(z)$ be the unique positive real function such that*

$$h(z) + h^*(z) = \frac{\Psi(z)}{Q(z)^2} \tag{6.8}$$

and $h(z_0)$ is real. Then the Hessian (6.7) is given by

$$H_{k\ell}(q) = \begin{cases} \frac{z_k}{z_\ell - z_k}h(z_k) + \frac{z_\ell}{z_k - z_\ell}h(z_\ell) + h(z_0) & \text{for } k \neq \ell; k, \ell > 0 \\ -z_k h'(z_k) - h(z_k) + h(z_0) & \text{for } k = \ell > 0 \\ h(z_k) - h(z_0) & \text{for } k > 0, \ell = 0 \\ h(z_\ell) - h(z_0) & \text{for } k = 0, \ell > 0 \\ 2h(z_0) & \text{for } k = \ell = 0, \end{cases} \tag{6.9}$$

where $h'(z)$ is the derivative of $h(z)$.

Next, we turn to the computational procedure, which will be based on Newton's method [31, 32]. We need an $f^{(0)} \in \mathcal{P}$, and a corresponding $Q^{(0)}$ defined via (6.2), as an initial condition. We may choose $Q^{(0)} = 1$. Each iteration in our procedure consists of four steps and updates the pair f, Q to $\hat{f}, \hat{Q}$, in the following way:

Step 1. Given f, let $\nabla \mathbb{J}_\Psi(q)$ be the gradient defined by (5.2) and (5.3).

Step 2. Determine the unique positive real function h satisfying (6.8), which is a linear problem of the same type as the one used to determine f from Q. In fact, exchanging $\alpha(z)$ for $\alpha(z)^2$ and $d(z, z^{-1})$ for $v(z, z^{-1}) = \tau(z)\tau^*(z)d(z, z^{-1})$ in (6.6) we obtain

$$h(z) = \frac{\beta(z)}{\alpha(z)^2} \quad \text{where } \beta = S(\alpha^2)^{-1}v.$$

The Hessian $H(q)$ is then determined from h as in Lemma 6.1.

Step 3. Update $Q(z)$ by applying Newton's method to the function $\mathbb{J}_\Psi$. A Newton step yields

$$\bar{q}_{\text{update}} = \bar{q} - \lambda H(q)^{-1}\nabla \mathbb{J}_\Psi(q),$$

where $\lambda \in (0, 1]$ needs to chosen so that

$$Q_{\text{update}}(e^{i\theta}) > 0 \quad \text{for all } \theta. \tag{6.10}$$

This positivity condition is tested in Step 4.

Step 4. Factor Q_{update} as in (6.3). This is also a test for condition (6.10). If the test fails, return to Step 3 and decrease the step size λ. If not, check whether the norm of $\nabla \mathbb{J}_\Psi(q_{\text{update}})$ is sufficiently small. Recall that this norm quantifies the interpolation error, as can be seen from Lemma 5.5. If this error is small, stop; otherwise, use the linear procedure above to determine the next iterate f_{update}. Then, set $f := f_{\text{update}}$ and return to Step 1.

The computations can be carried out quite efficiently using state space descriptions. We restrict our attention to the self-conjugate case, where both $\mathcal{Z}$ and $\mathcal{W}$ are self-conjugate and $w_k = \bar{w}_j$ whenever $z_k = \bar{z}_j$, and $\Psi(z)$ is real. (See Corollary 4.6.) In particular, we develop the steps of the algorithm so as to avoid complex arithmetic.

It is easy to see that, in this case,

$$\tau(z) := \prod_{k=1}^{n}(z - \bar{z}_k^{-1}) = z^n + \tau_1 z^{n-1} + \cdots + \tau_n \tag{6.11}$$

is a real polynomial and

$$B(z) = z^{-1}\frac{\tau_*(z)}{\tau(z)} \tag{6.12}$$

is a real function, where $\tau_*(z) := 1 + \tau_1 z + \cdots + \tau_n z^n$ is the reverse polynomial. For the rest of this section, we shall be concerned with real interpolation functions.

Any real function $h \in \mathcal{H}(B)$ admits a state space representation of the form

$$h(z) = h_0 + c(zI - A)^{-1}b_h, \tag{6.13}$$

where (A, b_h, c) are taken in the observer canonical form

$$A = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -\tau_n & -\tau_{n-1} & \dots & -\tau_1 \end{bmatrix} \qquad b_h = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \tag{6.14}$$

$$c = \begin{bmatrix} 1, 0, \dots, 0 \end{bmatrix},$$

$h_1, h_2, \dots, h_n$ being the Markov parameters in the Taylor expansion

$$h(z) = h_0 + h_1 z^{-1} + \cdots + h_n z^{-n} + \dots$$

about infinity. We shall use the compact notation

$$\mathbf{h} = \left[\begin{array}{c|c} A & b_h \\ \hline c & h_0 \end{array} \right]$$

for this representation, and keep A and c fixed when representing real functions in $\mathcal{H}(B)$. Since the function (6.13) is completely determined by the Markov parameters h_0, b_h, we shall refer to them as the *Markov coordinates* of the function (6.13). Alternatively, $h(z)$ can also be represented with respect to the standard basis in $\mathcal{H}(B)$ as

$$h(z) = h_0 + \sum_{j=1}^{n} \eta_j g_j(z), \tag{6.15}$$

where, of course, $\eta_1, \dots, \eta_n$ are complex numbers. Finally, any $h \in \mathcal{H}(B)$ can be uniquely identified by its values at $\mathcal{Z}$,

$$\{h(z_0), h(z_1), \dots, h(z_n)\}.$$

The correspondence between these three alternative representations is the content of the following lemma.

Lemma 6.2. *Let V be the Vandermonde matrix $V := [\bar{z}_k^{-j}]_{j,k}$, and G the matrix $G := [g_k(z_j)]_{j,k}$. Then, for any $h \in \mathcal{H}(B)$,*

$$b_h = V\eta,$$

where $\eta = (\eta_1, \eta_2, \ldots, \eta_n)'$ is defined via (6.15), and

$$\begin{bmatrix} h(z_1) - h_0 \\ h(z_2) - h_0 \\ \vdots \\ h(z_n) - h_0 \end{bmatrix} = G\eta.$$

Moreover, G and V are invertible.

Proof. The first correspondence follows immediately from (6.15) and the expansion

$$g_k(z) = \bar{z}_k^{-1} z^{-1} + \bar{z}_k^{-2} z^{-2} + \bar{z}_k^{-3} z^{-3} + \ldots.$$

The second correspondence also follows from (6.15). Moreover,

$$G = \left[\frac{\bar{z}_k^{-1} z_\ell^{-1}}{1 - \bar{z}_k^{-1} z_\ell^{-1}} \right]_{k,\ell=1}^n$$

can be written as $G = \frac{1}{2}Z^*PZ$, where Z is the diagonal matrix $\mathrm{diag}(z_1^{-1}, \ldots, z_n^{-1})$ and P is the Pick matrix for $\mathcal{Z} = \{z_1, \ldots, z_n\}$ and $\mathcal{W} = \{1, \ldots, 1\}$. Since there is a unique function, namely $f \equiv 1$, satisfying this interpolation data, P is positive definite, establishing the invertibility of G. Finally, the Vandermonde matrix V is invertible since the points in $\mathcal{Z}$ are distinct. $\square$

We now reformulate the steps of the algorithm given in Section 6 in terms of the real Markov coordinates of the relevant functions. We shall consistently work with functions in $\mathcal{H}(B)$. Therefore, as $f \notin \mathcal{H}(B)$, we form

$$\hat{f} := \Pi_{\mathcal{H}(B)} f,$$

where $\Pi_{\mathcal{H}(B)}$ denotes orthogonal projection onto $\mathcal{H}(B)$. Since $\hat{f} = f + Bg$ for a suitable $g \in \mathcal{H}^2$, it follows that

$$\hat{f}(z_k) = f(z_k) \quad \text{for } k = 0, 1, \ldots, n.$$

Next, define $w(z)$ to be the unique function in $\mathcal{H}(B)$ such that

$$w(z_k) = w_k \quad \text{for } k = 0, 1, \ldots, n. \tag{6.16}$$

This function has the form $w(z) = \pi(z)/\tau(z)$, where $\tau(z)$ is given by (6.11), and where the coefficients of the polynomial π, of degree at most n, can be determined by solving the linear (Vandermonde) system of equations defined by (6.16). The gradient of $\mathbb{J}_\Psi$ in Lemma 5.5 can then be expressed in terms of the "error function"

$$\rho(z) := w(z) - \Pi_{\mathcal{H}(B)} f(z), \tag{6.17}$$

which also belongs to $\mathcal{H}(B)$. In fact,

$$\rho(z_k) := w_k - f(z_k). \tag{6.18}$$

Moreover, we introduce an $\mathcal{H}(B)$-representation for any $Q \in \mathcal{S}$ and any given $\Psi \in \mathcal{S}$ by writing

$$Q(z) = q(z) + q^*(z), \quad \Psi(z) = \psi(z) + \psi^*(z),$$

where $q, \psi \in \mathcal{H}(B)$ are positive real. Finally, we represent q and ψ by their respective Markov coordinates (x, x_0) and (y, y_0), respectively, in the standard state-space representation described above, i.e.,

$$\mathbf{q} = \left[\begin{array}{c|c} A & x \\ \hline c & x_0 \end{array}\right] \quad \text{and} \quad \psi = \left[\begin{array}{c|c} A & y \\ \hline c & y_0 \end{array}\right].$$

We begin with the state-space implementation of Step 1 in the computational scheme described above. In this context, we have the following version of Lemma 5.5.

Proposition 6.3. *Given an $f \in \mathcal{P}$, let q be the positive real part of $Q := \mathcal{I}^{-1}f$, where $\mathcal{I}$ is defined as in (6.4). Moreover, let $\rho(z)$ be given by (6.17), and denote by (x_0, x) and (r_0, r) the Markov coordinates of $q(z)$ and $\rho(z)$, respectively. Then*

$$\frac{\partial \mathbb{J}_\Psi}{\partial x_0} = 4r_0,$$

$$\frac{\partial \mathbb{J}_\Psi}{\partial x} = (V^*)^{-1}GV^{-1}r,$$

where G and V be defined as in Lemma 6.2. The $n \times n$ matrix $(V^*)^{-1}GV^{-1}$ is a real matrix.

Proof. Since $q_0 = 2x_0$ and $w_0 - f(z_0) = \rho(z_0) = r_0$, the derivative with respect to x_0 follows immediately from (5.2). Next, applying Lemma 6.2, we see that

$$\left[\frac{\partial x_j}{\partial q_k}\right]_{j,k=1}^n = V,$$

and that $\rho(z_k) - r_0$ is the k:th entry in $GV^{-1}r$. Moreover, by (6.18), we have

$$\rho(z_k) - r_0 = [w_k - f(z_k)] - [w_0 - f(z_0)]$$

for $k = 1, 2, \ldots, n$. Finally, using equation (5.3) and defining $\bar{q} := (\bar{q}_1, \bar{q}_2, \ldots, \bar{q}_n)'$, we obtain

$$\frac{\partial \mathbb{J}_\Psi}{\partial x} = \left(\left[\frac{\partial x}{\partial \bar{q}}\right]^{-1}\right)' \frac{\partial \mathbb{J}_\Psi}{\partial \bar{q}} = (V^*)^{-1}GV^{-1}r$$

establishing the rest of the proposition. $\square$

It remains to determine the error function ρ. For this we need the projection $\hat{f} := \Pi_{\mathcal{H}(B)}f$.

Proposition 6.4. *Suppose $f \in \mathcal{P}$ has the state space representation*

$$\mathbf{f} = \left[\begin{array}{c|c} A_f & b_f \\ \hline c_f & f_0 \end{array}\right]$$

and that A and c are defined as in (6.14). Then $\hat{f} := \Pi_{\mathcal{H}(B)}f$ is given by

$$\Pi_{\mathcal{H}(B)}f = f_0 + c(zI - A)^{-1}R^{-1}Sb_f,$$

where R and S are the unique solutions of the Lyapunov equation

$$R = A'RA + c'c \tag{6.19}$$

and the Sylvester equation

$$S = A'SA_f + c'c_f, \tag{6.20}$$

respectively.

Proof. Since $\hat{f} := \Pi_{\mathcal{H}(B)} f \in \mathcal{H}(B)$ and $\hat{f}(\infty) = f(\infty)$, there is a representation

$$\hat{f}(z) = f_0 + c(zI - A)^{-1}k$$

for some $k \in \mathbb{R}^n$. Now, $f - \hat{f} \perp \mathcal{H}(B)$, and hence

$$\langle h, f - \hat{f} \rangle = 0, \quad \text{for all } h \in \mathcal{H}(B). \tag{6.21}$$

Since the functions $h \in \mathcal{H}(B)$ are precisely those of the form (6.13), (6.21) is equivalent to

$$(b_h)'(Sb_f - Rk) = 0, \quad \text{for all } b_h \in \mathbb{R}^n, \tag{6.22}$$

where

$$R = \frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{-i\theta}I - A')^{-1}c'c(e^{i\theta}I - A)^{-1}d\theta$$

and

$$S = \frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{-i\theta}I - A')^{-1}c'c_f(e^{i\theta}I - A_f)^{-1}d\theta.$$

It is well-known and straight-forward to show that R satisfies the Lyapunov equation (6.19) and S the Sylvester equation (6.20). In fact, they are the unique solutions of these equations, as the eigenvalues of both A and A_f are located in the open unit disc [16]. Since (c, A) is an observable pair, R is positive definite, and hence invertible. Then the proposition follows from (6.22). $\square$

Consequently, the state-space version of Step 1 amounts to solving first the Lyapunov equation (6.19) and the Sylvester equation (6.20) to obtain $\Pi_{\mathcal{H}(B)} f$ via Proposition 6.4. Then the gradient is determined from (6.17) as described in Proposition 6.3. Step 2 is developed along the same lines as in Step 1 by instead representing relevant functions in $\mathcal{H}(B^2)$. Then a Newton step is taken as described in Step 3. Alternatively, a gradient method is used, in which case Step 2 can be deleted. Finally, Step 4, i.e., determining f from q, amounts to solving a matrix Riccati equation and a Lyapunov equation, as seen from the following proposition.

Proposition 6.5. *Suppose that $q, \psi \in \mathcal{H}(B)$ are strictly positive real with Markov parameters (x, x_0) and (y, y_0), respectively. Let P be the unique solution to the algebraic Riccati equation*

$$P = APA' + (x - APc')(2x_0 - cPc')^{-1}(x - APc')', \tag{6.23}$$
$$d_1 := (2x_0 - cPc')^{1/2},$$
$$b_1 := (x - APc')d_1^{-1},$$

having the property that

$$\Gamma := A - b_1 d_1^{-1} c' \tag{6.24}$$

is stable, and let X be the unique solution of the Lyapunov equation

$$X = \Gamma X \Gamma' + y y_0^{-1} y' - (y - b_1 d_1^{-1} y_0) y_0^{-1} (y - b_1 d_1^{-1} y_0)', \tag{6.25}$$

$$d_2 := \tfrac{1}{2}(y_0 - cXc')d_1^{-1},$$

$$b_2 := [(y - Axc') - b_1 d_2]d_1^{-1}.$$

Then $f = \Im(q + q^*)$, defined as in (6.4), has the state-space representation

$$\mathbf{f} = \left[\begin{array}{c|c} \Gamma & b_1 d_1^{-1} d_2 - b_2 \\ \hline -d_1^{-1} c & d_1^{-1} d_2 \end{array} \right].$$

Proof. Observe that determining $a(z)$ from $q + q^* = aa^*$ is a standard spectral factorization problem [1, 14] with the unique minimum-phase solution given by

$$\mathbf{a} = \left[\begin{array}{c|c} A & b_1 \\ \hline c & d_1 \end{array} \right].$$

Then $b(z)$ is determined from the linear equation

$$\Psi = \psi + \psi^* = ab^* + ba^*$$

which, in the state-space formulation, becomes (6.25). Since Γ is stable, it has a unique solution X. Finally, the state space description of $f = a^{-1}b$ is obtained by direct computation. $\square$

Example 2: Maximal power transfer (continued). Consider a passive load with impedance

$$Z_\ell(s) = \frac{1 + RCs}{1 + (1 + R)Cs} + \frac{R_1 + L_1 s}{1 + R_1 + L_1 s},$$

where $R = 0.5\Omega$, $R_1 = 0.1\Omega$, $L_1 = 0.5H$, and $C = 0.01F$. This is a cascade connection of two (first order) filters, which are the parallel connections of a resistor $R = 1\Omega$ with a lossy capacitor and a lossy inductor respectively. The transmission zeros of $Z_\ell(s)$ are computed as the zeros of $Z_\ell(s) + Z_\ell(-s)$ to be ± 81.6429, ± 1.6249. The Blaschke factor

$$B(s) = \frac{(1 + R_1 - L_1 s)(1 - (1 + R)Cs)}{(1 + R_1 + L_1 s)(1 + (1 + R)Cs)}$$

evaluated at the transmission zeros provides the interpolation data

$$\rho(81.6429) = 0.0957, \quad \rho(1.6249) = 0.1432.$$

Translating the interpolation data to the z-domain we obtain $\hat{\rho}(-1.0248) = 0.0957$ and $\hat{\rho}(-4.2003) = 0.1432$. Thus, the interpolation conditions become $f(-1.0248) = 1.2116$ and $f(-4.2003) = 1.3342$. Suppose we want an effective power transmission characteristic, i.e., a power transmission gain Ψ close to one at low frequencies. Choice of spectral zeros in the neighborhood of 1 leads to low-pass gain transmission characteristic. Figure 5 shows the power transmission gain characteristics, i.e., $\Phi_{pg}(i\omega)$ vs. ω, for spectral zeros chosen at 0.6224, 0.9444 and 0.9987, respectively.
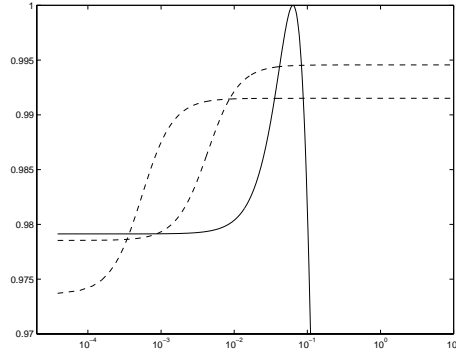
Figure 5: Power gain vs. frequency

Figure 6 shows the surface $\Phi_{pg}(i\omega)$ vs. $\log\omega$ vs. the choice of spectral zeros in the interval $[0.6224, 0.9987]$. Next, suppose that an additional lossy inductor is connected to the passive load with $L_2 = 0.2H$ and $R_2 = 0.5\Omega$. Applying the same analysis as before, $Z_\ell(s)$ is now of third order. A selection of two spectral zeros parametrizes the coupling network of dimension two. Selecting a double transmission zero at 0.9236, 0.9611, and 0.9932, respectively, leads to the lowpass characteristics shown in Figure 7 (dashed curves correspond to the first two choices while a continuous curve indicates the last one with a slighly wider bandwidth).
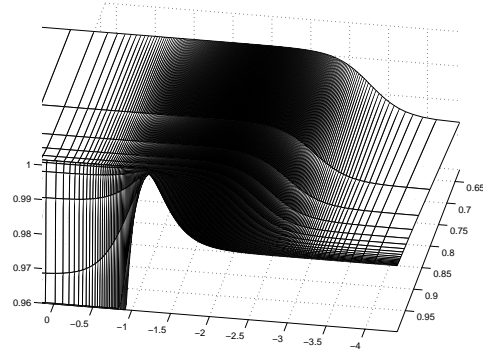

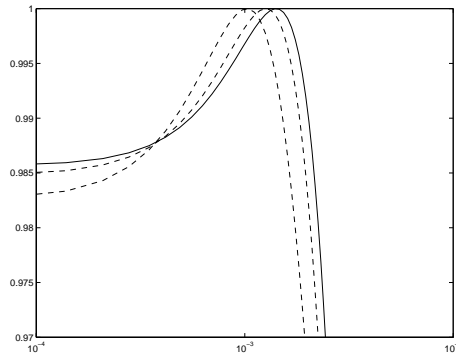Figure 6: Power gain vs. frequency vs. zero location

Figure 7: Power gain vs. frequency

At the present time, in high order cases, there is no systematic way to select transmission zeros that could produce the exact desired shape of the power transmission gain.

Example 3: Spectral estimation (continued). Consider a bank of three filters as in Figure 4, with $(z_0, z_1, z_2) = (\infty, 2, 1.5)$. Assume that the resulting values for $c_0(u_k)$, which specify $f(z)$ at these points, give interpolating values $(w_0, w_1, w_2) = (1, 1.2, 1.1)$. We would like to construct a model with an all-pole spectral density. Traditional techniques based on the Levinson algorithm are not applicable since the interpolation data are not in the form of a partial covariance sequence. Furthermore, the "central solution" corresponding to $\Psi(s) = 1$ leads to filters with spectral zeros at $z_1^{-1}, z_2^{-1}, \ldots, z_n^{-1}$, whereas we are interested in an AR model, i.e., one with all zeros at the origin. Selecting $\Psi(z) = \frac{1}{\tau(z)\tau(z^{-1})}$, where $\tau(z) = (z - 1/2)(z - 2/3)$, and using our algorithm, we obtain

$$f(z) = \frac{(z - 0.6829)(z + 0.8677)}{(z - 0.3612)^2 + 0.6797^2}.$$

Note that the zeros of $f(z)$ are at 0.6829 and at -0.8677, while there are no spectral zeros in the unit disc. The corresponding all-pole spectral density $f(z) + f(z^{-1})$ is depicted in Figure 8.
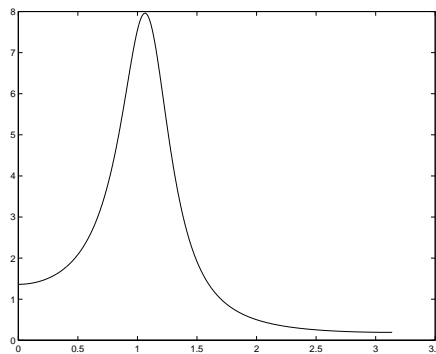

Figure 7: $|\Phi(e^{i\theta})|$ as a function of θ

A natural question regarding this example is why one would want to use Nevanlinna-Pick data for determinig an autoregressive model, when such a model can be obtained from traditional covariance data simply using the Levinson algorithm. The advantage in using Nevanlinna-Pick data is discussed in [3] where it is shown that a suitable selection of filterbank poles enhances resolution beyond what can be obtained with traditional covariance estimates. Intuitively, interpolation in the vicinity of an arc of the unit circle specifies more accurately the shape of f, and hence the spectral density, in that part of the spectrum.

7. Conclusions

In this paper, we have given a method for finding all solutions to the scalar, rational Nevanlinna-Pick interpolation problem, having degree less than or equal to n, in terms of the minima of a parameterized family of convex optimization problems.

While the problem has been posed for positive real interpolants, as would arise for the control of discrete-time systems, standard linear fractional transformations can adapt this generalized entropy criterion approach to positive real, or bounded-real, transfer functions for both continuous and discrete-time linear systems.

Appendix A. Proofs of deferred propositions and lemmas

Proof of Proposition 4.3. We note that $\mathcal{C}_+ \subset \mathcal{H}^2$, and we consider the representation

$$f(z) = \sum_{j=0}^{\infty} f_j g_j(z). \tag{A.1}$$

Based on our standing assumptions on $f(z)$, and our choice of the basis (3.1), (3.3), we have $f_0 = f(\infty)$ is real, while f_k, $k = 1, 2, \ldots$, are allowed to be complex. Thus, we identify $f(z)$ with the vector of coefficients $f := (f_0, f_1, \ldots)$, and define the set

$$\mathcal{F} = \{f \in \ell_2 \mid f_0 \in \mathbb{R}, f_1, f_2, \cdots \in \mathbb{C}, \sum_{j=0}^{\infty} f_j g_j(z) \in \mathcal{C}_+\}. \tag{A.2}$$

Since $B(z_k) = 0$ for $k = 0, 1, 2, \ldots, n$, we have $g_j(z_k) = 0$ for $j > n$, and consequently

$$f(z_k) = \sum_{j=0}^{n} f_j g_j(z_k). \tag{A.3}$$

Suppose that $\lambda \in \Lambda_+$. The function $f \to L(f, \lambda)$ is strictly concave, so, if it has a stationary point where the gradient is zero, it has a unique maximum there. Thus, we set $\frac{\partial L}{\partial f_k} = 0$ for all k. Since f_0 is real and $g_0 = 1$, we then have

$$\frac{\partial L}{\partial f_0} = 2\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi^{-1}(e^{i\theta})\Psi(e^{i\theta})d\theta - \lambda_0 - 2\mathrm{Re}\left\{\sum_{k=1}^{n} \bar{\lambda}_k\right\} = 0. \tag{A.4}$$

Furthermore, referring back to the discussion on function theory before Lemma 5.5, we recall that $\frac{\partial \bar{f}_k}{\partial f_k} = 0$ and $\frac{\partial \bar{f}_k}{\partial f_k} = 1$. Therefore, in view of (A.3), we obtain

$$\frac{\partial L}{\partial f_k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_k(e^{i\theta})\Phi^{-1}(e^{i\theta})\Psi(e^{i\theta})d\theta - \sum_{j=1}^{n} \bar{\lambda}_j g_k(z_j) = 0 \tag{A.5}$$

for $k = 1, 2, \ldots, n$, and

$$\frac{\partial L}{\partial f_k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_k(e^{i\theta})\Phi^{-1}(e^{i\theta})\Psi(e^{i\theta})d\theta = 0 \tag{A.6}$$

for $k = n+1, n+2, \ldots$, where we have used the orthogonality properties discussed in Section 3. Now, let $Q(z) := \Phi^{-1}(z)\Psi(z)$, and note that $Q^*(z) = Q(z)$. From (A.6),

$$\langle Q, g_k \rangle = 0 = \langle Q, g_k^* \rangle \quad \text{for } k = n+1, n+2, \ldots.$$

Hence $Q \in \mathcal{S}$, having a representation (4.10) with $q_0 \in \mathbb{R}$ and $q_1, \ldots, q_n \in \mathbb{C}$. By construction, (4.15) holds, and therefore it remains to show that $Q \in \mathcal{S}_+$ or, equivalently, that $q \in \mathcal{Q}_+$, to establish that $f \in \mathcal{C}_+$, proving the proposition.

From (A.4), we immediately see that

$$\lambda_0 = 2q_0 - 2\text{Re}\left\{\sum_{j=1}^{n}\lambda_j\right\}. \tag{A.7}$$

Next, taking the conjugate of (A.5) we obtain

$$\langle Q, g_k\rangle = \sum_{j=1}^{n}\lambda_j\bar{g}_k(\bar{z}_j) \tag{A.8}$$

for $k = 1, 2, \ldots, n$. On the other hand,

$$\langle Q, g_k\rangle = \sum_{j=1}^{n}q_j g_j(z_k). \tag{A.9}$$

Since $\bar{g}_k(\bar{z}_j) = g_j(z_k)$, by (A.8) and (A.9),

$$\begin{bmatrix} g_1(z_1) & g_2(z_1) & \cdots & g_n(z_1) \\ g_1(z_2) & g_2(z_2) & \cdots & g_n(z_2) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(z_n) & g_2(z_n) & \cdots & g_n(z_n) \end{bmatrix}\begin{bmatrix} \lambda_1 - q_1 \\ \lambda_2 - q_2 \\ \vdots \\ \lambda_n - q_n \end{bmatrix} = 0. \tag{A.10}$$

Now, since the the coefficient matrix of (A.10) is the matrix G of Lemma 6.2 and hence nonsingular,

$$\lambda_k = q_k \quad \text{for } k = 1, 2, \ldots, n. \tag{A.11}$$

Equations (A.7) and (A.11) establish that $\lambda = \boldsymbol{\lambda}(q)$. Therefore, since $\lambda \in \Lambda_+$, we have $q \in \mathcal{Q}_+$, as required. $\square$

Proof of Proposition 4.4. Applying the linear map (4.13), the dual functional (4.12) can be expressed in terms of $q := (q_0, q_1, \ldots, q_n)$. In fact,

$$\rho(\boldsymbol{\lambda}(q)) = -\frac{1}{2\pi}\int_{-\pi}^{\pi}\log[Q(e^{i\theta})]\Psi(e^{i\theta})d\theta + \frac{1}{2\pi}\int_{-\pi}^{\pi}\log[\Psi(e^{i\theta})]\Psi(e^{i\theta})d\theta$$
$$+ \left(2q_0 - 2\text{Re}\left\{\sum_{j=1}^{n}q_j\right\}\right)(w_0 - f_0) + 2\text{Re}\left\{\sum_{j=1}^{n}\bar{q}_j[w_j - f(z_j)]\right\}. \tag{A.12}$$

In this expression the sum of the two last terms turns out to be linear in q. To see this and eliminate the dependence of f's on the q's, consider the following:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Psi(e^{i\theta})d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}Q(e^{i\theta})\Phi(e^{i\theta})d\theta$$

$$= q_0\langle f + f^*, g_0\rangle + 2\text{Re}\left\{\sum_{j=1}^{n}\bar{q}_j\langle f + f^*, g_j\rangle\right\}$$

$$= 2q_0 f_0 + 2\text{Re}\left\{\sum_{j=1}^{n}\bar{q}_j(f(z_j) - f_0)\right\}.$$

Using this last expression, the dual function becomes

$$
\rho(\boldsymbol{\lambda}(q)) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \log[Q(e^{i\theta})]\Psi(e^{i\theta})d\theta + \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[\Psi(e^{i\theta})]\Psi(e^{i\theta})d\theta
$$

$$
- \frac{1}{2\pi} \int_{-\pi}^{\pi} \Psi(e^{i\theta})d\theta + 2q_0 w_0 + 2\mathrm{Re}\left\{ \sum_{j=1}^{n} \bar{q}_j(w_j - w_0) \right\}. \tag{A.13}
$$

In this expression, define c to be the sum of the second and third terms. Then, the proposition follows. $\square$

Proof of Proposition 5.1. We want to prove that $\mathbb{J}_\Psi(q)$ is finite when $q \neq 0$. Then the rest follows by inspection. Clearly, $\mathbb{J}_\Psi(q)$ cannot take the value $-\infty$; hence it remains to prove that $\mathbb{J}_\Psi(q) < \infty$. Since $q \neq 0$,

$$
\mu := \max_{\theta} Q(e^{i\theta}) > 0.
$$

Then, setting $P(z) := \mu^{-1}Q(z)$,

$$
\log P(e^{i\theta}) \leq 0 \tag{A.14}
$$

and

$$
\mathbb{J}_\Psi(q) = J(q) - \frac{1}{2\pi} \log \mu \int_{-\pi}^{\pi} \Psi(e^{i\theta})d\theta - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[P(e^{i\theta})]\Psi(e^{i\theta})d\theta,
$$

and hence the question of whether $\mathbb{J}_\Psi(q) < \infty$ is reduced to determining whether

$$
- \int_{-\pi}^{\pi} \log[P(e^{i\theta})]\Psi(e^{i\theta})d\theta < \infty.
$$

But, since $\Psi(e^{i\theta}) \leq M$ for some bound M, this follows from

$$
\int_{-\pi}^{\pi} \log P(e^{i\theta})d\theta > -\infty, \tag{A.15}
$$

which is the well-known Szegö condition: (A.15) is a necessary and sufficient condition for $P(e^{i\theta})$ to have a stable spectral factor [22]. But, since the rational function $P(z)$ belongs to $\mathcal{S}_+$, there is a function $\pi(z) \in \mathcal{H}(B)$ such that $\pi(z)\pi^*(z) = P(z)$. But then $\pi(z)$ is a stable spectral factor of $P(z)$, and hence (A.15) holds. $\square$

Proof of Proposition 5.2. Suppose $q^{(k)}$ is a sequence in $M_r := \mathbb{J}_\Psi^{-1}(-\infty, r]$. It suffices to show that $q^{(k)}$ has a convergent subsequence. The sequence $q^{(k)}$ defines a sequence of unordered n-tuples of zeros lying in the the unit disc, and a sequence of scalar multipliers. We wish to prove that both of these sequences cluster. To this end, each $Q^{(k)}$ may be factored as

$$
Q^{(k)}(z) = \lambda_k a_k(z)a_k^*(z) = \lambda_k \tilde{Q}^{(k)}(z),
$$

where λ_k is positive and $a_k(z)$ is a function in $\mathcal{H}(B)$ which is normalized so that $a_k(\infty) = 1$.

We shall first show that the sequence of zeros clusters. The corresponding sequence of the (unordered) set of n zeros of each $a_k(z)$ has a convergent subsequence, since all (unordered) sets of zeros lie in the closed unit disc. Denote by $a(z)$ the function in $\mathcal{H}(B)$ which vanishes at this limit set of zeros and which is normalized so that $a(\infty) = 1$. By reordering the sequence if necessary, we may assume the sequence

$a_k(z)$ tends to $a(z)$. Therefore the sequence $q^{(k)}$ has a convergent subsequence if and only if the sequence λ_k does.

We now show that the sequence of multipliers, λ_k, clusters. It suffices to prove that the sequence λ_k is bounded from above and from below away from zero. This will follow by analyzing the linear and the logarithmic growth in

$$\mathbb{J}_\Psi(q^{(k)}) = \lambda_k J(\tilde{q}^{(k)}) - \frac{1}{2\pi} \log \lambda_k \int_{-\pi}^{\pi} \Psi(e^{i\theta})d\theta - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[\tilde{Q}^{(k)}(e^{i\theta})]\Psi(e^{i\theta})d\theta$$

with respect to the sequence λ_k. Here $J(q)$ is the linear term (5.1) of $\mathbb{J}_\Psi(q)$. We first note that the sequence $J(\tilde{q}^{(k)})$, where $\tilde{q}^{(k)}$ is the vector corresponding to the pseudo-polynomial $\tilde{Q}^{(k)}$, is bounded from above because the normalized functions $a_k(z)$ lie in a bounded set. Similarly, by the proof of Lemma 5.3, the sequence $J(\tilde{q}^{(k)})$ is bounded from below, away from zero. In particular, the coefficient of λ_k in the first term for this expression for $\mathbb{J}_\Psi(q^{(k)})$ is bounded away from 0 and away from ∞. We also note that the coefficient of $\log \lambda_k$ in this expression for $\mathbb{J}_\Psi(q^{(k)})$ is independent of k. Next, the term

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log[\tilde{Q}^{(k)}(e^{i\theta})]\Psi(e^{i\theta})d\theta \tag{A.16}$$

in this expression for $\mathbb{J}_\Psi(q^{(k)})$ is independent of λ_k, and we claim that it remains bounded as a function of k. Indeed,

$$\tilde{Q}^{(k)}(e^{i\theta}) \to |a(e^{i\theta})|^2 = Q(z)$$

where $a(z)$ has all its zeros in the closed unit disc. In particular, if q in $\mathfrak{Q}$ corresponds to q, then the third term in the expression for $\mathbb{J}_\Psi(q^{(k)})$ converges to $\mathbb{J}_\Psi(q)$, which is finite since a is not identically zero.

Finally, observe that if a subsequence of λ_k were to tend to zero, then $\mathbb{J}_\Psi(q^{(k)})$ would exceed r. Likewise, if a subsequence of λ_k were to tend to infinity, $\mathbb{J}_\Psi$ would exceed r, since linear growth dominates logarithmic growth. $\quad\square$

Proof of Proposition 5.4. Denoting by $D_p\mathbb{J}_\Psi(q)$ the directional derivative of $\mathbb{J}_\Psi$ at q in the direction p, it is easy to see that

$$D_p\mathbb{J}_\Psi(q) := \lim_{\epsilon \to 0} \frac{\mathbb{J}_\Psi(q + \epsilon p) - \mathbb{J}_\Psi(q)}{\epsilon}$$
$$= J(p) - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{P(e^{i\theta})}{Q(e^{i\theta})} \Psi(e^{i\theta})d\theta, \tag{A.17}$$

where $P(z)$ is the pseudo-polynomial

$$P(z) = \bar{p}_n g_n^*(z) + \ldots + \bar{p}_1 g_1^*(z) + p_0 g_0(z) + p_1 g_1(z) + \ldots + p_n g_n(z)$$

corresponding to the vector $p \in \mathbb{C}^{n+1}$. In fact,

$$\frac{\log(Q + \epsilon P) - \log Q}{\epsilon} = \frac{P}{Q} \log\left[(1 + \epsilon \frac{P}{Q})^{\frac{1}{\epsilon}\frac{Q}{P}}\right] \to \frac{P}{Q}$$

as $\epsilon \to +0$, and hence (A.17) follows by dominated convergence.

Now, let $q \in \mathcal{Q}_+$ and $\tilde{q} \in \partial\mathcal{Q}$ be arbitrary. Then the corresponding pseudo-polynomials Q and $\tilde{Q}$ have the properties

$$Q(e^{i\theta}) > 0 \quad \text{for all } \theta \in [-\pi, \pi]$$

and

$$\tilde{Q}(e^{i\theta}) \geq 0 \quad \text{for all } \theta \text{ and } \tilde{Q}(e^{i\theta_0}) = 0 \text{ for some } \theta_0.$$

Since $q_\lambda := \tilde{q} + \lambda(q - \tilde{q}) \in \mathcal{Q}_+$ for $\lambda \in (0, 1]$, we also have for $\lambda \in (0, 1]$ that

$$Q_\lambda(e^{i\theta}) := \tilde{Q}(e^{i\theta}) + \lambda[Q(e^{i\theta}) - \tilde{Q}(e^{i\theta})] > 0, \quad \text{for all } \theta \in [-\pi, \pi],$$

and we may form the directional derivative

$$D_{\tilde{q}-q}\mathbb{J}_\Psi(q_\lambda) = J(\tilde{q} - q) + \frac{1}{2\pi} \int_{-\pi}^{\pi} h_\lambda(\theta) d\theta, \tag{A.18}$$

where

$$h_\lambda(\theta) = -\frac{Q(e^{i\theta}) - \tilde{Q}(e^{i\theta})}{Q_\lambda(e^{i\theta})} \Psi(e^{i\theta}).$$

Now,

$$\frac{d}{d\lambda} h_\lambda(\theta) = \frac{[Q(e^{i\theta}) - \tilde{Q}(e^{i\theta})]^2}{Q_\lambda(e^{i\theta})^2} \Psi(e^{i\theta}) \geq 0,$$

and hence $h_\lambda(\theta)$ is a monotonely nondecreasing function of λ for all $\theta \in [-\pi, \pi]$. Consequently h_λ tends pointwise to h_0 as $\lambda \to 0$. Therefore,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} h_\lambda(\theta) d\theta \to +\infty \quad \text{as } \lambda \to 0. \tag{A.19}$$

In fact, if

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} h_\lambda(\theta) d\theta \to \alpha < \infty \quad \text{as } \lambda \to 0, \tag{A.20}$$

then $\{h_\lambda\}$ is a Cauchy sequence in $L^1(-\pi, \pi)$ and hence has a limit in $L^1(-\pi, \pi)$ which must equal h_0 a.e. But h_0, having a pole at $\theta_0 \in [-\pi, \pi]$, is not summable and hence, as claimed, (A.20) cannot hold.

Consequently, by virtue of (A.18),

$$D_{\tilde{q}-q}\mathbb{J}_\Psi(q_\lambda) \to +\infty \quad \text{as } \lambda \to 0 \tag{A.21}$$

for all $q \in \mathcal{Q}_+$ and $\tilde{q} \in \partial\mathcal{Q}$, and hence, in view of Lemma 26.2 in [34], $\mathbb{J}_\Psi$ is essentially smooth. Then it follows from Theorem 26.3 in [34] that the subdifferential of $\mathbb{J}_\Psi$ is empty on the boundary of $\mathcal{Q}$, and therefore $\mathbb{J}_\Psi$ cannot have a minimum there. $\square$

Proof of Proposition 5.6. The proof follows directly from (A.21) $\square$

Proof of Lemma 6.1. For $k, \ell = 0, 1, \ldots, n$ we have

$$\frac{\partial^2 \mathbb{J}_\Psi}{\partial \bar{q}_k \partial \bar{q}_\ell} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_k^*(e^{i\theta}) g_\ell^*(e^{i\theta}) \frac{\Psi(e^{i\theta})}{Q(e^{i\theta})^2} d\theta \tag{A.22}$$

$$= \langle (h + h^*) g_\ell^*, g_k \rangle. \tag{A.23}$$

For $\ell = 0$ this becomes $\langle h, g_k \rangle + \langle h^*, g_k \rangle$, which, in view of (3.2), becomes $h(z_k) - h(z_0)$ if $k > 0$ and $2h(z_0)$ if $k = 0$. For $k, \ell > 0$, we have $\langle h^* g_\ell^*, g_k \rangle = 0$ and therefore

$$\frac{\partial^2 \mathbb{J}_\Psi}{\partial \bar{q}_k \partial \bar{q}_\ell} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_k^*(e^{i\theta}) g_\ell^*(e^{i\theta}) h(e^{i\theta}) d\theta.$$

There are two cases. First, suppose $k \neq \ell$. Then a simple calculation yields

$$g_k^*(z) g_\ell^*(z) = \frac{z_k}{z_\ell - z_k} g_k^*(z) + \frac{z_\ell}{z_k - z_\ell} g_\ell^*(z),$$

and hence

$$\frac{\partial^2 \mathbb{J}_\Psi}{\partial \bar{q}_k \partial \bar{q}_\ell} = \frac{z_k}{z_\ell - z_k} \langle h, g_k \rangle + \frac{z_\ell}{z_k - z_\ell} \langle h, g_\ell \rangle,$$

which, by (3.2), yields those elements of the Hessian for which $k \neq \ell$ and $k, \ell > 0$. Secondly, suppose that $k = \ell$. Since

$$g_k^*(z) = \frac{z}{z_k - z} = -1 + \frac{z_k}{z_k - z},$$

we obtain

$$\frac{\partial^2 \mathbb{J}_\Psi}{\partial^2 \bar{q}_k} = -\langle h, g_k \rangle + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{z_k e^{i\theta}}{(z_k - e^{i\theta})^2} h(e^{i\theta}) d\theta. \tag{A.24}$$

To compute the second term in (A.24), differentiate $h(z)$, which is given, as above, by the Cauchy formula

$$h(z) - h(z_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta}}{z - e^{i\theta}} h(e^{i\theta}) d\theta.$$

Then

$$h'(z) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta}}{(z - e^{i\theta})^2} h(e^{i\theta}) d\theta,$$

which, together with (A.24) and (3.2), proves the remaining part of the lemma. $\qquad \square$

References

1. B. D. O. Anderson, *The inverse problem of stationary convariance generation*, J. Statistical Physics **1** (1969), 133–147.
2. 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 **49** (April, 2001), 677–693.
3. C. I. Byrnes, T. T. Georgiou, and A. Lindquist, *A new approach to Spectral Estimation: A tunable high-resolution spectral estimator*, IEEE Trans. Signal Processing **48** (nov., 2000), 3189–3205.
4. C. I. Byrnes and A. Lindquist, *On duality between filtering and control*, in *Systems and Control in the Twenty-First Century*, C. I. Byrnes, B. N. Datta, D. S. Gilliam and C. F. Martin, editors, Birkhäuser, 1997, pp. 101–136.
5. C. I. Byrnes and A. Lindquist, *On the partial stochastic realization problem*, IEEE Trans. Automatic Control AC-42 (1997), 1049–1069.
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, SIAM J. Control and Optimization, **37** (1999), 211–229.

8. C. I. Byrnes, H. J. Landau and A. Lindquist, *On the well-posedness of the rational covariance extension problem*, in *Current and Future Directions in Applied Mathematics*, M. Alber, and B. Hu, J. Rosenthal, editors, Birhäuser Boston, 1997, pp. 83–108.

9. 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.

10. W-K. Chen, *Theory and Design of Broadband Matching Networks*, Pergamon Press, 1976.

11. 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.

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

13. J. C. Doyle, B. A. Francis and A. R. Tannenbaum, *Feedback Control Theory*, Macmillan Publ. Co., New York, 1992.

14. P. Faurre, M. Clerget, and F. Germain, *Opérateurs Rationnels Positifs*, Dunod, 1979.

15. B. A. Francis, *A Course in H_∞ Control Theory*, Springer-Verlag, 1987.

16. F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959.

17. T. T. Georgiou, *Partial realization of covariance sequences*, CMST, Univ. Florida, Gainesville, 1983.

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

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

20. T. T. Georgiou, *The interpolation problem with a degree constraint*, IEEE Transactions on Automatic Control **44(3)**: 631-635, 1999.

21. B. K. Ghosh, *Transcendental and interpolation methods in simultaneous stabilization and simultaneous partial pole placement problems*, SIAM J. Control and Optim. **24** (1986), 1091–1109.

22. U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, Almqvist & Wiksell, Stockholm, 1956.

23. S. Haykin, *Nonlinear Methods of Spectral Analysis*, Springer-Verlag, 1983.

24. H. Helson, *Lectures on Invariant Subspaces*, Academic Press, New York, 1964.

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

26. J. W. Helton, *The distance from a function to H^∞ in the Poincaré metric: Electrical power transfer*, J. of Functional Analysis, **38** (1980), 273-314.

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

28. S. A. Kassam and H. V. Poor, *Robust techniques for signal processing*, Proceedings IEEE **73** (1985), 433–481.

29. P. P. Khargonekar and A. Tannenbaum, *Non-Euclidean metrics and robust stabilization of systems with parameter uncertainty*, IEEE Trans. Automatic Control **AC−30** (1985), 1005–1013.

30. H. Kimura, *Robust stabilizability for a class of transfer functions* IEEE Trans. Automatic Control **AC−29** (1984), 788–793.

31. D. G. Luenberger, *Linear and Nonlinear Programming* (Second Edition), Addison-Wesley Publishing Company, Reading, Mass., 1984.

32. M. Minoux, Jr., *Mathematical Programming: Theory and Algorithms*, Wiley, 1986.

33. D. Mustafa and K. Glover, *Minimum Entropy H^∞ Control*, Springer Verlag, Berlin, 1990.

34. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

35. D. Sarason, *Generalized interpolation in H^∞*, Trans. Amer. Math. Soc. **127** (1967), 179–203.

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

37. A. Tannenbaum, *Modified Nevanlinna-Pick interpolation of linear plants with uncertainty in the gain factor*, Int. J. Control **36** (1982), 331–336.

38. J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, Amer. Math.Soc. Colloquium Publications, **20**, Providence, R. I., 1956.

39. D. C. Youla and M. Saito, *Interpolation with positive-real functions*, J. Franklin Institute **284** (1967), 77–108.

40. G. Zames and B. A. Francis, *Feedback, minimax sensitvity, and optimal robustness*, IEEE Transactions on Automatic Control **AC–28** (1983), 585–601.
41. K. Zhou, *Essentials of Robust Control*, Prentice-Hall, 1998.