A Convex Optimization Approach to ARMA Modeling

Tryphon T. Georgiou, *Fellow, IEEE*, and Anders Lindquist, *Fellow, IEEE*

***Abstract*—We formulate a convex optimization problem for approximating any given spectral density with a rational one having a prescribed number of poles and zeros (n poles and m zeros inside the unit disc and their conjugates). The approximation utilizes the Kullback–Leibler divergence as a distance measure. The stationarity condition for optimality requires that the approximant matches $n+1$ covariance moments of the given power spectrum and m cepstral moments of the corresponding logarithm, although the latter with possible slack. The solution coincides with one derived by Byrnes, Enqvist, and Lindquist who addressed directly the question of covariance and cepstral matching. Thus, the present paper provides an approximation theoretic justification of such a problem. Since the approximation requires only moments of spectral densities and of their logarithms, it can also be used for system identification.**

***Index Terms*—ARMA modeling, cepstral coefficients, convex optimization, covariance matching.**

I. Introduction

FINITE-DIMENSIONAL models are central to most modern-day techniques for analysis and design of control systems. Yet many problems concerning modeling of linear systems remain open. A case in point is ARMA (autoregressive moving average) modeling of time-series. In ARMA modeling, least-squares techniques often lead to inadmissible (unstable) models while most other approaches require non-convex optimization and are based on iterative procedures whose global convergence is not guaranteed. Hence, at least from a theoretical point of view, the problem of ARMA modeling remains open [21, p. 103].

In the present paper, we formulate a convex optimization for identifying a finite-dimensional approximant of a given power spectrum. The formulation involves a frequency domain representation and data which represent certain statistical moments of the time series. In fact, for constructing an ARMA(n,m) model, the data consist of n autocorrelation lags and m moments of the logarithm of the power spectral density (cepstral coefficients). The approximation uses the Kullback–Leibler divergence as a criterion of fit.

It is well known that model approximation and system identification are closely related subjects. In either, the objective is to match to high degree of accuracy given data by a lower order model. The difference between the two subjects is that in model approximation the data is thought to be exact, whereas in system identification their statistical nature is of paramount importance. It is for this reason that some system identification techniques rely on estimated statistics as the preferred form in which data is to be supplied. Since the optimality conditions for our approximation problem are only in terms of moments, our results are relevant to system identification as well.

The present paper provides an approximation-theoretic justification of a procedure proposed in [3] and [4] that directly addressed the question of covariance and cepstral matching. This previous work has been, in turn, part of a broad program on convex optimization for interpolation and moment problems initiated in [8] (also see [9]) and continued in [5]–[7]. In a quite different but yet related direction, in [15] we considered the following problem. Find a spectral density that best approximates an *a priori* spectral estimate in the Kullback–Leibler sense and, at the same time, matches a window of prescribed covariance lags. Though this earlier work uses related concepts and methods, it does not lead to a model reduction procedure since the solution in [15] is in general more complex than the a priori estimate.

In Section II, we formulate and motivate the basic approximation problem. The optimality conditions for the approximant are given in Section III. Some of the technical arguments are deferred to Section V. Then, in Section IV, we consider a regularized version of the approximation problem which is more amenable to numerical computation. Connections to the problem in [3] and [4] will be elaborated upon in Section VIII. In a final section, Section VII, we illustrate the approximation results with several examples.

II. Formulation of the Problem

We begin by considering the Kullback–Leibler divergence [18], [10]

$$D(\hat{\Phi}\,\|\,\Phi) := \int_{-\pi}^{\pi} \hat{\Phi}(e^{i\theta}) \log \frac{\hat{\Phi}(e^{i\theta})}{\Phi(e^{i\theta})} \frac{d\theta}{2\pi}$$

as a distance measure between power spectral densities $\hat{\Phi}$ and Φ. The functional $(\hat{\Phi}, \Phi) \mapsto D(\hat{\Phi}\,\|\,\Phi)$ is convex [10] and, as-

Manuscript received November 17, 2006; revised June 29, 2007 and August 29, 2007. Published August 27, 2008 (projected). Recommended by Associate Editor J.-F. Zhang. This research was supported by grants from AFOSR, NSF, VR, SSF, and the Göran Gustafsson Foundation.

T. T. Georgiou is with the Department of Electrical & Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: tryphon@umn.edu).

A. Lindquist is with the Department of Mathematics, Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden (e-mail: alq@math.kth.se).

Color versions of one or more figures are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2008.923684

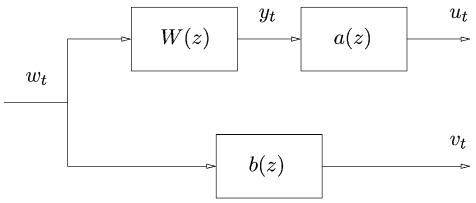


Fig. 1. Model matching problem.

suming that both arguments are normalized in the sense that

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

$D(\hat{\Phi} \,\|\, \Phi) \geq 0$ with equality if and only if $\hat{\Phi} = \Phi$. Although this is not a *bona fide* metric, it induces a metric topology and, in fact, a Riemannian metric structure on normalized positive functions; i.e., on probability densities [1], [2] (also see [14]). Therefore, it appears natural to consider the problem of approximating a given Φ by a $\hat{\Phi}$ belonging to a suitable class of admissible spectral densities.

From a systems-theoretic point of view it is natural to consider approximating Φ by a rational spectral density

$$\hat{\Phi} = \frac{P}{Q},$$

with P, Q trigonometric polynomials of at most degrees m and n, respectively. In the special case of MA approximation, when $n = 0$, this approximation problem reduces to minimizing $D(P \,\|\, \Phi)$ subject to normalization. This is a convex optimization problem in m coefficients of the trigonometric polynomial P. By way of contrast, while in general $\hat{\Phi} \mapsto D(\hat{\Phi} \,\|\, \Phi)$ is a convex functional in infinitely many variables, the functional $(P, Q) \mapsto D((P/Q) \,\|\, \Phi)$ has finitely many variables but is not convex. Therefore, to emulate the MA case, instead of minimizing $D((P/Q) \,\|\, \Phi)$, we consider the problem to minimize

$$\mathbb{S}(P, Q) := D(P \,\|\, Q\Phi),$$

subject to normalization of P and $Q\Phi$ as in (1). This is a convex optimization problem in $m + n + 1$ variables.

This approximation problem can be regarded as a model matching problem. We illustrate this in Fig. 1, where W is the outer spectral of Φ; i.e.

$$\Phi(e^{i\theta}) = |W(e^{i\theta})|^2$$

with W analytic and invertible in the open unit disc $\mathbb{D} := \{z : |z| < 1\}$

$$\begin{aligned} a(z) &= a_0 + a_1 z^{-1} + \cdots + a_n z^{-n} \\ b(z) &= b_0 + b_1 z^{-1} + \cdots + b_m z^{-m} \end{aligned}$$

are transfer functions of two finite-impulse-response filters, and $\{y_t : t \in \mathbb{Z}\}$ is a stationary random process produced by passing a white noise input $\{w_t\}$ through a filter with transfer function W. The approximation amounts to matching the power spectra of the two outputs $\{u_t\}$ and $\{v_t\}$.

To see this, first note that if $a(z)W(z) = b(z)$, then the power spectra of the outputs u_t and v_t will be the same and

$$\left| \frac{b(e^{i\theta})}{a(e^{i\theta})} \right|^2 = \Phi(e^{i\theta}). \tag{2}$$

More generally, we seek to determine the coefficients of $a(z)$ and $b(z)$ so as to minimize the Kullback–Leibler divergence

$$\int_{-\pi}^{\pi} |b(e^{i\theta})|^2 \log \left(\frac{|b(e^{i\theta})|^2}{\Phi(e^{i\theta}) |a(e^{i\theta})|^2} \right) \frac{d\theta}{2\pi}$$

between the power spectral densities $|b(e^{i\theta})|^2$ and $\Phi(e^{i\theta})|a(e^{i\theta})|^2$ of u_t and v_t, respectively, with the coefficients of $b(z)$ and $a(z)$ being normalized so that the variances $\int_{-\pi}^{\pi} |b(e^{i\theta})|^2 d\theta/2\pi$ and $\int_{-\pi}^{\pi} \Phi(e^{i\theta}) |a(e^{i\theta})|^2 d\theta/2\pi$ are both equal to one. With this normalization, the Kullback–Leibler divergence is nonnegative, and it is zero if and only if (2) holds, as indicated earlier.

The above formalism has already been considered in the context of least-squares estimation [19] (see also [20]). In this earlier work, a model matching problem is also motivated based on the configuration in Fig. 1. The matching criterion is to minimize the error variance $\mathrm{E}\{|u_t - v_t|^2\}$, instead of a distance between the power spectra of u_t and v_t, as done in the present work. Interestingly, the authors of [19] discovered that this least-squares model matching yields a stable filter. However, their problem requires as data the Markov parameters of the filter in addition to spectral moments. An interesting wrinkle to this connection is that for minimum-phase filters there is a bijective relation between Markov parameters and cepstral coefficients [4] (i.e., moments of the logarithm of the spectral density). It is this latter set of coefficients that enter in the present work.

As we shall see below, the problem of minimizing the distance between the spectral densities $|b(e^{i\theta})|^2$ and $\Phi(e^{i\theta})|a(e^{i\theta})|^2$ becomes a convex problem when expressed in terms of the coefficients of the pseudo-polynomials (trigonometric polynomials)

$$\begin{aligned} Q(z) &= a(z)a(z^{-1}) \\ &= q_0 + \frac{1}{2} \sum_{k=1}^{n} q_k (z^k + z^{-k}), \end{aligned} \tag{3a}$$

and

$$\begin{aligned} P(z) &= b(z)b(z^{-1}) \\ &= p_0 + \frac{1}{2} \sum_{k=1}^{m} p_k (z^k + z^{-k}). \end{aligned} \tag{3b}$$

Also, as we shall see in the next section, the optimal rational approximant turns out to be specified by a number of moment constraints of the given power spectrum and of its logarithm, which need to be matched (exactly or approximately in a suitable sense). Indeed, the optimality conditions are expressed in terms of the covariance lags

$$r_k := \int \cos k\theta \, \Phi(e^{i\theta}) \, d\mu, \quad k = 0, 1, \ldots, n \tag{4a}$$

and the cepstral coefficients

$$c_k := \int \cos k\theta \log \Phi(e^{i\theta})\, d\mu, \quad k = 0, 1, \dots, m \tag{4b}$$

where from this point on, for notational convenience, we use $d\mu$ to denote the normalized Lebesgue measure $d\theta/2\pi$, and we suppress the limits of integration, which will always be from $-\pi$ to π. Interestingly, the optimality conditions (4) coincide with those for a different optimization problem stated in [3], [4], as we have already indicated, and will further elaborate upon in Section VIII.

III. Optimality Conditions

Let $\mathcal{P}$ and $\mathcal{Q}$ be the closed convex sets

$$\mathcal{P} := \left\{ P \,|\, P(e^{i\theta}) = 1 + \sum_{k=1}^{m} p_k \cos k\theta \geq 0, \forall \theta \right\} \tag{5a}$$

and

$$\mathcal{Q} := \left\{ Q \,|\, Q(e^{i\theta}) = q_0 + \sum_{k=1}^{n} q_k \cos k\theta \geq 0, \forall \theta \right\} \tag{5b}$$

of nonnegative trigonometric polynomials. Then, the problem at hand is to minimize the convex functional

$$\mathbb{S}(P, Q) = \int P \log \left(\frac{P}{\Phi Q} \right) d\mu$$

over $\mathcal{P} \times \mathcal{Q}$ subject to the normalization condition

$$\int \Phi Q d\mu = 1. \tag{6a}$$

The normalization

$$p_0 = \int P d\mu = 1 \tag{6b}$$

also prescribed in Section II is already included in the definition of $\mathcal{P}$.

We denote by $\mathcal{P}_+$ and $\mathcal{Q}_+$ the subsets of $\mathcal{P}$ and $\mathcal{Q}$, respectively, for which $P(e^{i\theta}) > 0$ and $Q(e^{i\theta}) > 0$ for all θ. We also denote by $\partial\mathcal{P} := \mathcal{P} \setminus \mathcal{P}_+$ and $\partial\mathcal{Q} := \mathcal{Q} \setminus \mathcal{P}_+$ the corresponding boundaries. For future reference, we also define the hyperplane

$$\mathcal{H} := \left\{ Q \,|\, Q(e^{i\theta}) = q_0 + \sum_{k=1}^{n} q_k \cos k\theta \right. \quad \left. \text{and} \quad \int Q\Phi \, d\mu = 1 \right\}. \tag{7}$$

Moreover, we shall say that the pair $(P, Q) \in \mathcal{P} \times \mathcal{Q}$ has *maximal degree* if either $\deg P = m$ or $\deg Q = n$ or both.

Theorem 1: Consider the functional

$$\mathbb{S}(P, Q) : \mathcal{P} \times \mathcal{Q} \mapsto \mathbb{R}$$

defined on nonnegative trigonometric polynomials P, Q of degrees m and n, respectively, as specified in (5), and the set of minimizing solutions

$$\mathcal{S} := \{(\hat{P}, \hat{Q}) \,|\, \mathbb{S}(\hat{P}, \hat{Q}) = \min \mathbb{S}(P, Q) \text{ subject to (6a)}\}.$$

Then the following hold.

i) The set $\mathcal{S}$ is nonempty.
ii) The ratio P/Q defines the same function for all $(P, Q) \in \mathcal{S}$.
iii) If a pair $(P, Q) \in \mathcal{S}$ is relatively prime and has maximal degree, then $\mathcal{S}$ contains only this one element.
iv) If a pair $(P, Q) \in \mathcal{S}$ does not have maximal degree and $P \in \mathcal{P}_+$, then $\mathcal{S}$ contains more than one element.
v) If a pair $(P, Q) \in \mathcal{S}$ is not coprime and $P \in \mathcal{P}_+$, then $\mathcal{S}$ contains more than one element.

The following observation is needed throughout.

Lemma 2: The sets $\mathcal{P}$ and $\mathcal{Q} \cap \mathcal{H}$ are compact.

Proof: Since the elements of $\mathcal{P}$ are nonnegative and their constant term is one, their remaining coefficients are also bounded by one. Therefore, $\mathcal{P}$ is compact. Next we show that $\mathcal{Q} \cap \mathcal{H}$ is also compact. Since Φ is the power spectral density of a purely nondeterministic process, the Toeplitz matrix formed out of the first $n+1$ moments $r_0, r_1, \dots, r_n$ is positive definite, and the corresponding nth Szegö polynomial $\varphi_n(z)$ (see, e.g., [16]) is devoid of roots on the circle. Moreover, since

$$\int e^{ik\theta} \frac{1}{|\varphi_n(e^{i\theta})|^2} d\mu = r_k, \quad k = 0, 1, \dots, n$$

and Q is a trigonometric polynomial of degree n, it follows that

$$\begin{aligned} 1 &= \int Q(e^{i\theta}) \Phi(e^{i\theta}) d\mu \\ &= \int Q(e^{i\theta}) \frac{1}{|\varphi_n(e^{i\theta})|^2} d\mu \geq q_0 \frac{1}{M} \end{aligned}$$

where $M = \max\{|\varphi_n(e^{i\theta})|^2 : \theta \in [0, \pi]\} < \infty$. Hence, q_0 is bounded by M and so are the remaining coefficients of Q. Consequently, $\mathcal{Q} \cap \mathcal{H}$ is compact as claimed. ■

We are now in a position to prove statements i)–iii) of Theorem 1. Proofs of the remaining statements are deferred to Section V.

Proof: [Proof of Theorem 1, statements i)–iii)] Since $\mathbb{S}$ is convex and $\mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$ is compact (Lemma 1), there exists a minimizing point. This establishes statement i).

Since $\mathbb{S}$ is a convex functional, the set $\mathcal{S}$ of minimizers is convex. If (P_0, Q_0) and (P_1, Q_1) are both in $\mathcal{S}$, then so is the whole interval $(P_\tau, Q_\tau) \in \mathcal{S}$ for $\tau \in [0, 1]$. Therefore, the second derivative of $\mathbb{S}(P_\tau, Q_\tau)$ with respect to τ, at any point of the interval, must be zero. In particular, for $\tau = 0$ and for $\delta P = P_1 - P_0$ and $\delta Q = Q_1 - Q_0$, we compute that the second derivative is

$$\int \frac{(Q_0 \delta P - P_0 \delta Q)^2}{P_0 Q_0^2} d\mu.$$

If this integral vanishes, then, since the integrand is nonnegative, $P_0 \delta Q - Q_0 \delta P = 0$. Hence, $P_0/Q_0 = P_1/Q_1$ for any two elements (P_0, Q_0) and (P_1, Q_1) in $\mathcal{S}$. This proves statement (ii).

If a pair $(P, Q) \in \mathcal{S}$ is coprime and if at least one of them has maximal degree, i.e., at least one of $\deg P = m$ or $\deg Q = n$ holds, then any other element $(P_1, Q_1) \in \mathcal{S}$ must satisfy $P/Q = P_1/Q_1$. If P_1 and Q_1 have a nontrivial common factor, then one of the two must exceed the corresponding allowable degree and contradicts the assumption that $(P_1, Q_1) \in \mathcal{P} \times \mathcal{Q}$. This completes the proof of statement (iii). ■

The next theorem lists the appropriate optimality conditions under the same assumptions and notation as in Theorem 1. The proof will be deferred to Section V.

Theorem 3: Suppose a window of covariance lags $(r_0, r_1, \dots, r_n)$ and a window of cepstral coefficients $(c_0, c_1, \dots, c_m)$ are computed from Φ as in (4). Then, any $(P, Q) \in \mathcal{S}$ satisfies the covariance matching conditions

$$\int \cos k\theta \frac{P}{Q} d\mu = r_k, \quad k = 0, 1, \dots, n. \tag{8}$$

If in addition $P \in \mathcal{P}_+$, then (P, Q) satisfies the cepstral matching conditions

$$\int \cos k\theta \log \frac{P}{Q} d\mu = c_k, \quad k = 1, \dots, m. \tag{9}$$

Conversely, any $(P, Q) \in \mathcal{P} \times \mathcal{Q}$ which satisfies both sets of the moment conditions (8) and (9) belongs to $\mathcal{S}$. In general, allowing for the case where $P \in \partial\mathcal{P}$, any $(P, Q) \in \mathcal{S}$ satisfies the modified cepstral matching condition

$$\int \cos k\theta \log \frac{P}{Q} d\mu = c_k + g_k, \quad k = 1, \dots, m \tag{10}$$

where the slack variables $g_1, g_2, \dots, g_m$ and

$$g_0 := -\sum_{k=1}^{m} p_k g_k \tag{11}$$

take the same values for all $(P, Q) \in \mathcal{S}$. Moreover, $g_0 \geq 0$, and, if for some $(\hat{P}, \hat{Q}) \in \mathcal{S}$, $\hat{P} \in \mathcal{P}_+$, then $g_k = 0$ for $k = 0, 1, \dots, m$. On the other hand, if $\hat{P} \in \partial\mathcal{P}_+$ with zeros at $e^{i\theta_j}$, $j = 1, 2, \dots, \ell$, then there are nonnegatve numbers $\alpha_1, \alpha_2, \dots, \alpha_\ell$ such that

$$g_0 = 2\sum_{j=1}^{\ell} \alpha_j \tag{12a}$$

$$g_k = 2\sum_{j=1}^{\ell} \alpha_j \cos k\theta_j, \quad k = 1, 2, \dots, m. \tag{12b}$$

Finally, for any $(P, Q) \in \mathcal{S}$, the optimal value is

$$\begin{aligned} \mathbb{S}(P, Q) &= \int \log \frac{P}{Q} d\mu - \int \log \Phi \, d\mu - g_0 \\ &= \int \log \frac{P}{Q} d\mu - c_0 - g_0. \end{aligned} \tag{13}$$

In the last statement of the theorem we see that the minimal Kullback–Leibler divergence equals the difference in entropy gain between P/Q and Φ or, equivalently, the difference between their respective zeroth cepstral coefficients, both modified by the slack variable $g_0 \geq 0$ whenever $P \in \partial\mathcal{P}$. Moreover, we see from (8), that Q may not have a root on the unit circle unless P has the same root so that cancellation ensures integrability of the fraction P/Q.

IV. Regularization and Cepstral Slackness

The computation of the minimizer for $\mathbb{S}$ is considerably more complicated when $\hat{P}$ has roots on the unit circle; i.e., when $\hat{P}$ belongs to the boundary of $\mathcal{P}$. There is therefore a need for a suitable regularization that will keep the solution in the interior of $\mathcal{P} \times \mathcal{Q}$. Such strategies have been considered for the problem in [3] and [4] by Per Enqvist in his Ph.D. dissertation [11], [12]. Here we consider the same regularization for our Kullback–Leibler functional

$$\mathbb{S}_\varepsilon(P, Q) := \mathbb{S}(P, Q) - \int \varepsilon \log P(e^{i\theta}) d\mu. \tag{14}$$

By Szegö's theorem [16] $\log P$ is integrable on the unit circle, but its derivative is not. This forces the minimizing solution to lie in the interior of $\mathcal{P} \times \mathcal{Q}$.

Theorem 4: Suppose a window of covariance lags $(r_0, r_1, \dots, r_n)$ and a window of cepstral coefficients $(c_0, c_1, \dots, c_m)$ are computed from Φ as in (4a) and (4b), and let $\varepsilon > 0$. The problem of minimizing the functional

$$\mathbb{S}_\varepsilon(P, Q) : \mathcal{P} \times \mathcal{Q} \mapsto \mathbb{R}$$

subject to (6a) has a unique solution $(\hat{P}, \hat{Q})$. This solution belongs to $\mathcal{P}_+ \times \mathcal{Q}_+$ and satisfies the covariance matching conditions

$$\int \cos k\theta \frac{\hat{P}}{\hat{Q}} d\mu = r_k, \quad k = 0, 1, \dots, n$$

as well as the modified cepstral matching conditions

$$\int \cos k\theta \, \log \frac{\hat{P}}{\hat{Q}} d\mu = c_k + \varepsilon \int \cos k\theta \frac{1}{\hat{P}} d\mu, \quad k = 1, \dots, m. \tag{15}$$

Moreover, the optimal value is

$$\mathbb{S}_\varepsilon(\hat{P}, \hat{Q}) = \int \log \frac{\hat{P}}{\hat{Q}} d\mu - c_0 + \varepsilon \int \left(1 - \frac{1}{\hat{P}} - \log \hat{P}\right) d\mu. \tag{16}$$

Proof: We compute the first variation of $\mathbb{S}_\varepsilon$

$$\begin{aligned} \delta\mathbb{S}_\varepsilon(P, Q; \delta P, \delta Q) = &\int \left(1 + \log \frac{P}{Q\Phi}\right) \delta P \, d\mu \\ &- \int \frac{P}{Q} \delta Q \, d\mu - \varepsilon \int \frac{1}{P} \delta P \, d\mu \end{aligned}$$

and the second variation

$$\delta^2\mathbb{S}_\varepsilon(P, Q; \delta P, \delta Q) = \int \frac{(Q\delta P - P\delta Q)^2}{PQ^2} d\mu + \varepsilon \int \frac{(\delta P)^2}{P^2} d\mu.$$

The second variation is positive unless $\delta P = 0$ and $\delta Q = 0$. Consequently, $\mathbb{S}_\varepsilon$ is strictly convex.

The optimization problem in the theorem is equivalent to finding a pair $(P,Q)\in\mathcal{P}\times(\mathcal{Q}\cap\mathcal{H})$, with $\mathcal{H}$ as in (7), that minimizes $\mathbb{S}_\varepsilon$. We have shown earlier that $\mathcal{P}\times(\mathcal{Q}\cap\mathcal{H})$ is compact. Since in addition $\mathbb{S}_\varepsilon$ is strictly convex the optimization problem has a unique solution $(\hat{P},\hat{Q})$.

The functional $\mathbb{S}_\varepsilon$ remains finite on $\mathcal{P}\times(\mathcal{Q}\cap\mathcal{H})$, yet its first variation becomes unbounded on the boundary. This implies, in fact, that $(\hat{P},\hat{Q})$ lies in the interior $\mathcal{P}_+\times(\mathcal{Q}_+\cap\mathcal{H})$. To see this, we provide the following brief argument. Assume that for some $e^{i\theta_0}$, with $\theta_0\in[-\pi,\pi]$, one or both of $\hat{P},\hat{Q}$ vanish. Since the set $\mathcal{P}\times(\mathcal{Q}\cap\mathcal{H})$ is obviously convex with a nontrivial interior, there exist trigonometric polynomials δP and δQ so that

$$(\hat{P}+\delta P,\hat{Q}+\delta Q)\in\mathcal{P}_+\times(\mathcal{Q}_+\cap\mathcal{H})$$

and, in particular, $\delta P(e^{i\theta_0})>0$ and $\delta Q(e^{i\theta_0})>0$. But then, perturbing in the direction $(\delta P,\delta Q)$

$$\delta\mathbb{S}_\varepsilon(\hat{P},\hat{Q};\delta P,\delta Q)=\int\left(1+\log\frac{\hat{P}}{\hat{Q}\Phi}\right)\delta P\,d\mu - \int\frac{\hat{P}}{\hat{Q}}\delta Q\,d\mu-\varepsilon\int\frac{1}{\hat{P}}\delta P\,d\mu$$

gives $-\infty$ since the first term is finite, whereas at least one of the other two terms gives $-\infty$. This argument can become more rigorous by considering a path (P_α,Q_α), with $\alpha\in[0,1]$, which starts at an interior point and ends at the assumed minimizing point $(\hat{P},\hat{Q})$ on the boundary. The contradiction is drawn by evaluating the derivative of $\mathbb{S}_\varepsilon(P_\alpha,Q_\alpha)$ at the end point which turns out to be $+\infty$. In a similar context, this is done in detail in [8, p. 225].

In order to establish the moment conditions satisfied at the minimizing point $(\hat{P},\hat{Q})$, we consider the Lagrangian

$$L(P,Q,\lambda):=\int\left(P\log\frac{P}{Q\Phi}-\varepsilon\log P\right)d\mu+\lambda\left(\int Q\Phi\,d\mu-1\right)\quad(17)$$

on all of $\mathcal{P}\times\mathcal{Q}$, where $\lambda\in\mathbb{R}$ is a Lagrange multiplier. For all fixed $\lambda>0$, the Lagrangian has compact sublevel sets

$$M_r:=\{(P,Q)\,|\,L(P,Q,\lambda)\le r\}$$

with $r\in\mathbb{R}$. To see this let $Q=\eta Q_0$ with $Q_0\in\mathcal{Q}\cap\mathcal{H}$. Clearly, $\eta>0$. Then, since

$$L(P,Q,\lambda)=\int\left(P\log\frac{P}{Q_0\Phi}-P\log\eta-\varepsilon\log P\right)d\mu+\lambda(\eta-1),$$

$(P,Q)\in M_r$ if and only if

$$\lambda\eta-\log\eta\le r+\lambda-\mathbb{S}_\varepsilon(P,Q_0).$$

Comparing linear and logarithmic growth, this shows that η is bounded from above (and also bounded away from 0). This shows that the sublevel sets are bounded. Since the Lagrangian is continuous they are all also closed, and hence compact.

Because the Lagrangian differs from $\mathbb{S}_\varepsilon(P,Q)$ by a linear term, it is strictly convex as well. Therefore, for each $\lambda>0$, the Lagrangian has also a unique minimizing point $(P_\lambda,Q_\lambda)\in\mathcal{P}\times\mathcal{Q}$. This minimizing point is not located on the boundary. To see this note that the first variation

$$\delta L(P,Q,\lambda;\delta P,\delta Q)=\delta\mathbb{S}_\varepsilon(P,Q;\delta P,\delta Q)+\lambda\int\Phi\delta Q d\mu.$$

A similar argument as the one given above for the case of $\mathbb{S}_\varepsilon$ applies to show that $(P_\lambda,Q_\lambda)\in\mathcal{P}_+\times\mathcal{Q}_+$. Therefore, (P_λ,Q_λ) is the unique solution of the stationarity conditions

$$\frac{\partial L}{\partial p_k}=\int\cos k\theta\left(1+\log\frac{P}{Q\Phi}-\frac{\varepsilon}{P}\right)d\mu=0,\quad k=1,2,\dots,m\quad(18a)$$

$$\frac{\partial L}{\partial q_k}=\int\cos k\theta\left(-\frac{P}{Q}+\lambda\Phi\right)d\mu=0,\quad k=0,1,\dots,n.\quad(18b)$$

The set of (18a) can be rewritten as follows:

$$\int\cos k\theta\log\frac{P}{Q}d\mu=c_k+\int\cos k\theta\frac{\varepsilon}{P}d\mu,\quad k=1,2,\dots,m.\quad(19)$$

Using (18b), we can determine the unique value for λ which is in harmony with the side condition (6a). To this end, we form

$$\begin{aligned}\sum_{k=0}^{n}q_k\int\cos k\theta\left(-\frac{P}{Q}+\lambda\Phi\right)d\mu&=-\int Pd\mu+\lambda\int Q\Phi\,d\mu\\&=-1+\lambda\int Q\Phi\,d\mu\end{aligned}$$

which equals zero due to (18b), and therefore

$$\lambda=1.\quad(20)$$

Conversely, for $\lambda=1$, the unique solution satisfying the stationarity conditions also satisfies (6a), and therefore

$$L(P_1,Q_1,1)=\mathbb{S}_\varepsilon(P_1,Q_1).\quad(21)$$

With the choice $\lambda=1$ (18b) becomes

$$\int\cos k\theta\frac{P}{Q}d\mu=r_k,\quad k=0,1,\dots,n.\quad(22)$$

From the optimization of the Lagrangian, we have

$$L(P,Q,1)\ge L(P_1,Q_1,1)$$

for all $(P,Q)\in\mathcal{P}\times\mathcal{Q}$. If in addition $Q\in\mathcal{H}$, then

$$\mathbb{S}_\varepsilon(P,Q)=L(P,Q,1)\ge L(P_1,Q_1,1)$$

and hence, from (21), we conclude that

$$\mathbb{S}_\varepsilon(P,Q)\ge\mathbb{S}_\varepsilon(P_1,Q_1)\quad(23)$$

for all $(P,Q)\in\mathcal{P}\times(\mathcal{Q}\cap\mathcal{H})$.

Finally, we establish (16). From (18a), we see that

$$\sum_{k=1}^{m} \hat{p}_k \int \cos k\theta \left(\log \frac{\hat{P}}{\hat{Q}\Phi} - \frac{\varepsilon}{\hat{P}} \right) d\mu = 0$$

which can be rewritten as

$$\int (\hat{P} - 1) \left(\log \frac{\hat{P}}{\hat{Q}\Phi} - \frac{\varepsilon}{\hat{P}} \right) d\mu = 0.$$

Therefore

$$\begin{aligned} \mathbb{S}_\varepsilon(\hat{P}, \hat{Q}) &= \int \hat{P} \log \frac{\hat{P}}{\hat{Q}\Phi} d\mu - \int \varepsilon \log \hat{P} d\mu \\ &= \int \log \frac{\hat{P}}{\hat{Q}\Phi} d\mu + \varepsilon \int \frac{\hat{P} - 1}{\hat{P}} d\mu \\ &\quad - \varepsilon \int \log \hat{P} d\mu \end{aligned}$$

from which (16) readily follows. ∎

V. Proofs of Theorems 1 and 3

In order to prove Theorem 3, we consider what happens to conditions of Theorem 4 when $\varepsilon \to 0$. We then use Theorem 3 to complete the proof of Theorem 1.

We first note that, by Theorem 4, for each ε there is a unique minimizer of $\mathbb{S}_\varepsilon$ in $\mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$, which we denote by $(P_\varepsilon, Q_\varepsilon)$. Since $\mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$ is compact (Lemma 2), there is a sequence in the set $\{(P_\varepsilon, Q_\varepsilon) \,|\, \epsilon > 0\}$ that converges to a limit point (P_0, Q_0). For any point in this sequence, we have that

$$\mathbb{S}_\varepsilon(P, Q) \geq \mathbb{S}_\varepsilon(P_\varepsilon, Q_\varepsilon)$$

holds for all (P, Q) in $\mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$. By continuity $\mathbb{S}_\varepsilon(P, Q) \to \mathbb{S}(P, Q)$ and $\mathbb{S}_\varepsilon(P_\varepsilon, Q_\varepsilon) \to \mathbb{S}(P_0, Q_0)$ as $\varepsilon \to 0$. Therefore

$$\mathbb{S}(P, Q) \geq \mathbb{S}(P_0, Q_0) \text{ for all } (P, Q) \in \mathcal{P} \times (\mathcal{Q} \cap \mathcal{H}),$$

and hence, any such limit point (P_0, Q_0), which need not be unique, is a minimizer of $\mathbb{S}$. By Theorem 3, any such limit point (P_0, Q_0) satisfies the covariance matching conditions (8); i.e.

$$\int \cos k\theta \frac{P_0}{Q_0} d\mu = \int \cos k\theta \, \Phi \, d\mu, \qquad k = 0, 1, \ldots, n. \tag{24a}$$

In view of Theorem 1, ii), any $(P, Q) \in \mathcal{S}$ also satisfies (8). The quantities

$$g_k := \int \cos k\theta \log \frac{P_0}{Q_0} d\mu - \int \cos k\theta \log \Phi \, d\mu, \qquad k = 1, 2, \ldots, m \tag{24b}$$

are well defined, since the logarithms of trigonometric polynomials are always integrable. These are the variables in (10) in Theorem 3. Then, again by Theorem 1, ii), any $(P, Q) \in \mathcal{S}$ satisfies (10). Next, choose an arbitrary $(\hat{P}, \hat{Q}) \in \mathcal{S}$ and suppose that $\hat{P} \in \mathcal{P}_+$. Since the function

$$F : P \mapsto \mathbb{S}(P, \hat{Q})$$

is convex and $\hat{P}$ is a minimizing point, the stationarity condition

$$\frac{\partial F}{\partial p_k} = \int \cos k\theta \left(\log \frac{P}{\hat{Q}} - \log \Phi \right) d\mu = 0$$

holds for $P = \hat{P}$ and $k = 1, \ldots, m$. Therefore (9) holds at the point $(\hat{P}, \hat{Q})$, and hence, whenever $\hat{P} \in \mathcal{P}_+$, $g_k = 0$, $k = 1, 2, \ldots, m$.

Conversely, suppose that $(\hat{P}, \hat{Q})$ satisfies (8) and (10) for some $\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_m$. Then, in view of (8)

$$\begin{aligned} \int \hat{Q}\Phi \, d\mu &= \sum_{k=1}^{n} \hat{q}_k \int \cos k\theta \frac{\hat{P}}{\hat{Q}} d\mu \\ &= \int \hat{P} d\mu = 1 \end{aligned} \tag{25}$$

which implies that $\hat{Q} \in \mathcal{H}$. Next we show that $(\hat{P}, \hat{Q}) \in \mathcal{S}$. To this end, define the Lagrangian

$$\begin{aligned} L(P, Q, \lambda, g_0) := &\int (P \log P - P \log Q - P \log \Phi) d\mu \\ &+ \lambda \left(\int Q\Phi \, d\mu - 1 \right) - \int GP d\mu \end{aligned} \tag{26}$$

where $\lambda \in \mathbb{R}$ is a Lagrange multiplier as is the constant term g_0 in the trigonometric polynomial

$$G(e^{i\theta}) = g_0 + 2 \sum_{k=1}^{m} \hat{g}_k \cos k\theta \tag{27a}$$

chosen so that

$$\langle G, P \rangle := \int GP d\mu \geq 0. \tag{27b}$$

A straight-forward inspection shows that, for $\lambda > 0$, $(\hat{P}_\lambda, \hat{Q}_\lambda) := (\hat{P}, \lambda^{-1}\hat{Q})$ satisfied the stationarity conditions

$$\begin{aligned} \frac{\partial L}{\partial p_k} &= \int \cos k\theta \left(\log \frac{P}{Q} - \log \Phi \right) d\mu - \hat{g}_k \\ &= 0, \quad k = 1, 2, \ldots, m \end{aligned} \tag{28a}$$

$$\begin{aligned} \frac{\partial L}{\partial q_k} &= \int \cos k\theta \left(-\frac{P}{Q} + \lambda\Phi \right) d\mu \\ &= 0, \quad k = 0, 1, \ldots, n \end{aligned} \tag{28b}$$

and hence $(\hat{P}_\lambda, \hat{Q}_\lambda)$ is a minimizer of the (convex) Lagrangian. Forming the dual functional

$$\begin{aligned} J(\lambda, g_0) &:= \min_{(P,Q) \in \mathcal{P} \times \mathcal{Q}} L(P, Q, \lambda, g_0) \\ &= L(\hat{P}, \lambda^{-1}\hat{Q}, \lambda, g_0) \end{aligned}$$

we have

$$J(\lambda, g_0) = \int \left(\log \frac{\hat{P}}{\hat{Q}} - \log \Phi \right) d\mu - g_0 + \log \lambda - \lambda + \sum_{k=1}^{m} \hat{p}_k \left[\int \cos k\theta \left(\log \frac{\hat{P}}{\hat{Q}} - \log \Phi \right) d\mu - \hat{g}_k \right] + \int \hat{Q}\Phi \, d\mu. \tag{29}$$

By the assumption that $(\hat{P}, \hat{Q})$ satisfies (10) with $g_k = \hat{g}_k$, $k = 1, 2, \ldots, m$, the quantity within square brackets in the second line of (29) equals zero. Therefore, in view of (25)

$$J(\lambda, g_0) = \log \lambda - \lambda - g_0 + \int \log \frac{\hat{P}}{\hat{Q}\Phi} \, d\mu + 1 \tag{30}$$

which should be maximized over all $\lambda > 0$ and g_0 satisfying (27). Clearly, there is a unique maximum at $(\hat{\lambda}, \hat{g}_0)$, where $\hat{\lambda} = 1$ and

$$\hat{g}_0 = \int (1 - \hat{P}) \log \frac{\hat{P}}{\hat{Q}\Phi} d\mu = \int \log \frac{\hat{P}}{\hat{Q}\Phi} d\mu - \mathbb{S}(\hat{P}, \hat{Q}) \tag{31}$$

is chosen so that

$$\hat{g}_0 + \sum_{k=1}^{m} \hat{p}_k \hat{g}_k = \int \hat{G}\hat{P} d\mu = 0. \tag{32}$$

By Theorem 1, ii), the first term in (31) takes the same value for all $(\hat{P}, \hat{Q}) \in \mathcal{S}$, and so of course does the second term. Hence $\hat{g}_0$ is independent of the particular $(\hat{P}, \hat{Q})$. This also establishes (13).

Moreover, from (30) and (31), we have $J(\hat{\lambda}, \hat{g}_0) = \mathbb{S}(\hat{P}, \hat{Q})$, as expected. Consequently, the Lagrangian (26) has a saddle point at $(\hat{P}, \hat{Q}, 1, \hat{g}_0)$. In particular,

$$L(P, Q, 1, \hat{g}_0) \geq J(1, \hat{g}_0) = \mathbb{S}(\hat{P}, \hat{Q})$$

which proves that $(\hat{P}, \hat{Q}) \in \mathcal{S}$, as claimed.

To prove (12b), we first introduce the trigonometric polynomials

$$\gamma_j(e^{i\theta}) := 1 + 2 \sum_{k=1}^{m} \cos(k\theta_j) \cos(k\theta) \tag{33}$$

$j = 1, \ldots, \ell$. These satisfy

$$\langle \gamma_j, P \rangle = P(e^{i\theta_j}), \quad j = 1, 2, \ldots, \ell \tag{34}$$

for any trigonometric polynomial P of degree m. We also denote by Γ_+ the nonnegative cone

$$\Gamma_+ := \left\{ \sum_{j=1}^{\ell} \alpha_j \gamma_j(e^{i\theta}) \text{ with } \alpha_j \geq 0, l = 1, \ldots, \ell \right\}.$$

Suppose now that $(\hat{P}, \hat{Q}) \in \mathcal{S}$ and that $\hat{P}(e^{i\theta_j}) = 0$ for $j = 1, 2, \ldots, \ell$. Then perturbing $\mathbb{S}$ in admissible directions $(\delta P, 0)$, we obtain

$$\delta \mathbb{S} = \int \delta P \left(1 + \log \frac{\hat{P}}{\hat{Q}\Phi} \right) d\mu \geq 0$$

which, in view of the fact that the δP is a trigonometric polynomial without constant term, is the same as

$$\langle \hat{G}, \delta P \rangle \geq 0. \tag{35}$$

Since any variation δP of $\hat{P} \in \mathcal{P}$ must preserve positivity, the admissible perturbations are precisely those for which either $\delta P(e^{i\theta_j}) > 0$ in all the points where $\hat{P}$ vanishes or $\delta P(e^{i\theta_j}) \geq 0$ with the extra condition that, when it vanishes at some of those same points, the derivative vanishes as well. We only need to consider the interior of the set of admissible perturbations. This is

$$\mathcal{D} := \{ \delta P \,|\, \langle \delta P, \gamma_j \rangle > 0, \quad j = 1, 2, \ldots \ell \}. \tag{36}$$

It is clear that $\mathcal{D}$ is precisely the interior of the dual cone of Γ_+; i.e., $\mathcal{D} = \text{int}(\Gamma_+{}^{\text{dual}})$. Since, therefore, (35) holds for all $\delta P \in \text{int}(\Gamma_+{}^{\text{dual}})$, it follows that $\hat{G} \in (\text{int}(\Gamma_+{}^{\text{dual}}))^{\text{dual}} = \Gamma_+$, as claimed. This establishes (12b). Equation (12a) follows by substituting the values of the optimal $g_1, g_2, \ldots, g_n$ given by (12b) into (11).

To wrap up the proof of Theorem 1 and explain some connections, we return to the regularized problem of Section IV. From (16) we have

$$\mathbb{S}_\varepsilon(P_\varepsilon, Q_\varepsilon) = \int \log \frac{P_\varepsilon}{Q_\varepsilon \Phi} d\mu + \varepsilon \int \left(1 - \frac{1}{P_\varepsilon} - \log P_\varepsilon \right) d\mu. \tag{37}$$

By taking limits as $\varepsilon \to 0$, and using (31), which of course in particular holds for $(\hat{P}, \hat{Q}) = (P_0, Q_0)$, we see that

$$g_0 = \lim_{\varepsilon \to 0} \varepsilon \int \frac{1}{P_\varepsilon} d\mu \tag{38a}$$

which of course is finite, since all other terms in (37) have finite limits. The value of g_0 is also nonnegative as seen from (38a), since $P_0 \geq 0$. Similarly, from (15) and (24b), we also see that

$$g_k = \lim_{\varepsilon \to 0} \varepsilon \int \cos k\theta \frac{1}{P_\varepsilon} d\mu \tag{38b}$$

which certainly is consistent with (12b). From (38), we can again deduce that, when $P_0 \in \mathcal{P}_+$, $g_k = 0$ for $k = 0, 1, \ldots, m$. This concludes the proof of Theorem 3.

The remaining statements (iv) and (v) of Theorem 1 now follow directly from the fact that any (P, Q) that satisfies the moment conditions (8) and (9) is also a minimizer (Theorem 3). Hence, if $P \in \mathcal{P}_+$ and either $(P, Q) \in \mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$ have a common factor or they are both degree deficient, then we can introduce or alter existing common factors without violating the moment conditions or the normalization (6a) (which can be seen

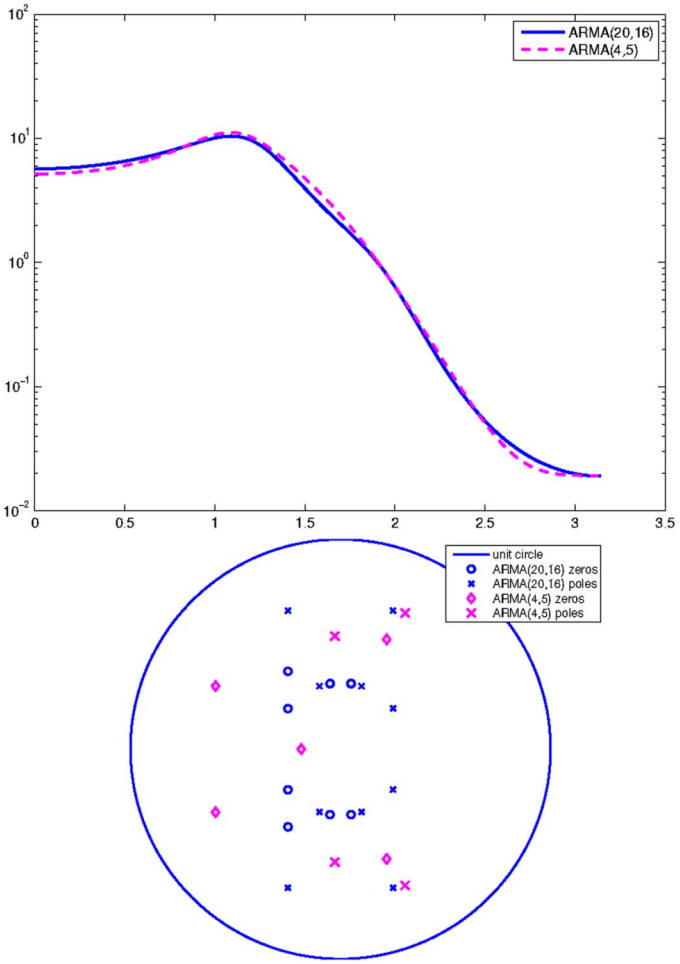


Fig. 2. High-order power spectrum and ARMA$(4,5)$ approximant (above) and poles/zeros pattern (below).

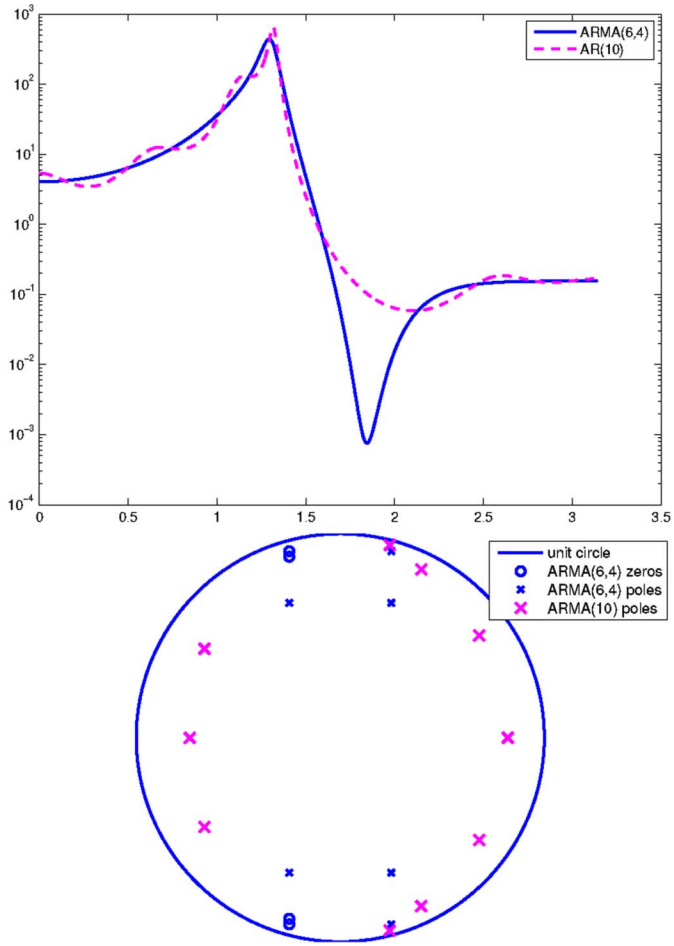


Fig. 3. Power spectrum and AR(10) approximant (above) and poles/zeros pattern (below).

as before, by summing the covariance moment conditions after we multiply by q_k, respectively). Notice that this may not be possible when $P \in \partial\mathcal{P}$ because of the slack variables. This concludes the proof of Theorem 1.

VI. Numerical Computation

The proof of Theorem 3 in the previous section, suggests a numerical scheme for approximating the optimal solution. We collect the relevant facts in the following proposition.

Proposition 5: Let $(P_\varepsilon, Q_\varepsilon)$ be the unique minimizer of $\mathbb{S}_\varepsilon$ in $\mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$, as per Theorem 4. Then, as $\varepsilon \to 0$

$$(P_\varepsilon, Q_\varepsilon) \to (P_0, Q_0)$$

where (P_0, Q_0) is a minimizer of $\mathbb{S}(P,Q)$ in $\mathcal{P} \times (\mathcal{Q} \cap \mathcal{H})$.

It is shown in Section IV that $(P_\varepsilon, Q_\varepsilon)$ is the unique minimizer of the strictly convex functional (Lagrangian)

$$\begin{aligned}\psi(x) &:= L(P,Q,1)\\ &= \int \left(P \log \frac{P}{Q\Phi} - \varepsilon \log P \right) d\mu\\ &\quad + \left(\int Q\Phi \, d\mu - 1 \right)\end{aligned}$$

where

$$x = (q_0, q_1, \dots, q_n, p_1, \dots, p_m)$$

is the vector of coefficients of $P = 1 + \sum_{k=1}^m p_k \cos(k\theta)$ and $Q = \sum_{k=0}^n q_k \cos(k\theta)$. As indicated in (20), the Lagrange multiplier in (17) is set to the optimal value $\lambda = 1$. The computation of the minimizer can be done using Newton's method. More specifically, choosing a suitably small $\varepsilon > 0$, iterate

$$x_{k+1} = x_k - \nu(\psi_{xx}(x_k))^{-1}\psi_x(x_k)$$

starting from initial conditions $x_0 = (1, 0, \dots, 0)'$ and a suitable step size ν. Here, ψ_x is the gradient

$$\psi_x = \begin{bmatrix} \chi \\ \pi \end{bmatrix}$$

where the entries of the column vectors χ and π are

$$\chi_k = r_k - \int \cos k\theta \frac{P}{Q} d\mu$$

with $k = 0, \dots, n$, and

$$\pi_k = \int \cos k\theta \log \frac{P}{Q} d\mu - c_k - \epsilon \int \cos k\theta \frac{d\mu}{P}$$

with $k = 1, \dots, m$. Next, ψ_{xx} is the Hessian

$$\psi_{xx}(x) = \begin{bmatrix} A & B \\ B' & C \end{bmatrix}$$

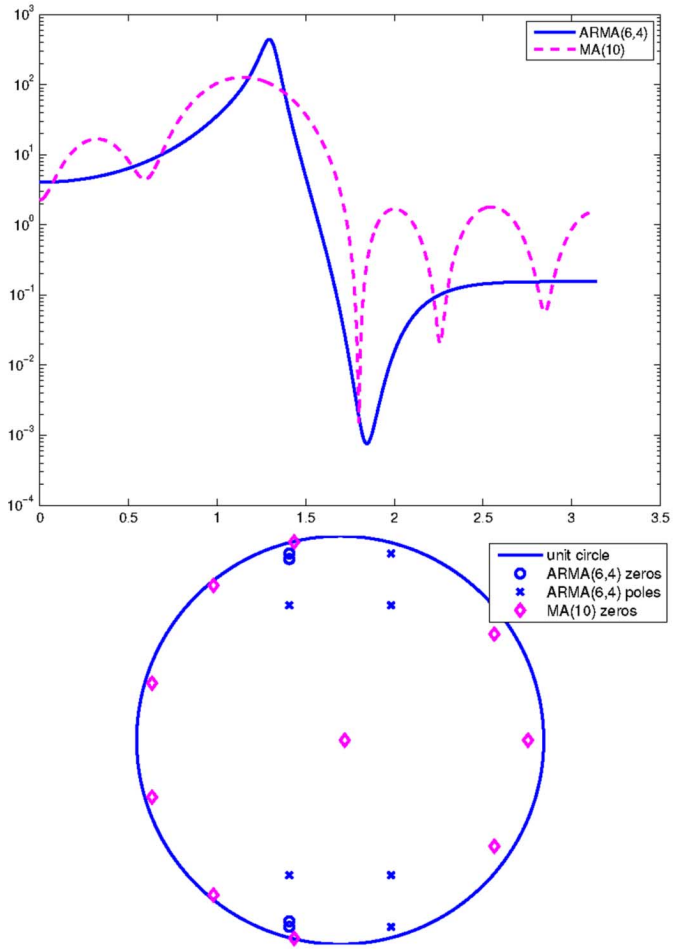


Fig. 4. Power spectrum and MA(10) approximant (above) and poles/zeros pattern (below).

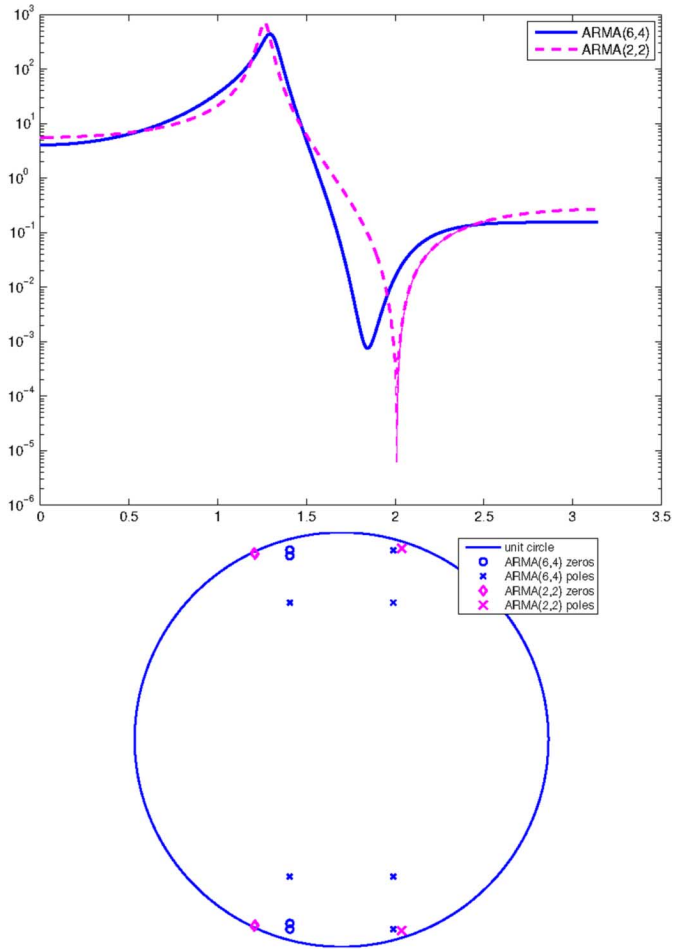


Fig. 5. Power spectrum and ARMA$(2, 2)$ approximant (above) and poles/zeros pattern (below)

where A, B, C take values in $\mathbb{R}^{(n+1)\times(n+1)}$, $\mathbb{R}^{(n+1)\times m}$, and $\mathbb{R}^{m\times m}$, respectively, and have entries

$$a_{k\ell} = \int \cos k\theta \cos \ell\theta \, \frac{d\mu}{Q^2}$$

with $k, \ell = 0, \ldots, n$,

$$b_{k\ell} = -\int \cos k\theta \cos \ell\theta \, \frac{d\mu}{Q}$$

with $k = 0, \ldots, n$, and $\ell = 1, \ldots, m$,

$$c_{k\ell} = \int \cos k\theta \cos \ell\theta \, \frac{(\epsilon + P)}{P^2} d\mu$$

with $k, \ell = 1, \ldots, m$.

VII. Spectral Approximation: Case Studies

We begin with a power spectral density Φ of high order shown in Fig. 2, which corresponds to an ARMA$(20, 16)$ model. The poles and zeros of the corresponding canonical spectral factor are also shown in the same figure. This is an example of a rather tame "high order" power spectrum which can be easily approximated by a low order one. For the appoximation we select $n = 4$, $m = 5$, and hence the approximant corresponding to an ARMA$(4, 5)$ model. This indeed is capable of matching perfectly the set of the first five covariance samples as well as the set of the first five cepstral coefficients. This "low order" power spectrum and its corresponding pole/zero pattern are superimposed with those of Φ in the same figures.

When a power spectrum has a number of poles and zeros near the unit circle, then it may be impossible to match perfectly all relevant cepstral coefficients with a low order model (i.e., m of them for an ARMA(n, m) approximant). We highlight the ability of low order approximants to follow the "shape" of Φ in a series of representative cases. The power spectral density Φ that we have selected corresponds to an ARMA$(6, 4)$ model with pairs of poles and zeros near each other in the unit disc. We display Φ together with approximating spectra of lower order ones, AR(10), MA(10), ARMA$(2, 2)$, and ARMA$(4, 5)$, respectively, along with the corresponding pole/zero patterns superimposed with those of Φ, in separate graphs in Figs. 3–6. In all these cases, the approximating power spectra have zeros on the boundary and are unable to match the cepstral coefficients. This is examplified for the case that corresponds to an ARMA$(2, 5)$ model in Fig. 7 with the corresponding power spectrum and poles/zeros pattern shown in Fig. 8. It is worthing noting the improvement in matching the actual "shape" of Φ

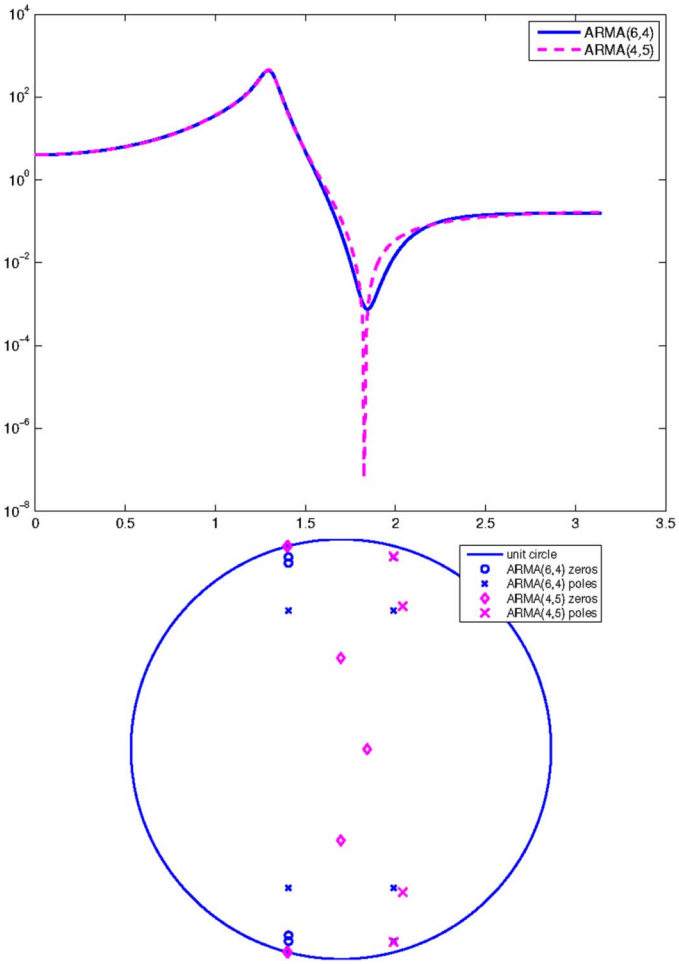


Fig. 6. Power spectrum and ARMA$(4,5)$ approximant (above) and poles/zeros pattern (below)

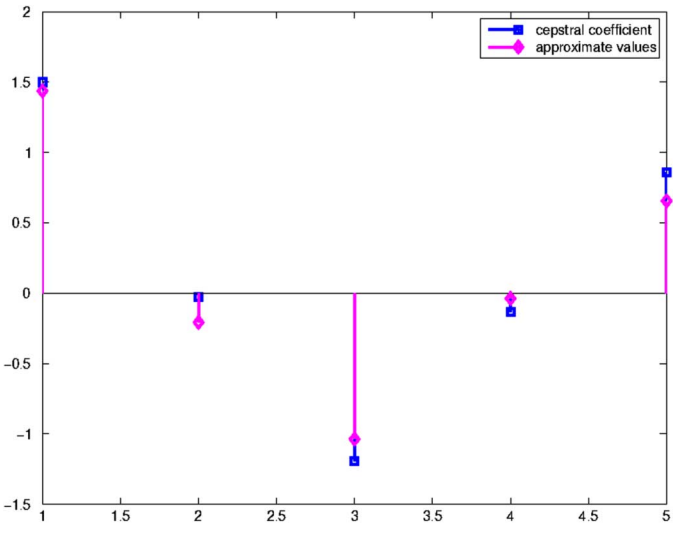


Fig. 7. Cepstral coefficients of power spectrum and of ARMA$(2,5)$ approximant.

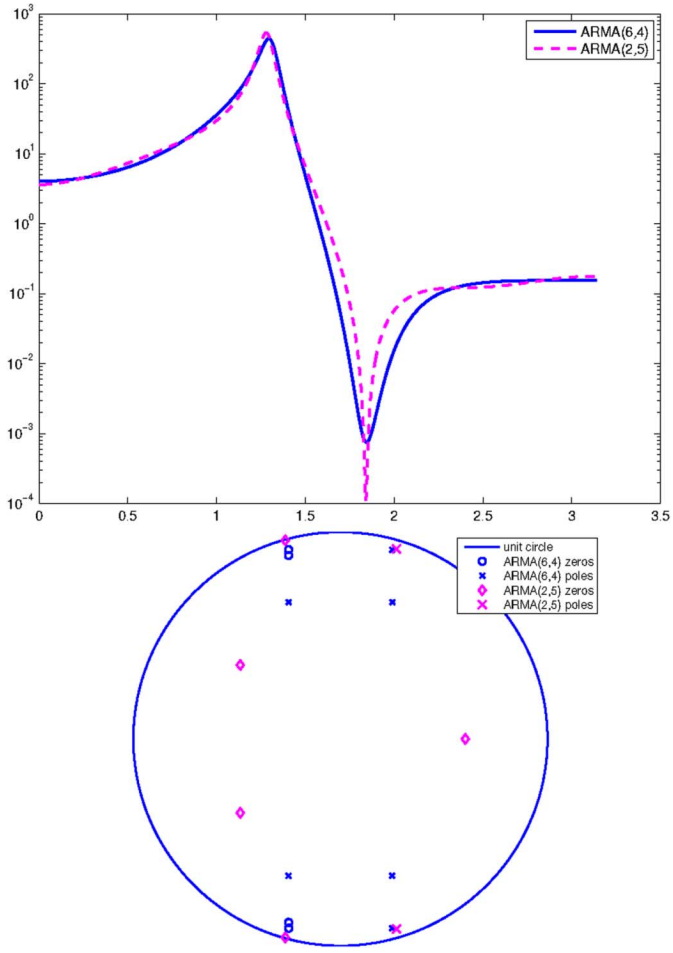


Fig. 8. Power spectrum and ARMA$(2,5)$ approximant (above) and poles/zeros pattern (below).

while comparing these cases. Interestingly enough, the peak can be reproduced relatively accurately with one pair of complex poles. This is not the case for the "valley" in Φ because it is produced by two pairs of complex zeros. For matching the "valley," a higher order MA-part is needed. However, despite the fact that the MA-part is of order five in the two examples in Figs. 6 and 8, the approximants do not match the specific zero pattern.

An alternative set of examples is displayed in Figs. 9 and 10. In these, we observe the inability of AR models to match the "shape" of a rather flat power spectrum with significant "valleys." Despite the fact that the original spectrum now corresponds to an ARMA$(6,6)$ model, relatively good fit is achieved with an MA(8) model, as shown in the last plot.

VIII. Connections to Cepstral Approximation

We proceed to explain the connection between the approximation problem in the present work and a seemingly unrelated problem in [3] and [4]. A central theme in [3] and [4] was to determine a rational power spectral density which matches a set of prescribed covariance lags while, at the same time, approximates a set of given cepstral coefficients. To this end, the functional

$$\mathbb{I}(\Psi) := \sum_{k=1}^{m} \left| \int \cos k\theta \log \Psi(e^{i\theta}) d\mu - c_k \right| - \int \log \Psi(e^{i\theta}) d\mu$$

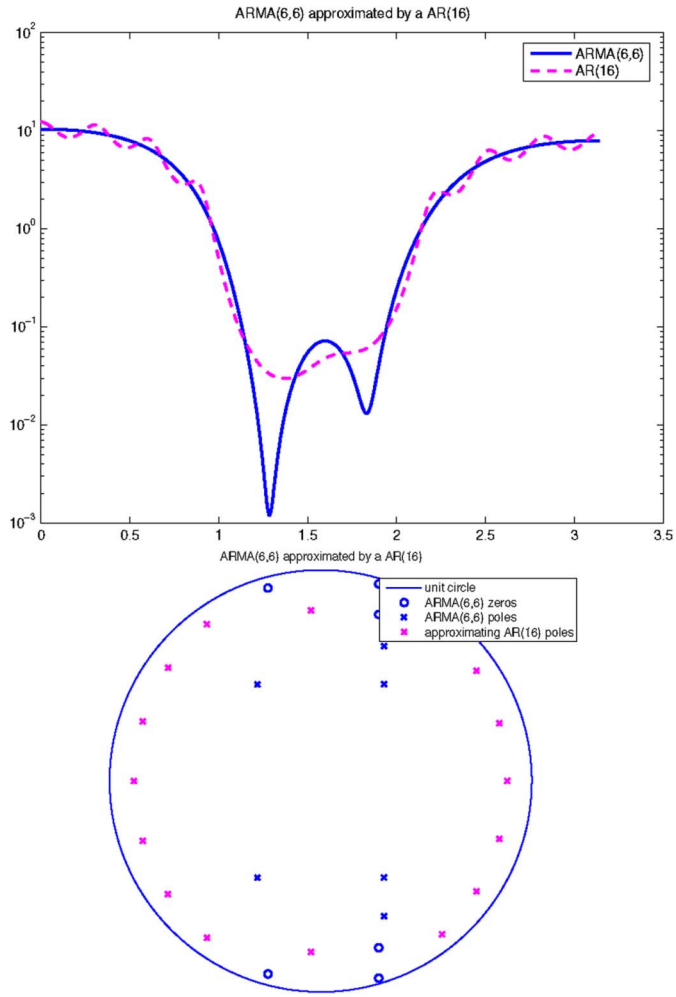


Fig. 9. Power spectrum and AR(16) approximant (above) and pole/zero patterns (below).

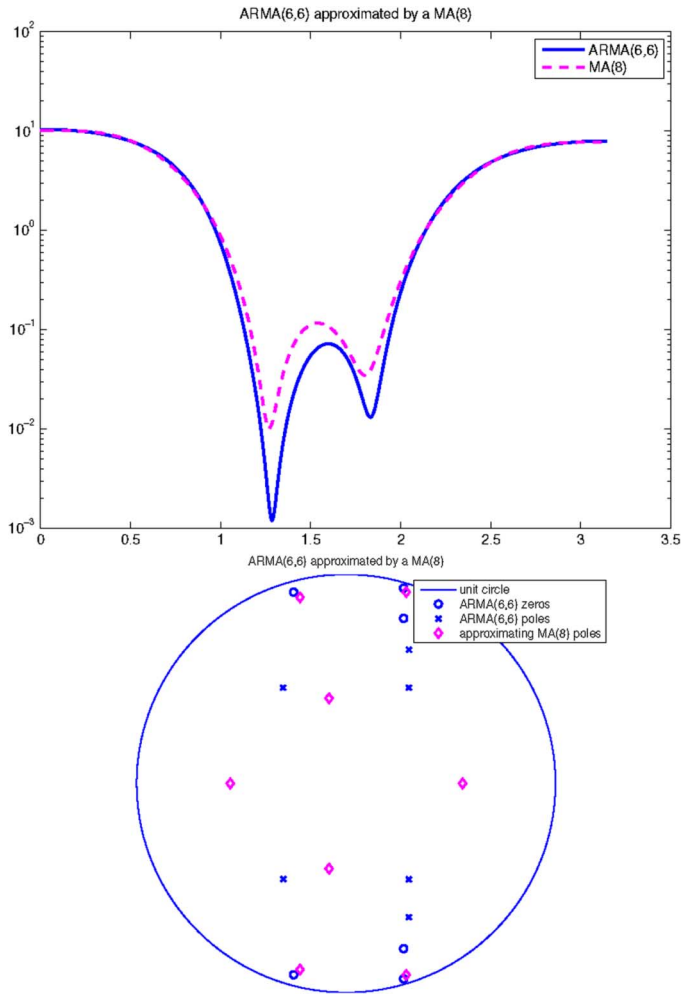


Fig. 10. Power spectrum and MA(8) approximant (above) and pole/zero pattern (below).

was introduced in [3] and [4], to be minimized subject to the matching conditions

$$\int \cos k\theta \Psi(e^{i\theta}) d\mu = r_k, \quad k = 0, 1, \ldots, n$$

for a given set of covariance lags $r_0, r_1, \ldots, r_n$ and cepstral coefficients $c_1, \ldots, c_m$. This cost functional trades off maximization of entropy gain against approximating the cepstral coefficients.

The connection to our present work is through the dual of this optimization problem, namely to maximize the concave functional

$$\mathbb{J}(P,Q) = c_1 p_1 + \cdots + c_m p_m - r_0 q_0 - \cdots - r_n q_n - \int P(e^{i\theta}) \log \frac{P(e^{i\theta})}{Q(e^{i\theta})} d\mu \quad (39)$$

over $\mathcal{P} \times \mathcal{Q}$. It was shown in [4, Th. 5.3] that there exists at least one solution $(\hat{P}, \hat{Q})$ to this problem, and $\hat{\Psi} = (\hat{P}/\hat{Q})$ is the solution of the primal problem. For any such maximizer $(\hat{P}, \hat{Q})$, we have $\hat{Q} \in \mathcal{Q}_+$ and exact covariance matching. If $\hat{P} \in \mathcal{P}_+$, there is also exact cepstral matching. In this case $(\hat{P}, \hat{Q})$ is unique, and this happens if and only if $\hat{P}$ and $\hat{Q}$ are coprime. Hence there is an analogous set of conclusions for this pair of dual optimization problems to those in Theorem 1.

As explained in Section V, the optimal solution $(\hat{P}, \hat{Q})$ of the optimization problem of Section III is obtained by minimizing $L(P, Q, 1, \hat{g}_0)$, where the Lagrangian L is given by (26), and $\hat{g}_0$ is the optimal solution (31) of the corresponding dual problem. However

$$L(P,Q,1,\hat{g}_0) = \int P \log \frac{P}{Q} d\mu - \sum_{k=0}^{m} p_k \int \cos k\theta \, \log \Phi \, d\mu + \sum_{k=0}^{n} q_k \int \cos k\theta \, \Phi \, d\mu - 1 - \sum_{k=0}^{m} \hat{g}_k p_k$$

which, in view of (4), can be written

$$L(P,Q,1,\hat{g}_0) = r_0 q_0 + \cdots + r_n q_n - (c_1 + \hat{g}_1) p_1 - \cdots - (c_m + \hat{g}_m) p_m + \int P \log \frac{P}{Q} d\mu - 1 - c_0 - \hat{g}_0.$$

Since c_0 and $\hat{g}_0$ are constant, minimizing $(P,Q) \mapsto L(P,Q,1,\hat{g}_0)$ over $\mathcal{P} \times \mathcal{Q}$ is equivalent to maximizing

$$(c_1+\hat{g}_1)p_1+\cdots+(c_m+\hat{g}_m)p_m - r_0q_0-\cdots-r_nq_n-\int P\log\frac{P}{Q}\,d\mu$$

which is precisely the dual problem (39) with the cepstral coefficients appropriately modified to account for slack variables. This explains the congruity of the optimality conditions between the two problems.

IX. Concluding Remarks

We have presented a model matching approach to spectral density approximation using the Kullback–Leibler divergence as a criterion for goodness of fit. The approach yields a *convex optimization* procedure for ARMA modeling. The optimality conditions are given in terms of moments of the spectral density and its logarithm. This fact makes the approach potentially useful to system identification. Moments of spectral density functions are routinely computed in applications requiring spectral estimation [21]. While statistical estimation of covariance lags is reasonably well studied [17], the estimation of cepstral coefficients remains a topic of current research (see, e.g., [13]).

The current paper provides a motivation for the study in [3] and [4]. Indeed, while [3] and [4] focuses on covariance and cepstral matching, the present work provides an approximation theoretic justification.

Acknowledgment

The authors would like to thank Dr. Per Enqvist for bringing [19] to our attention.

References

[1] S. Amari, *Differential-Geometrical Methods in Statistics*. Berlin, Germany: Springer-Verlag, 1985, Lecture notes in statistics.

[2] S. Amari and H. Nagaoka, *Methods of Information Geometry*, ser. Transactions of Mathematical Monographs. Providence, RI: American Mathematical Society, 2000, vol. 191.

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

[4] C. I. Byrnes, P. Enqvist, and A. Lindquist, "Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach," *SIAM J. Contr. Optim.*, vol. 41, pp. 23–59, 2002.

[5] C. Byrnes, T. T. Georgiou, and A. Lindquist, "A generalized entropy criterion for Nevanlinna-Pick interpolation: A convex optimization approach to certain problems in systems and control," *IEEE Trans. Autom. Control*, vol. 45, no. 6, pp. 822–839, Jun. 2001.

[6] 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*, vol. 48, no. 11, pp. 3189–3206, Nov. 2000.

[7] C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev, "A complete parameterization of all positive rational extensions of a covariance sequence," *IEEE Trans. Autom. Control*, vol. 40, pp. 1841–1857, 1995.

[8] C. I. Byrnes, S. V. Gusev, and A. Lindquist, "A convex optimization approach to the rational covariance extension problem," *SIAM J. Contr. Opt.*, vol. 37, pp. 211–229, 1999.

[9] C. I. Byrnes, S. V. Gusev, and A. Lindquist, "From finite covariance windows to modeling filters: A convex optimization approach," *SIAM Rev.*, vol. 43, pp. 645–675, Dec. 2001.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[11] P. Enqvist, "Spectral Estimation by Geometric, Topological and Optimization Methods," Ph.D. dissertation, Optimization and Systems Theory, KTH, Stockholm, Sweden, 2001.

[12] P. Enqvist, "A convex optimization approach to ARMA(n,m) model design from covariance and cepstrum data," *SIAM J. Contr. Optim.*, vol. 43, no. 3, pp. 1011–1036, 2004.

[13] Y. Ephraim and M. Rahim, "On second-order statistics and linear estimation of cepstral coefficients," *IEEE Trans. Speech Audio Processing*, vol. 7, no. 2, pp. 162–176, 1999.

[14] T. T. Georgiou, "Distances and Riemannian metrics for spectral density functions," *IEEE Trans. Signal Processing*, vol. 55, no. 8, pp. 3995–4003, Aug. 2007.

[15] T. T. Georgiou and A. Lindquist, "Kullback-Leibler approximation of spectral density functions," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, Nov. 2003.

[16] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Berkeley, CA: University of California Press, 1958.

[17] M. G. Kendall, A. Stuart, and J. K. Ord, *Advanced Theory of Statistics*, 4th ed. New York: McMillan, 1983, vol. 2.

[18] S. Kullback, *Information Theory and Statistics*, 2nd ed. New York: Dover, 1968.

[19] C. T. Mullis and R. A. Roberts, "The use of second-order information in the approximation of discrete-time linear systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, no. 3, pp. 226–238, Jun. 1976.

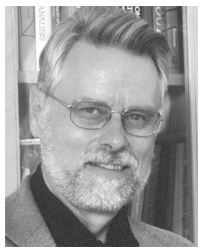
[20] P. A. Regalia and P. Stoica, "Stability of multivariable least-squares models," *IEEE Signal Processing Lett.*, vol. 2, no. 10, pp. 195–196, Oct. 1995.

[21] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice Hall, 1997.

Tryphon T. Georgiou (M'79–SM'99–F'00) was born in Athens, Greece, on October 18, 1956. He received the Diploma in mechanical and electrical engineering from the National Technical University of Athens, Greece, in 1979, and the Ph.D. degree from the University of Florida, Gainesville, in 1983.

He has served on the faculty of Florida Atlantic (1983–1986) and Iowa State (1986–1989) Universities. Since 1989, he has been with the University of Minnesota, Minneapolis, where he holds the Vincentine Hermes-Luh Chair of Electrical Engineering.

Dr. Georgiou has been a recipient of the George S. Axelby Outstanding Paper award of the IEEE Control Systems Society three times, for the years 1992, 1999, and 2003. In 1992 and 1999, he received the award for joint work with Prof. Malcolm C. Smith (Cambridge Univ., U.K.), and in 2003 for joint work with Professors Chris Byrnes (Washington Univ., St. Louis) and Anders Lindquist (KTH, Stockholm). He has served as an Associate Editor for the IEEE Transactions on Automatic Control (1991–1992), the *SIAM Journal on Control and Optimization* (1988–1995), the *Systems and Control Letters* (1995-present), and the IEEE Transactions on Signal Processing (currently). He has served on the Board of Governors of the Control Systems Society of the IEEE (2002–2005).



Anders Lindquist (M'77–SM'86–F'89) received the Ph.D. degree from the Royal Institute of Technology, Stockholm, Sweden, in 1972.

From 1972 to 1974, he held visiting positions at the University of Florida, Brown University, and State University of New York at Albany. In 1974, he became an Associate Professor, and in 1980, a Professor at the University of Kentucky, Lexington, where he remained until 1983. He is now a Professor at the Royal Institute of Technology, where in 1982, he was appointed to the Chair of Optimization and Systems Theory. Presently, he is the Head of the Mathematics Department and Director of the Strategic Research Center for Industrial and Applied Mathematics, both at the Royal Institute of Technology. Since 1989, he has also been an Affiliate Professor of Optimization and Systems Theory at Washington University, St. Louis, MO.

Dr. Lindquist is a Member of the Royal Swedish Academy of Engineering Sciences, a Foreign Member of the Russian Academy of Natural Sciences, and an Honorary Member the Hungarian Operations Research Society. He has served on many editorial boards of journals and book series. He is the recipient (together with C. I. Byrnes and T. T. Georgiou) of the George S. Axelby Outstanding Paper Award of the IEEE Control Systems Society (CSS) for the year 2003.