Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

A new algorithm for Kullback-Leibler approximation of spectral densities

Michele Pavon and Augusto Ferrante

Abstract— **The Georgiou-Lindquist problem of Kullback-Leibler approximation of a spectrum with spectra that are consistent with given state-covariance statistics is considered. An alternative solution algorithm is presented. Simulation indicates that the algorithm is extremely effective.**

I. INTRODUCTION

In a recent paper [1], Georgiou and Lindquist have introduced and solved the problem of approximating spectra in the Kullback-Leibler pseudo-distance with spectra that are consistent with given state-covariance statistics. The authors show that their mathematical framework accomodates a variety of important problems as special cases (see the examples in the following section). In [1], the (convex) optimization problem was approached via duality theory. It was shown that the dual problem is also convex and admits a unique solution in a certain prescribed set of Hermitian matrices. In general, the solution cannot be obtained in closed form. It is then necessary to solve the dual problem numerically. In [1], it is observed that such a numerical solution can be based on Newton's method along the lines of the methods developed in [4], [16], [17] for similar problems. As observed in [1], these methods all have advantages and disadvantages, the latter being possible unboundness of the gradient at the boundary or loss of global convexity.

In this paper, which is a shortened version without detailed proofs of [2], we introduce an alternative algorithm to solve the approximation problem based on a nonlinear iteration scheme for the Lagrange multiplier matrix. Simulation results indicate that this iteration converges extremely fast.

The outline of the paper is as follows. In Sections II, III and IV we review, with some minor differences, the problem setting, the variational analysis and the corresponding dual problem of Georgiou and Lindquist, respectively. In Section V, we introduce the new algorithm. A theoretical analysis of the algorithm is provided in Section VI. Simulation results are presented in Section VII.

Work partially supported by the MIUR-PRIN Italian grant "Identification and Control of Industrial Systems"

M. Pavon is with the Dipartimento di Matematica Pura e Applicata, Università di Padova, via Belzoni 7, 35131 Padova, Italy pavon@math.unipd.it

A. Ferrante is with the Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy augusto@dei.unipd.it

II. THE APPROXIMATION PROBLEM

In order to facilitate the reader, we shall adopt, with a few exceptions, the same notation as in [1]. We denote by $C_+(\mathbb{T})$ the everywhere positive, continuous functions on the unit circle. We consider the rational transfer function

$$G(z) = (I - zA)^{-1}B, A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{n\times 1},$$

where A is a stability matrix, i.e. has all its eigenvalues in the open unit disc, and (A, B) is a reachable pair. Namely. the controllability matrix $\mathcal{C}_n = [B, AB, A^2B, \ldots, A^{n-1}B]$ has full rank. Let $\Psi \in C_+(\mathbb{T})$ represent the *a priori* estimate of the spectrum of an underlying zero-mean, wide-sense stationary stochastic process $\{y(n), n \in \mathbb{Z}\}$. Consider the situation where new data become available in the form of an estimate Σ of the state covariance of the system

The Hermitian matrix Σ is assumed to be positive definite. It is obtained by feeding y to the measurement device (a bank of filters) modeled by G until it reaches steady state, and then estimating the state covariance. If Ψ is not consistent with Σ, we need to find Φ in $C_+(\mathbb{T})$ that is closest to Ψ in a suitable sense among spectra consistent with Σ. In [1], the following measure of distance for spectra in $C_+(\mathbb{T})$ was introduced:

$$\mathbb{S}(\Psi\|\Phi) = \int \Psi \log\left(\frac{\Psi}{\Phi}\right) = \int_{-\pi}^{\pi} \Psi(e^{i\theta}) \log\left(\frac{\Psi(e^{i\theta})}{\Phi(e^{i\theta})}\right) \frac{d\theta}{2\pi}.$$

The Kullback-Leibler pseudo-distance for probability distributions is defined similarly. It is also known as *divergence*, *relative entropy*, *information distance*. etc., and it is also denoted by $\mathbb{D}(\cdot,\cdot)$ and $H(\cdot,\cdot)$. If

$$\int \Phi = \int \Psi,$$

we have $\mathbb{S}(\Psi\|\Phi) \geq 0$. The choice of $\mathbb{S}(\Psi\|\Phi)$ as a distance measure, even for spectra that have different zeroth moment, is justified in [1, Section III]. Notice that minimizing $\mathbb{S}(\Psi\|\Phi)$ rather than $\mathbb{S}(\Phi\|\Psi)$ is unusual with respect to the statistics-probability-information theory world. Besides leading to a more tractable problem, however, it also includes as special case ($\Psi \equiv 1$) maximization of entropy. As in [1], we consider the following

0-7803-9568-9/05/$20.00 ©2005 IEEE

Problem 2.1: (Approximation problem) Let $\Psi \in C_+(\mathbb{T})$, and let $\Sigma \in \mathbb{C}^{n\times n}$ satisfy $\Sigma = \Sigma^* > 0$. Find $\hat{\Phi}$ that solves

$$\text{minimize} \quad \mathbb{S}(\Psi\|\Phi) \tag{1}$$

$$\text{over} \quad \left\{\Phi \in C_+(\mathbb{T}) \mid \int G\Phi G^* = \Sigma\right\}, \tag{2}$$

where star denotes transposition plus conjugation.

The constraint (2) expresses the fact that Φ is consistent with Σ. It is not necessary to require that Φ satisfies the *parahermitian* property $\Phi(e^{i\theta})^* = \Phi(e^{i\theta}), \forall\theta$. Indeed, whenever $\Psi \in C_+(\mathbb{T})$ is a spectrum, namely it satisfies the parahermitian property, the solution of the approximation problem also turns out to be a spectrum. First of all, one needs to worry about existence of $\Phi \in C_+(\mathbb{T})$ satisfying constraint (2). It was shown in [13], [14] that such family is nonempty if and only if there exists $H \in \mathbb{C}^{1\times n}$ such that

$$\Sigma - A\Sigma A^* = BH + H^*B^*,$$

or, equivalently, the following rank condition holds

$$\text{rank}\begin{pmatrix} \Sigma - A\Sigma A^* & B \\ B^* & 0 \end{pmatrix} = \text{rank}\begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix} \tag{3}$$

We now show, following [1], that this problem includes as special cases a variety of important applications. In all three examples below, the family of spectral densities consistent with the data is nonempty if $\Sigma \geq 0$ and contains infinitely many elements if $\Sigma > 0$.

Example 1: (*Covariance extension problem*) Let

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{4}$$

so that the k-th component of $G(z)$ is $G_k(z) = z^{n-k}$. Also let Σ be the Toeplitz matrix:

$$\Sigma = \begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ \bar{c}_1 & c_0 & \dots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{c}_{n-1} & \bar{c}_{n-2} & \dots & c_0 \end{bmatrix} \tag{5}$$

where

$$c_k := E\{y(n)\bar{y}(n+k)\}.$$

Thus, the information available on the process y is the finite sequence of covariance lags $c_0, c_1, \dots, c_{n-1}$. If $\Psi \equiv 1$, then Problem 2.1 is just the maximum-entropy covariance extension problem, see e.g. [18]. When Ψ is a positive polynomial of degree n, the solution of Problem 2.1 has degree $2n$ [10], [12], [8], [7], [5], [6]

Example 2: In this case

$$A = \begin{bmatrix} p_1 & 0 & 0 & \dots & 0 \\ 0 & p_2 & 0 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & p_n \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \tag{6}$$

so that the k-th element of $G(z)$ is $G_k(z) = \frac{1}{1-p_k z}$. The matrix Σ is a Pick matrix with elements $\Sigma_{i,j} = \frac{w_i + \bar{w}_j}{1-p_i\bar{p}_j}$ where

$$w_k = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{e^{-i\omega} + p_k}{e^{-i\omega} - p_k}\Phi(e^{i\omega})d\omega, \quad k = 1, 2, \dots, n. \tag{7}$$

In this case the problem is a *Nevanlinna-Pick interpolation problem* [11], [12], [3], [4] whose solution permits *spectral estimation by selective harmonic amplification* [3], [4], [13]. Similarly for Example 3 below.

Example 3: In this case

$$A = \begin{bmatrix} p & 1 & 0 & \dots & 0 \\ 0 & p & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & p \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{8}$$

so that the k-th element of $G(z)$ is $G_k(z) = \frac{z^{n-k}}{(1-pz)^{n+1-k}}$. The matrix Σ is of the form

$$\Sigma = \frac{1}{2}(WE + EW^*) \tag{9}$$

where W is an upper triangular toeplitz matrix and E is the reachability gramian of the pair (A, B).

III. VARIATIONAL ANALYSIS

For $\Lambda \in \mathbb{C}^{n\times n}$, we consider the *Lagrangian function*

$$\begin{aligned} L(\Phi, \Lambda) &= \mathbb{S}(\Psi\|\Phi) + \text{trace}\left(\Lambda\left(\int G\Phi G^* - \Sigma\right)\right) \\ &= \mathbb{S}(\Psi\|\Phi) + \int G^*\Lambda G\Phi - \text{trace}\,(\Lambda\Sigma). \end{aligned} \tag{10}$$

In [1], the matrix Λ is immediately restriced to the subset of matrices satisfying $G^*\Lambda G > 0$ on all of $\mathbb{T}$, and belonging to the range of the operator Γ defined on $C(\mathbb{T})$ by

$$\Gamma(\Phi) = \int G\Phi G^*. \tag{11}$$

For reasons that will be apparent below, we prefer, for the moment, not to pose any restriction on Λ. Next, we consider the unconstrained minimization of $L(\Phi, \Lambda)$.

$$\text{minimize}\{L(\Phi, \Lambda)|\Phi \in C_+(\mathbb{T})\} \tag{12}$$

Remark 3.1: First of all, observe that $C_+(\mathbb{T})$ is a convex subset of $C(\mathbb{T})$. Second, notice that, for each Λ, the functional

$$\Phi \to L(\Phi, \Lambda)$$

is strictly convex. Thus, (12) is a convex optimization problem. Furthermore, we observe:

- any $\Phi \in C_+(\mathbb{T})$ has a positive minimum on $\mathbb{T}$ (is bounded away from zero). Hence, for any continuous function $\delta\Phi$, there exists $\epsilon(\delta\Phi) > 0$ such that $(\Phi + \epsilon\delta\Phi) \in C_+(\mathbb{T}), \forall|\epsilon| < \epsilon(\delta\Phi)$. Hence, any Φ is an internal point of $C_+(\mathbb{T})$ in any direction $\delta\Phi$ (in other

words, $C_+(\mathbb{T})$ is an open subset of $C(\mathbb{T})$ with respect to the norm topology, where $\|\Phi\| = \sup_\theta |\Phi(e^{i\theta})|$);

- the smoothness of the integrands in (10) makes $L(\Phi, \Lambda)$ Gâteaux differentiable at Φ in any direction $\delta\Phi \in C(\mathbb{T})$.

It then follows from a fundamental result of convex optimization that $\hat{\Phi} \in C_+(\mathbb{T})$ solves problem (12) if and only if it satisfies the condition

$$\delta L(\hat{\Phi}, \Lambda; \delta\Phi) = 0, \quad \forall \delta\Phi \in C(\mathbb{T}). \tag{13}$$

Here, $\delta L(\hat{\Phi}, \Lambda; \delta\Phi)$, the first variation of L at $\hat{\Phi}$ in direction $\delta\Phi$, is defined by

$$\delta L(\hat{\Phi}, \Lambda; \delta\Phi) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[L(\hat{\Phi} + \epsilon\delta\Phi, \Lambda) - L(\hat{\Phi}, \Lambda) \right].$$

Proposition 3.2: $\hat{\Phi} \in C_+(\mathbb{T})$ *solves problem (12) if and only if it satisfies the condition*

$$\hat{\Phi} G^* \Lambda G = \Psi. \tag{14}$$

Proof: We compute

$$\delta L(\hat{\Phi}, \Lambda; \delta\Phi) = \int \left(G^* \Lambda G - \frac{\psi}{\hat{\Phi}} \right) \delta\Phi.$$

By (13), $\hat{\Phi} \in C_+(\mathbb{T})$ solves (12) if and only if

$$\int \left(G^* \Lambda G - \frac{\psi}{\hat{\Phi}} \right) \delta\Phi = 0, \quad \forall \delta\Phi \in C(\mathbb{T}). \tag{15}$$

In particular, if $\hat{\Phi}$ is optimal, this must be true for

$$\delta\bar{\Phi} := G^* \Lambda G - \frac{\psi}{\hat{\Phi}}.$$

We get

$$\int \left[G^* \Lambda G - \frac{\psi}{\hat{\Phi}} \right]^2 = 0.$$

Hence, recalling that $\hat{\Phi}$ is positive, (14) must be verified. Conversely, (14) implies immediately (15). Q.E.D.

Since both $\hat{\Phi}$ and Ψ are positive functions, we can have a solution of problem (12) only if $G^*\Lambda G > 0$ on $\mathbb{T}$. Hence, it is natural to restrict our "multiplier" Λ to Hermitian matrices satisfying this condition. We now come to an elementary, albeit fundamental, result in the spirit of Lagrange.

Theorem 3.3: Suppose $\hat{\Lambda} = \hat{\Lambda}^*$ is such that

$$G^* \hat{\Lambda} G > 0, \quad \forall e^{i\theta} \in \mathbb{T}, \tag{16}$$

$$\int G \frac{\Psi}{G^* \hat{\Lambda} G} G^* = \Sigma. \tag{17}$$

Then $\hat{\Phi}$ given by

$$\hat{\Phi} = \frac{\Psi}{G^* \hat{\Lambda} G} \tag{18}$$

is the unique solution of the approximation problem (1)-(2).

Proof: Let $\Phi \in C_+(\mathbb{T})$ satisfy the constraint

$$\int G \Phi G^* = \Sigma.$$

By Theorem (3.3), and by the strict convexity of the functional $L(\cdot, \hat{\Lambda})$, we get

$$\mathbb{S}(\Psi, \Phi) = L(\Phi, \hat{\Lambda}) > L(\hat{\Phi}, \hat{\Lambda}) = \mathbb{S}(\Psi, \hat{\Phi}).$$

Q.E.D.

Thus, the original problem (1)-(2) is now reduced to finding $\hat{\Lambda}$ satisfying (16)-(17). This is accomplished in [1] via duality theory.

IV. THE DUAL PROBLEM

Consider, as in [1], the dual functional

$$\Lambda \to \inf\{L(\Phi, \Lambda) | \Phi \in C_+(\mathbb{T})\}.$$

In view of Proposition 3.2, for Λ satisfying (16), the dual functional takes the form

$$\Lambda \to L(\frac{\Psi}{G^* \Lambda G}, \Lambda) = \int \Psi \log G^* \Lambda G - \text{trace } (\Lambda\Sigma) + \int \Psi. \tag{19}$$

We consider the maximization of the dual functional (19) over the set

$$\mathcal{L}_+ := \{\Lambda = \Lambda^* | G^* \Lambda G > 0, \forall e^{i\theta} \in \mathbb{T}\}. \tag{20}$$

Notice that we do not impose, as in [1], the condition that Λ should be in the range of the operator Γ. To avoid any confusion, we shall adopt here the following notation

$$\mathcal{L}_+(\Gamma) := \mathcal{L}_+ \cap \text{range } \Gamma.$$

Let, as in [1],

$$J_\Psi(\Lambda) := -\int \Psi \log G^* \Lambda G + \text{trace } (\Lambda\Sigma)$$

Our dual problem is then equivalent to

$$\text{minimize} \quad \{J_\Psi(\Lambda) | \Lambda \in \mathcal{L}_+\}, \tag{21}$$

whereas in [1] the dual problem is equivalent to minimize $\{J_\Psi(\Lambda) | \Lambda \in \mathcal{L}_+(\Gamma)\}$.

Remark 4.1: We observe that $\mathcal{L}_+$ is convex. Moreover, $\Lambda \to J_\Psi(\Lambda)$ is convex on $\mathcal{L}_+$, but, in general, *not strictly convex*. Thus, our dual problem is a convex optimization problem that, when solvable, may have many (even infinitely many) solutions. Moreover, $\mathcal{L}_+$ is an open subset of the set $\mathcal{H}$ of all $n \times n$ Hermitian matrices and J_Ψ is Gâteaux differentiable in any direction $\delta\Lambda \in \mathcal{H}$.

Hence, $\hat{\Lambda} \in \mathcal{L}_+$ solves (21) if and only if, for all $\delta\Lambda \in \mathcal{H}$, we have

$$\delta J_\psi(\hat{\Lambda}; \delta\Lambda) = \text{trace } \left[\left(\Sigma - \int G \frac{\Psi}{G^* \hat{\Lambda} G} G^* \right) \delta\Lambda \right] = 0. \tag{22}$$

Arguing as in the minimization of the Lagrangian, we get, the following result.

Proposition 4.2: $\hat{\Lambda} \in \mathcal{L}_+$ solves (21) if and only if it satisfies (17).

Remark 4.3: In [1], *existence* of a minimum of J_Ψ on $\mathcal{L}_+(\Gamma)$ is established via a profound homeomorphism result for continuous maps between open, connected sets of the

same dimension [9]. This result implies that, under assumption (3), there exists (a unique) $\hat{\Lambda}$ in $\mathcal{L}_+(\Gamma)$ satisfying (17). Such a $\hat{\Lambda}$ then provides the optimal solution of the primal problem (1)-(2) via (18). This, of course, establishes existence also for our dual problem (21) in view of Proposition (4.2).

It is observed in [1] that, in general, to solve their dual problem, it is necessary to resort to an iterative algorithm, and that the latter, after a suitable parametrization of $\mathcal{L}_+(\Gamma)$ can be based on Newton's method. In the following section, we propose an alternative strategy.

V. A NEW ALGORITHM

To simplify the writing, we assume, without loss of generality, that $\Sigma = I$. Indeed, if $\Sigma \neq I$, it suffices to replace G by $G' := \Sigma^{-1/2}G$ and (A, B) with $(A' = \Sigma^{-1/2}A\Sigma^{1/2}, B' = \Sigma^{-1/2}B)$. Moreover, if $\hat{\Lambda}'$ has been found such that

$$\int G' \frac{\Psi}{G'^* \hat{\Lambda}' G'} G'^* = I,$$

then

$$\int G \frac{\Psi}{G'^* \hat{\Lambda}' G'} G^* = \Sigma.$$

Hence, $\hat{\Phi}$ in (18) may also be obtained directly from G' and $\hat{\Lambda}'$ as

$$\hat{\Phi} = \frac{\Psi}{G'^* \hat{\Lambda}' G'}. \tag{23}$$

In order to have a nonempty feasible set for problem (1)-(2), we assume

$$\text{rank} \begin{pmatrix} I - AA^* & B \\ B^* & 0 \end{pmatrix} = \text{rank} \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix} \tag{24}$$

We assume, moreover, w.l.o.g., $\int \Psi = 1$.
Let

$$\mathcal{M} := \{M \in \mathcal{L}_+ | M \geq 0, \text{trace } [M] = 1, M \leq I\}, \tag{25}$$

$$\mathcal{M}_+ := \{M \in \mathcal{M} | M > 0\}. \tag{26}$$

For $M \in \mathcal{M}$, define the map Θ by

$$\Theta(M) := \int M^{1/2} G \left[\frac{\Psi}{G^* M G}\right] G^* M^{1/2}. \tag{27}$$

Theorem 5.1: Θ maps $\mathcal{M}$ into $\mathcal{M}$ and $\mathcal{M}_+$ into $\mathcal{M}_+$. The proof may be found in [2].

The algorithm: Let $M_0 = \frac{1}{n} I$. Note that $M_0 \in \mathcal{M}_+$. Define the sequence $\{M_k\}_{k=0}^{\infty}$ by

$$M_{k+1} := \Theta(M_k). \tag{28}$$

Notice that, by Theorem5.1, $M_k \in \mathcal{M}_+$ for all k. Moreover, since $M_k \in \mathcal{M}, \forall k$, the sequence is bounded. Hence it has at least one accumulation (limit) point in the closure $\bar{\mathcal{M}}$ of $\mathcal{M}$.

Proposition 5.2: Suppose that the sequence $\{M_k\}_{k=0}^{\infty}$ has a limit $\hat{M} \in \mathcal{M}_+$. Then $\hat{M}$ satisfies (16)-(17), and therefore provides the optimal solution of (1)-(2) via (18).

Proof: $\hat{M}$ is a fixed point of the map Θ.

$$\hat{M} = \Theta(\hat{M}). \tag{29}$$

Since $\hat{M} > 0$, (16) is satisfied and (29) gives (17). Q.E.D.

VI. THEORETICAL ANALYSIS OF THE ALGORITHM

First of all, we identify a large family of fixed points of Θ in $\bar{\mathcal{M}}$ that do not, in general, yield the optimal solution $\hat{\Phi}$ via (18).

Proposition 6.1: Let $x \in \mathbb{C}^n$ be such that $x^* G(e^{i\theta})$ is not identically zero. Then the orthogonal projection $\Pi_x \in \bar{\mathcal{M}}$ is a fixed point of the map Θ.

Proof: Recall that

$$\Theta(M) = \int \Pi_{M^{1/2}G} \Psi.$$

Notice that $x^* G$ is a rational function. If it is not identically zero, then it has at most finitely many zeroes on $\mathbb{T}$. Hence, $\Pi_{\Pi_x G} = \Pi_x$ almost everywhere on $\mathbb{T}$. Thus,

$$\Theta(\Pi_x) = \int \Pi_{\Pi_x G} \Psi = \Pi_x \int \Psi = \Pi_x.$$

Q.E.D.
Let us introduce $\mathcal{M}_0 := \{M \in \mathcal{M}_+ | \Theta(M) = M\}$, and $\bar{\mathcal{M}}_0$ the closure of $\mathcal{M}_0$.

Theorem 6.2: The set $\mathcal{M}_0$, and consequently its closure $\bar{\mathcal{M}}_0$, is convex.
For the proof see [2]. Now let M be a singular matrix in $\bar{\mathcal{M}}_0$. Then, there exist a sequence $\{M_k\}_{k=0}^{\infty}$ with elements in $\mathcal{M}_0$ such that $M_k \to M$. It follows that $G^* M_k G \to G^* M G$ at each point on $\mathbb{T}$. Since $G^* M_k G$ does not vary with k, it also follows that $G^* M G \equiv G^* M_k G$. We conclude that *any* matrix $M \in \bar{\mathcal{M}}_0$ yields the optimal solution of the approximation problem via the formula

$$\Phi^o := \frac{\Psi}{G^* M G} \tag{30}$$

Hence, it suffices that M_k defined in (28) converges to any element of $\bar{\mathcal{M}}_0$ to obtain the solution of the primal problem (1)-(2).
Now let

$$N_k = \int \frac{G \Psi G^*}{G^* M_k G}.$$

In order to facilitate the analysis of the algorithm, rewrite the recursion (28) as follows

$$M_{k+1} = M_k^{1/2} N_k M_k^{1/2}, \qquad M_0 = \frac{1}{n} I \tag{31}$$

$$N_{k+1} = \int \frac{G \Psi G^*}{G^* M_k^{1/2} N_k M_k^{1/2} G}, \quad N_0 = n \int \frac{G \Psi G^*}{G^* G} \tag{32}$$

Notice that M_k and N_k generated by the recursion (31)-(32) are all positive definite. Also notice that if $N_k = I$, then $M_{k+1} = M_k$ and $N_{k+1} = N_k$, namely we have a fixed point of the dynamical system. Moreover, if $N_k = \alpha I, \alpha \in$

$\mathbb{R}_+$ is a scalar matrix, then $N_{k+1} = I$. Finally, notice that, since $M \leq I$, the N_K have the uniform lower bound

$$N_k \geq \int \frac{G\Psi G^*}{G^*G}.$$

Let us denote by

$$\sigma_k^- = \sigma_{\min}(N_k), \quad \sigma_k^+ = \sigma_{\max}(N_k),$$

the minimum and maximum eigenvalues of N_k. We then have the following result.

Proposition 6.3:

$$\frac{1}{\sigma_k^+} N_k \leq N_{k+1} \leq \frac{1}{\sigma_k^-} N_k, \forall k \geq 0. \tag{33}$$

Proof: Since $\sigma_k^- I \leq N_k$, we get from (32)

$$N_{k+1} \leq \frac{1}{\sigma_k^-} \int \frac{G\Psi G^*}{G^* M_k G} = \frac{1}{\sigma_k^-} N_k.$$

Similarly for the other inequality. Q.E.D.

Corollary 6.4: For all $k \geq 1$, we have

$$\sigma_k^- \ \leq \ 1 \ \leq \ \sigma_k^+; \tag{34}$$

$$\frac{\sigma_k^-}{\sigma_k^+} \leq \sigma_{k+1}^- \leq \sigma_{k+1}^+ \leq \frac{\sigma_k^+}{\sigma_k^-} \tag{35}$$

Let $M \in \bar{\mathcal{M}} \backslash \bar{\mathcal{M}}_0$ and assume that M is not a fixed point of Θ. We conjecture (but cannot quite prove) that the corresponding sequence $\{N_k\}$ given by (32) converges to the identity matrix thereby producing the optimal solution of the approximation problem.

VII. SIMULATION RESULTS

In all three examples below, we have iterated the algorithm (28) with the stopping condition $\|M - \Theta(M)\| < 10^{-5}$.

Consider first the following instance of Example 1. Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The (estimated) finite covariance sequence is $c_0 = 4$, $c_1 = -1$, $c_2 = 2$, so that

$$\Sigma = \begin{bmatrix} 4 & -1 & 2 \\ -1 & 4 & -1 \\ 2 & -1 & 4 \end{bmatrix}$$

Also let

$$\Psi(z) = \frac{1.25}{(z + 1.5)(z^{-1} + 1.5)}$$

Notice that $\int \Psi = 1$. After the renormalization described at the beginning of Section V to transform Σ into the identity, we get

$$A' = \begin{bmatrix} -0.1036 & 1.1023 & -0.0351 \\ 0.2515 & -0.0220 & 0.9837 \\ 0.0571 & -0.2935 & 0.1256 \end{bmatrix}, \quad B' = \begin{bmatrix} -0.1431 \\ 0.0484 \\ 0.5640 \end{bmatrix}$$

The algorithm converged in 11 steps and the optimal matrix $M^o = \hat{\Lambda}'$ is given by

$$M^o = \begin{bmatrix} 0.3225 & -0.2534 & 0.0924 \\ -0.2534 & 0.3549 & -0.2534 \\ 0.0924 & -0.2534 & 0.3225 \end{bmatrix}$$

The optimal solution Φ^o may now be computed through (23). The function $\Phi^o(e^{i\omega})$ is depicted together with $\Psi(e^{i\omega})$ (bold line) in Figure 1.

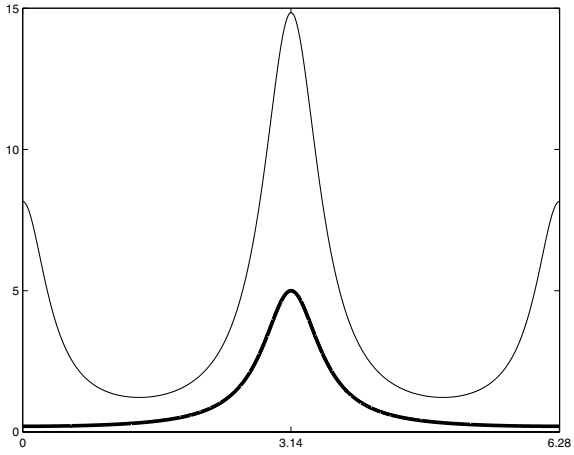

Fig. 1. Graphics of $\Psi(e^{i\omega})$ (bold line) and $\Phi^o(e^{i\omega})$ as functions of ω.

Consider next one instance of Example 2. Let

$$A = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

and

$$\Sigma = \begin{bmatrix} 4/3 & 6/5 & 8/7 \\ 6/5 & 9/8 & 12/11 \\ 8/7 & 12/11 & 16/15 \end{bmatrix}$$

Finally let

$$\Psi(z) = \frac{624}{(z + 25)(z^{-1} + 25)},$$

and notice that $\int \Psi = 1$. After the usual renormalization, we get

$$A' = \begin{bmatrix} 0.8842 & 0.2778 & 0.1660 \\ -0.4550 & 0.5423 & 0.4964 \\ -0.0837 & -0.4929 & -0.3431 \end{bmatrix}, \quad B' = \begin{bmatrix} 0.3368 \\ 0.5025 \\ 0.7952 \end{bmatrix}$$

The algorithm converged in 15 steps and the optimal matrix $M^o = \hat{\Lambda}'$ is given by

$$M^o = \begin{bmatrix} 0.1233 & 0.0193 & -0.0023 \\ 0.0193 & 0.1982 & -0.0059 \\ -0.0023 & -0.0059 & 0.6784 \end{bmatrix}$$

The optimal solution $\Phi^o(e^{i\omega})$ is depicted together with $\Psi(e^{i\omega})$ (bold line) in Figure 2.

Finally, we consider one case of Example 3. Let

$$A = \begin{bmatrix} 1/2 & 1 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 1/2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

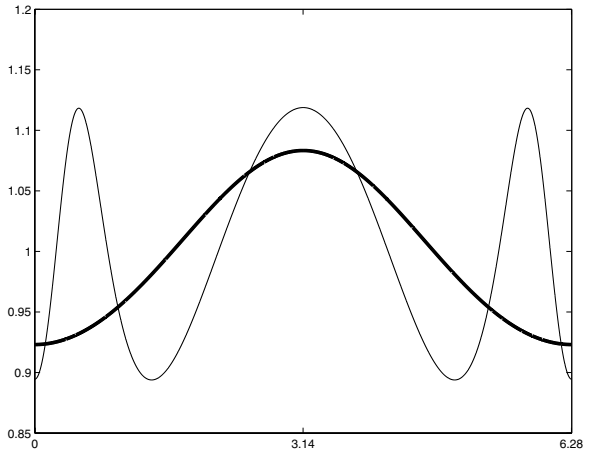

Fig. 2. Graphics of $\Psi(e^{i\omega})$ (bold line) and $\Phi^o(e^{i\omega})$ as functions of ω.

and

$$\Sigma = \begin{bmatrix} 19.3086 & 8.1111 & 1.5741 \\ 8.1111 & 6.3704 & 2.1111 \\ 1.5741 & 2.1111 & 2.6667 \end{bmatrix}$$

Let

$$\Psi(z) = \frac{624}{(z+25)(z^{-1}+25)},$$

where, as usual, $\int \Psi = 1$. After the usual renormalization, we get

$$A' = \begin{bmatrix} 0.8418 & 0.5227 & -0.1292 \\ -0.1325 & 0.4415 & 0.8741 \\ 0.0097 & -0.0340 & 0.2167 \end{bmatrix}, \quad B' = \begin{bmatrix} 0.0320 \\ -0.1949 \\ 0.7126 \end{bmatrix}$$

The algorithm converged in 32 steps and the optimal matrix $M^o = \hat{\Lambda}'$ is given by

$$M^o = \begin{bmatrix} 0.0645 & -0.1494 & 0.1817 \\ -0.1494 & 0.3748 & -0.4530 \\ 0.1817 & -0.4530 & 0.5608 \end{bmatrix}$$

The optimal solution $\Phi^o(e^{i\omega})$ is depicted together with $\Psi(e^{i\omega})$ (bold line) in Figure 3.

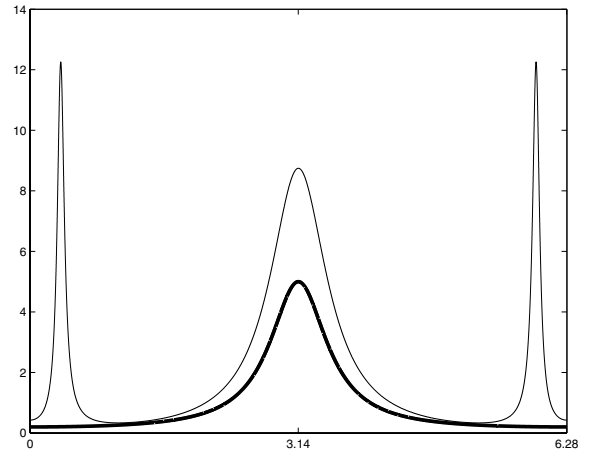

Fig. 3. Graphics of $\Psi(e^{i\omega})$ (bold line) and $\Phi^o(e^{i\omega})$ as functions of ω.

We have tested the algorithm, with similar results, also in many other cases of Examples 1-3 (see also [2]). The algorithm never failed to rapidly converge to a positive definite M which, by Proposition 5.2, provides an optimal solution to the original problem. This happened even when the algorithm was initialized with an M_0 very close to one of the one-dimensional projections that are fixed points of the map Θ (Proposition 6.1). The latter, as observed before, that do not, in general, yield the optimal solution $\hat{\Phi}$ via (18).

REFERENCES

[1] T. Georgiou and A. Lindquist, Kullback-Leibler approximation of spectral density functions, *IEEE Trans. Inform. Theory*, vol.49, pp. 2910-2917, 2003.

[2] M. Pavon and A. Ferrante, On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities, Sept. 2005, submitted for publication. Also (online) at http://www.dei.unipd.it/ augusto/preprints/SpectApp-preprint.pdf.

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

[4] C. I. Byrnes, T. Georgiou, and A. Lindquist, A new approach to spectral estimation: A tunable high-resolution spectral estimator, *IEEE Trans. Signal Processing* vol. 49, pp. 3189-3205, 2000.

[5] C. I. Byrnes, S. Gusev, and A. Lindquist, A convex optimization approach to the rational covariance extension problem, *SIAM J. Control and Opimization*, vol. 37, pp. 211-229, 1999.

[6] C. I. Byrnes, S. Gusev, and A. Lindquist, From finite covariance windows to modeling filters: A convex optimization approach, *SIAM Review*, vol. 43, pp. 645-675, 2001.

[7] 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, B. Hu and J Rosenthal (editors), Birkh"user Boston, pp. 81-108, 1997.

[8] 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. Automat. Contr.* vol. 40 , pp. 1841-1857, 1995.

[9] C. I. Byrnes and A. Lindquist, Interior point solutions of variational problems and global inverse function theorems, submitted for publication, Tech. Report TRITA/MAT-01-OS13, 2001. Also (online) at http://www.math.kth.se/alq/

[10] T. Georgiou, Realization of power spectra from partial covariance sequences, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 438-449, 1987.

[11] T. Georgiou, A topological approach to Nevanlinna-Pick interpolation, *SIAM J. on Mathematical Analysis*, vol. 18, pp. 1248-1260, 1987.

[12] T. Georgiou, The interpolation problem with a degree constraint, *IEEE Trans. on Automatic Control*, vol. 44, pp. 631-635, 1999.

[13] T. Georgiou, Spectral estimation by selective harmonic amplification, *IEEE Trans. Automat. Contr.*, vol.46, pp.29-42, 2001.

[14] T. Georgiou, The structure of state covariances and its relation to the power spectrum of the input, *IEEE Trans. Automat. Contr.*, vol.47, pp. 1056-1066, 2002.

[15] T. Georgiou, Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parameterization, *IEEE Trans. on Automatic Control*, vol. 47, pp.1811-1823, 2002.

[16] P. Enquist, A homotopy approach to rational covariance extension with degree constraint, *Int. J. Appl. Math. and Comp. Sci.*, vol. 11, pp.1173-1201, 2001.

[17] R. Nagamune, A robust solver using a continuation method for Nevanlinna-Pick interpolation with degree constraint, *Ieee Trans. Aut. Contr.*, vol 48, pp. 113-117, 2003.

[18] S. Haykin, *Nonlinear Methods of Spectral Analysis*, Springer-Verlag, 1979.