Some New Methods and Concepts in High-resolution Spectral Estimation

Plenary Lecture at the Third Asian Control Conference Shanghai, China, July 4–7, 2000

by

Anders Lindquist
Royal Institute of Technology
100 44 Stockholm, Sweden

1 Introduction

This lecture is a survey of some recent results in spectral estimation, and it is based on joint work together with C. I. Byrnes, P. Enqvist, T. T. Georgiou, S. V. Gusev, and others. We shall generalize traditional LPC filter design in several directions to obtain higher spectral resolution. This is done in the context of new paradigms for spectral estimation, based on analytic interpolation.

We begin by reviewing some basic concepts.

Shaping filters
It is common to model a signal $\{y(t) \mid t \in \mathbb{Z}\}$ as a convolution

$$y(t) = \sum_{k=-\infty}^{t} w_{t-k} u_k$$

of some excitation signal $\{u(t) \mid t \in \mathbb{Z}\}$. In the language of systems and control, this amounts to passing the excitation signal u through a linear filter with the transfer function

$$w(z) = \sum_{k=0}^{\infty} w_k z^{-k},$$

which is assumed to be rational, thus obtaining the signal y as the output, as depicted in Figure 1.

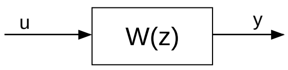

Figure 1: Representing a signal as the output of a black box.

More specifically, we take $w(z)$ to be rational with $w_0 \neq 0$ and all zeros and poles in the open unit disc. In other words, the system in Figure 1 is stable and minimum-phase. Such a filter will be called a *shaping filter*.

Consequently, the signal y is modeled by a choice of shaping filter and a choice of excitation signal u.

Let us begin by considering signals y for which the excitation signal u is white noise, i.e., $\mathrm{E}\{u(t)u(s)\} = \delta_{ts}$, where δ_{ts} is one if $t = s$ and zero otherwise. Then y is a stationary stochastic process with rational spectral density

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

which is positive for all θ. It is well-known that the spectral density has a Fourier expansion

$$\Phi(e^{i\theta}) = r_0 + 2\sum_{k=1}^{\infty} r_k \cos k\theta,$$

where the Fourier coefficients

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

are the covariance lags $r_k = \mathrm{E}\{y(t+k)y(t)\}$.

The positive real part of Φ
The spectral density $\Phi(z)$ is analytic in an annulus containing the unit circle and has there the representation

$$\Phi(z) = f(z) + f(z^{-1}),$$

where f is a rational function with all its poles and zeros in the open unit disc. Hence, in particular, f is analytic outside the unit disc, and

$$f(z) = \frac{1}{2}r_0 + r_1 z^{-1} + r_2 z^{-2} + r_3 z^{-3} + \dots$$

Moreover,

$$\Phi(e^{i\theta}) = 2\mathrm{Re}\{f(e^{i\theta})\} > 0,$$

for all θ, and therefore f is a real function which maps $\{|z| \geq 0\}$ into the right half-plane $\{\mathrm{Re}\, z > 0\}$; such a function is called *positive real.* For this to hold, the Toeplitz matrices

$$T_k = \begin{bmatrix} r_0 & r_1 & \cdots & r_k \\ r_1 & r_0 & \cdots & r_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_k & r_{k-1} & \cdots & r_0 \end{bmatrix} \tag{2}$$

must be positive definite for $k = 0, 1, 2, \ldots$.

Modeling from data

We would like to model the process y as the output of a shaping filter (Figure 1), when the available information about y is a finite record of observed data

$$y_0, y_1, \ldots, y_N. \tag{3}$$

As an example, to which we shall return several times in this lecture, let us consider a 30 ms frame of speech from the voiced nasal phoneme [ng], depicted in Figure 2. Here $N = 250$, a typical sample length for a mobile telephone.

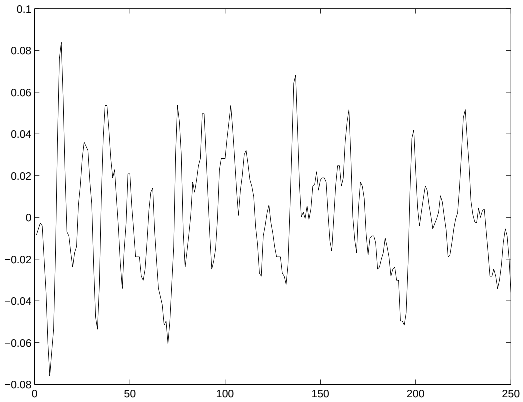

Figure 2: A frame of speech for the voiced nasal phoneme [ng].

Figure 3 depicts a periodogram of this signal, i.e., a spectral estimate obtained by fast Fourier transform. This spectral estimate can be modeled as a smooth spectral envelope perturbed by contributions from an excitation signal. The spectral envelope corresponds to the shaping of the vocal tract, which is described by the shaping filter.

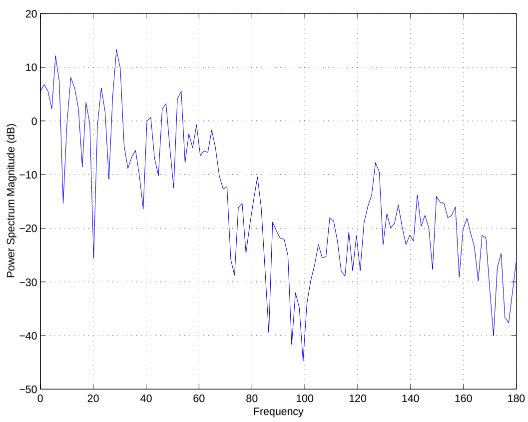

Figure 3: Periodogram for the voiced nasal phoneme [ng].

The contributions of the shaping filter and the excitation signal to the spectral estimate are multiplicative. If we consider the logarithm of the spectral density Φ, the *cepstrum*, instead of the Φ itself, the contribution of the excitation signal is superimposed on the that of the shaping filter. The Fourier coefficients

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \log \Phi(e^{i\theta}) d\theta \tag{4}$$

of the cepstrum

$$\log \Phi(e^{i\theta}) = c_0 + 2 \sum_{k=1}^{\infty} c_k \cos k\theta$$

are called the *cepstral coefficients* and can be estimated from data.

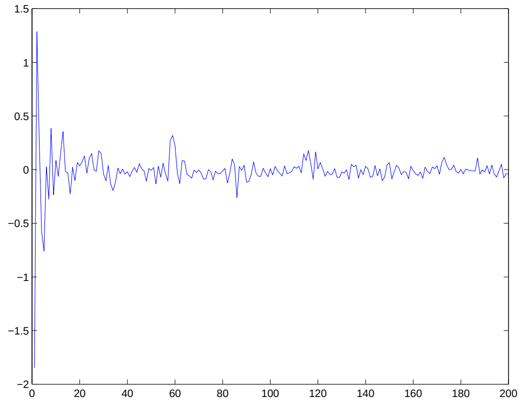

Figure 4: Cepstrum of voice speech signal.

Figure 4 shows the estimated cepstral coefficients of a frame of voiced speech. A contribution of the excitation signal is seen as spikes at multiples of the pitch period, corresponding to approximately $n_0 = 57$ in Figure 4. The spectral envelope can be estimated from a finite window

$$c_0, c_1, \ldots, c_n \tag{5}$$

of cepstral coefficients, where $n < n_0$.

First, however, we shall consider the problem of estimating the shaping filter from covariance data.

Estimating shaping filters from covariance estimates

Given a finite observed record (3) of the process y, a limited number of covariance lags $r_k := \mathrm{E}\{y(t+k)y(t)\}$ can be estimated via some ergodic estimate

$$r_k = \frac{1}{N+1-n} \sum_{t=0}^{N-n} y_{t+k} y_k. \tag{6}$$

However, we can only estimate

$$r_0, r_1, \ldots, r_n, \tag{7}$$

where $n << N$, with some precision. Therefore, in order to determine Φ, we must determine

$$r_{n+1}, r_{n+2}, r_{n+3}, \ldots$$

subject to the condition that f is positive real. This is the *covariance extension problem* or the *Carathéodory problem.* In order for this problem to have a solution we must have $T_n > 0$, a condition which is guaranteed by the ergodic estimates given above. Then the covariance extension problem has infinitely many (meromorphic) solutions [33].

However, we are only interested in rational solutions of low order. It can be shown [5, 20] that the smallest degree that can be guaranteed is n, so we shall be interested in f of degree at most n. Given such an f, there is a unique stable minimum-phase rational function w, of degree at most n, such that

$$|w(e^{i\theta})|^2 = \Phi(e^{i\theta}) := 2\text{Re}\{f(e^{i\theta})\}. \quad (8)$$

We shall call such a w a modeling filter for (7). More precisely, noting (1), a *modeling filter* for $r_0, r_1, \ldots, r_n$ is a rational, stable, minimum-phase function w such that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}|w(e^{i\theta})|^2 d\theta = r_k, \quad k = 0, 1, \ldots, n. \quad (9)$$

Clearly, this is the type of shaping filter we are looking for.

2 Linear predictive filtering

The most popular modeling filter is obtained by solving the *normal equations*

$$\begin{bmatrix} r_0 & r_1 & \cdots & r_{n-1} \\ r_1 & r_0 & \cdots & r_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n-1} & r_{n-2} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} \varphi_{nn} \\ \varphi_{n,n-1} \\ \vdots \\ \varphi_{n1} \end{bmatrix} = \begin{bmatrix} r_n \\ r_{n-1} \\ \vdots \\ r_1 \end{bmatrix}$$

for the unique solution $\varphi_{n1}, \varphi_{n2}, \ldots, \varphi_{nn}$ (recall that $T_{n-1} > 0$) and then forming

$$\rho_n = \sum_{j=0}^{n} r_j \varphi_{nj}.$$

The polynomial

$$\varphi_n(z) = z^n + \varphi_{n1} z^{n-1} + \ldots + \varphi_{nn}$$

is the n:th *Szegö polynomial* (of the first kind), and it can be determined via the *Levinson algorithm*

$$\varphi_{k+1}(z) = z\varphi_k(z) - \gamma_k z^k \varphi_k(z^{-1}), \quad \varphi_0(z) = 1, \quad (10)$$

where the *Schur parameters* $\{\gamma_k\}$ are given by

$$\gamma_k = \frac{1}{\rho_k}\sum_{j=0}^{k} \varphi_{k,k-j} r_{j+1}.$$

In signal processing, the Schur parameters are often called the *PARCOR parameters.* The condition $T_n > 0$ is equivalent to the condition that $|\gamma_k| < 1$ for $k = 0, 1, \ldots, n-1$.

Then, defining the *Szegö polynomials of the second kind*, $\{\psi_k\}$, via

$$\psi_{k+1}(z) = z\psi_k(z) + \gamma_k z^k \psi_k(z^{-1}), \quad \psi_0(z) = 1,$$

the rational function

$$f(z) = r_0 \frac{\psi_n(z)}{\varphi_n(z)}$$

is a solution of degree n to the covariance extension problem. The corresponding modeling filter

$$w(z) = \frac{\sqrt{\rho_n} z^n}{\varphi_n(z)}$$

is called the *linear predictive filter* or *LPC filter* and is standard in, for example, speech processing [31]. However, since the corresponding spectral density

$$\Phi(e^{i\theta}) = \frac{\rho_n}{|\varphi_n(e^{i\theta})|^2}$$

has no zeros, this method has some problems picking up the valleys in the spectrum, thus producing a "flat speech" in speech processing. Often this is compensated by choosing an appropriate excitation signal from a code book.

Let us now return to the speech data in Figure 2. Figure 5 depicts the periodogram of Figure 3 together with the spectral envelope determined by a tenth order LPC filter, based on ergodic estimates of $r_0, r_1, \ldots, r_{10}$ from the data in Figure 2.

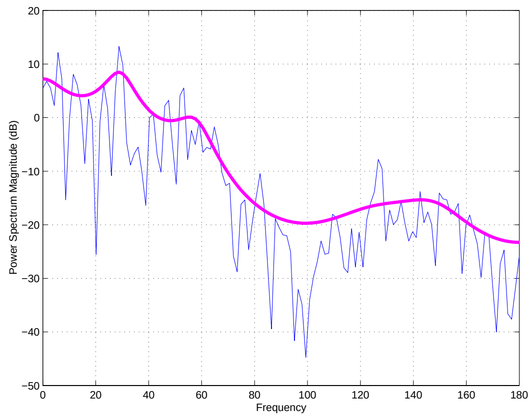

Figure 5: Spectral envelope of 10th order LPC filter.

It is seen that this estimate of the spectral envelope does not reproduce the notches of the spectrum very well. This is due to the fact that the zeros of the modeling filter are as far away as possible from the unit circle. This is one of the shortcomings of LPC filtering which we shall want to rectify.

3 Modeling filters with arbitrary zeros

The LPC solution is one of infinitely many solutions of the covariance extension problem. In the case that the positive real function f is not required to be rational, the family of solutions was completely parameterized by Schur [33]. In fact, each choice of free Schur parameters $\gamma_n, \gamma_{n+1}, \gamma_{n+2}, \ldots$ such that

$$|\gamma_k| < 1$$

corresponds to one and only one covariance extension, and all covariance extensions are obtained in this way. This extension is given by

$$r_{k+1} = \rho_k \gamma_k - \sum_{j=0}^{k-1} \varphi_{k,k-j} r_{j+1}, \quad k = n, n+1, \ldots,$$

where $\rho_{k+1} = (1 - \gamma_k^2)\rho_k$ and the coefficients of the Szegö polynomials are given by the Levinson-Szegö recursion (10). In particular, the LPC solution is the solution obtained by choosing

$$\gamma_k = 0, \quad k = n, n+1, \ldots.$$

It is also the solution which maximizes the the entropy gain

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \Phi(e^{i\theta}) d\theta,$$

which is precisely the zeroth cepstral coefficient. Therefore the LPC solutions is also called the *maximum-entropy solution*.

Parameterization of all modeling filters

However, if we want to parameterize the subset of all solutions to the covariance extension problem of degree at most n, or any other maximal degree for that matter, there is no way to do this in terms of the Schur parameters. Instead, we have the following result. Recall that a real polynomial is said to be *stable* if it has all its roots in the open unit disc $\{|z| < 1\}$.

Theorem 1 *Let $r_0, r_1, \ldots, r_n$ be a partial covariance sequence, i.e., real numbers such that the Toeplitz matrix (2) is positive definite. Then, to any stable polynomial*

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \ldots + \sigma_{n-1} z + \sigma_n$$

of degree n, there corresponds a unique pair of real stable polynomials

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

of degree n such that

$$a(z)b(z^{-1}) + b(z)a(z^{-1}) = \sigma(z)\sigma(z^{-1}), \tag{11}$$

and the rational function

$$f(z) := \frac{a(z)}{b(z)} \tag{12}$$

is a solution to the covariance extension problem, i.e., is positive real and satisfies

$$f(z) = \frac{1}{2} r_0 + r_1 z^{-1} + \ldots + r_n z^{-n} + \ldots \tag{13}$$

Georgiou, who had proven the existence part of Theorem 1 in [20], conjectured in [21] that uniqueness would also hold. This conjecture remained open until 1993, when in [4] we proved a stronger version of Theorem 1.

From (11) and (12) we see that

$$f(z) + f(z^{-1}) = w(z)w(z^{-1}),$$

where

$$w(z) = \frac{\sigma(z)}{a(z)}. \tag{14}$$

Consequently, given the partial covariance sequence (7), Theorem 1 states that to each choice of zero polynomial $\sigma(z)$ there is a unique pole polynomial $a(z)$ such that (14) is a modeling filter for (7). Hence, we have a complete parameterization of all modeling filters in terms of zeros.

The geometry of positive real functions

Theorem 1 is a corollary of a more fundamental result [4,11] about the space $\mathcal{P}_n$ of all positive real functions of degree at most n, which is a manifold of dimension $2n$. For $n = 1$, and a suitable choice of coordinates, this manifold can be represented as the interior of the diamond depicted in Figure 6, and it is divided into disjoint submanifolds (hyperbolas in the figure).

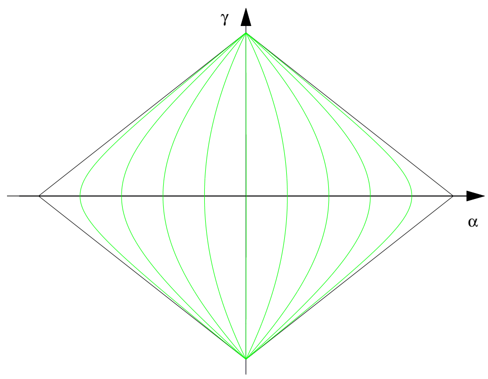

Figure 6: Filtering foliation of $\mathcal{P}_1$.

Such a decomposition is called a *foliation* in differential geometry, and the submanifolds are called *leaves*, provided certain smoothness conditions are satisfied, as is the case here. We call this foliation the *filtering foliation*, since the leaves are the stable manifolds of a certain fast algorithm for Kalman filtering [15, 28, 29]. There is precisely one leaf for each choice of zero polynomial $\sigma(z)$. An analogous foliation result holds for any n.

There is also another foliation of $\mathcal{P}_n$ with one leaf for each choice of covariance data (7). For the case $n = 1$, it is depicted in Figure 7 (with horizontal lines). The leaves of this foliation intersect the leaves of the filtering foliation transversely (under nonzero angle) so that each leaf of one foliation intersects each leaf of the other in one, and only one, point in $\mathcal{P}_n$; see [4, 11]. This point determines the unique pole polynomial $a(z)$ in Theorem 1. It also insures that the bijection of Theorem 1 is a diffeomorphism.

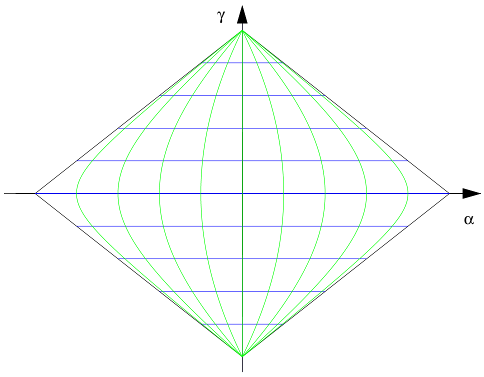

Figure 7: Filtering and covariance matching foliations of $\mathcal{P}_1$.

One might ask whether it is possible to prescribe the poles instead of the zeros, but already the case $n = 1$ provides a counterexample as seen in Figure 8. In fact, choosing an arbitrary leaf from the covariance matching foliation and an arbitrary leaf from the pole foliation (diagonal lines), they do not necessarily intersect.

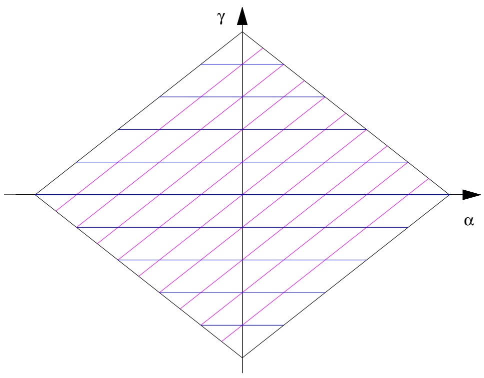

Figure 8: Pole and covariance matching foliations of $\mathcal{P}_1$.

A convex optimization approach to modeling-filter design

The proof of Theorem 1 in [4], as well as the existence proof in [20, 21], was nonconstructive. A convex optimization algorithm was given in [8]. Next, following [9], we shall introduce this algorithm as the dual of the problem to maximize a linear combination of coefficients in a cepstral window, subject to covariance matching.

Recall that the LPC solution is the modeling filter which maximizes the zeroth cepstral coefficient. Suppose that we maximize instead a linear combination

$$p_0 c_0 + p_1 c_1 + \ldots + p_n c_n \tag{15}$$

of the cepstral coefficients in the window (5). In view of (4), this may be written as a generalized entropy gain

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

where P is the symmetric pseudopolynomial

$$P(z) = p_0 + \tfrac{1}{2} p_1 (z + z^{-1}) + \ldots + \tfrac{1}{2} p_n (z^n + z^{-n}). \tag{16}$$

We shall say that $P \in \mathcal{D}$ if P is nonnegative on the unit circle and $P \in \mathcal{D}_+$ if it is positive there. We note that the covariance matching condition (9) becomes

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\theta = r_k, \quad k = 0, 1, \ldots, n \tag{17}$$

in terms of $\Phi(e^{i\theta}) = |w(e^{i\theta})|^2$.

We state the following theorem from [9].

Theorem 2 *The problem to maximize (15) subject to (17) has a finite solution only if the pseudo-polynomial (16) belongs to $\mathcal{D}$. If $P \in \mathcal{D}_+$, there is a unique solution Φ, and this solution has the form*

$$\Phi(z) = \frac{P(z)}{Q(z)}, \tag{18}$$

where

$$Q(z) = q_0 + \tfrac{1}{2} q_1 (z + z^{-1}) + \ldots + \tfrac{1}{2} q_n (z^n + z^{-n})$$

belongs to $\mathcal{D}_+$.

Connecting Theorems 1 and 2, we see that if we take P to be

$$P(z) = \sigma(z)\sigma(z^{-1}),$$

then the unique $a(z)$ in Theorem 1 is the stable polynomial satisfying

$$Q(z) = a(z)a(z^{-1}).$$

Consequently, if we can determine $Q(z)$ in Theorem 2, we have also determined $a(z)$ in Theorem 1, and hence the modeling filter corresponding to $\sigma(z)$.

As we shall see below, it turns out that the algorithm needed to determine Q is precisely the convex optimization algorithm presented in [8]. In fact, as shown in [9], the dual problem, in the sense of mathematical programming, of the maximization problem of Theorem 2 is the problem to minimize

$$\mathbb{J}_P(q_0, q_1, \ldots, q_n) = r_0 q_0 + r_1 q_1 + \ldots + r_n q_n - \frac{1}{2\pi}\int_{-\pi}^{\pi} P(e^{i\theta}) \log Q(e^{i\theta}) d\theta \quad (19)$$

over all $Q \in \mathcal{D}$. The functional $\mathbb{J}_P$ is strictly convex. The following theorem was proven in [8].

Theorem 3 *The problem*

$$\min_{Q \in \mathcal{D}} \mathbb{J}_P(Q)$$

has a unique solution, and it belongs to $\mathcal{D}_+$.

Since $\mathbb{J}_P$ takes its minimum in an interior point,

$$\frac{\partial \mathbb{J}_P}{\partial q_k} = r_k - \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta} \frac{P(e^{i\theta})}{Q(e^{i\theta})} d\theta$$

equals zero there. This stationarity condition is precisely the covariance matching condition. The dual problem is easily solved by Newton's method [8,9]. We shall call a modeling filter obtained from this procedure an *Ladder-Lattice-Notch (LNN) filter*, since it can be implemented by a ladder-lattice filter [9] and represents the notches in the spectrum better.

Examples

Let us now return to the voice sample in Figure 2. The periodogram of this voiced nasal phoneme [ng] is depicted in Figure 3, and the spectral envelope of a 10th order LPC filter in Figure 5.

In order to compute an LLN filter we first need to estimate the zero polynomial $\sigma(z)$, or equivalently $P(z)$. Since this estimate does not have to be very accurate – it is anyway better than choosing all zeros at the origin as the LPC filter – we may use one of several *ad hoc* methods, which are not sufficiently accurate for determining the filter itself. In the present example, using points on a smoothed periodogram $\hat{\Phi}$, we find the P and Q which minimize

$$\max_k |Q(e^{i\theta_k})\hat{\Phi}(e^{i\theta_k}) - P(e^{i\theta_k})|$$

for equidistant points on the unit circle. We use this P to form $\mathbb{J}_P$ and the Q as an initial condition in the convex minimization problem of Theorem 3 to compute the optimal Q corresponding to P.

In Figure 9 we show the spectral envelope of a sixth order LLN filter, which should be compared with that of the 10th order LPC filter in Figure 5. Clearly the LLN filter reproduces the notches much better although it is lower order. If we take a 10th order LLN filter, we can pick up more of the fine structure, as seen in Figure 10.

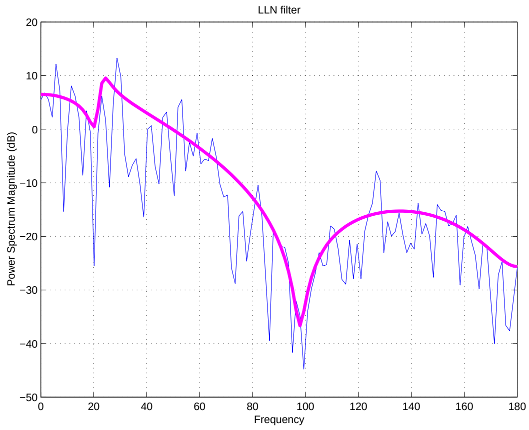

Figure 9: Spectral envelope of a 6th order LLN filter.

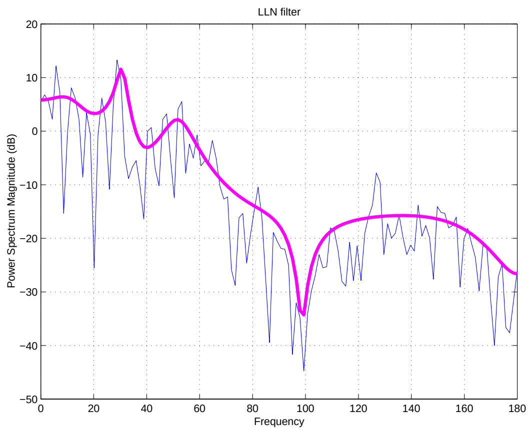

Figure 10: Spectral envelope of 10th order LLN filter.

4 Simultaneous covariance and cepstral matching

We have shown that modeling filters, i.e., stable, minimum-phase filters, which match a window of covariance lags of length $n+1$, where n is an upper bound for the degree of the filter, can be completely parameterized in terms of its zeros, but not in terms of its poles. It would be interesting to find out whether it is possible to parameterize modeling filters in terms of a window of cepstral coefficients, which like the covariance lags can be estimated from data. At a first glance, this might be suggested by the following theorem, which we quote from [9].

Theorem 4 *Each modeling filter (14) of degree n determines and is uniquely determined by its window $r_0, r_1, \ldots, r_n$ of covariance lags and its window $c_1, c_2, \ldots, c_n$ of cepstral coefficients.*

In other words, the very nontrivial statement of this theorem is that there is a one-to-one correspondence between the $2n + 1$ coefficients $r_0, r_1, \ldots, r_n, c_1, c_2, \ldots, c_n$ of the modeling filter (14) and the $2n+1$ coefficients $a_0, a_1, \ldots, a_n, \sigma_1, \sigma_2, \ldots, \sigma_n$ of the denominator and numerator polynomials of (14), provided the degree of w is *exactly* n. In fact, $r_0, r_1, \ldots, r_n, c_1, c_2, \ldots, c_n$ form local coordinates for the space of pole-zero filters of degree n. The proof is given in [10].

To this end, given observed records of $r_0, r_1, \ldots, r_n$ and $c_1, c_2, \ldots, c_n$, consider the problem of finding a spectral density Φ which minimizes the "cepstral error"

$$\sum_{k=0}^{n} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \log \Phi(e^{i\theta}) d\theta - c_k \right|$$

subject to the covariance-lag matching

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\theta = r_k, \quad k = 0, 1, \ldots, n.$$

As seen from (8), finding a minimizing Φ is equivalent to finding a minimizing modeling filter w or a positive real function f.

As shown in [10] (also see, [19]), the dual problem (in the sense of mathematical programming) is to maximize the concave, but not strictly concave, functional

$$\begin{aligned} J(p,q) &= c_1 p_1 + c_2 p_2 + \ldots + c_n p_n \\ &\quad - r_0 q_0 - r_1 q_1 - \ldots - r_n q_n \\ &\quad - \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{ik\theta}) \log \frac{P(e^{ik\theta})}{Q(e^{ik\theta})} d\theta \end{aligned}$$

over all $P \in \mathcal{D}$ such that $p_0 = 1$ and all $Q \in \mathcal{D}$. It can be shown [10] that there is a, possibly nonunique, maximizing solution such that $Q \in \mathcal{D}_+$, whereas we cannot in general insure that $P \in \mathcal{D}_+$.

Now,

$$\begin{aligned} \frac{\partial J}{\partial q_k} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \frac{P(e^{i\theta})}{Q(e^{i\theta})} d\theta - r_k \\ \frac{\partial J}{\partial p_k} &= c_k - \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \log \frac{P(e^{i\theta})}{Q(e^{i\theta})} d\theta, \end{aligned}$$

and hence we always have covariance matching, while we may or may not have cepstral matching, depending on whether $P \in \mathcal{D}_+$ or not.

Figure 11 shows the periodogram of a frame of speech for the phoneme [s] together with a 10th order spectral envelope produced by this method. In this case, $P \in \mathcal{D}_+$, so there is both covariance and cepstral matching. In general, however, this is not the case, and then there may be a duality gap. This can be seen already in the case $n = 1$. In Figure 12 the covariance matching foliation (straight lines) is depicted together with the cepstral matching foliation (curved). As can be seen, a leaf in one foliation in general does not intersect all leaves in the other. Therefore, this problem, unlike that of Theorem 2, is not well-posed.

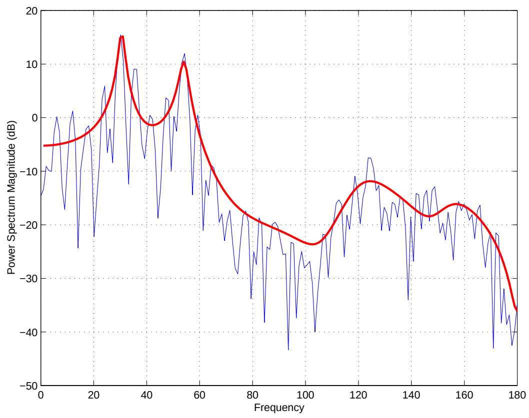

Figure 11: Spectral envelope of 10th order cepstral match filter.

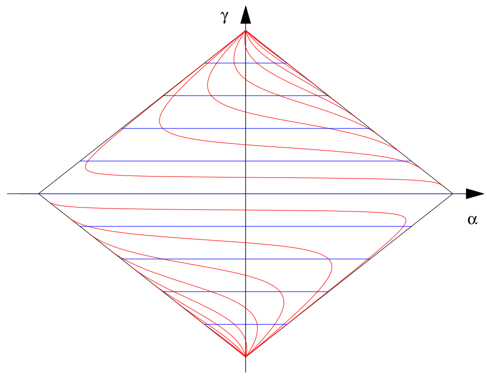

Figure 12: Cepstral and covariance matching foliations of $\mathcal{P}_1$.

5 A tunable high-resolution spectral estimator

So far we have been looking for shaping filters w with the covariance-lag matching property (9). We recall (8) to see that this is the same as looking for a positive real f such that

$$f(z) = \frac{1}{2} r_0 + r_1 z^{-1} + \ldots + r_n z^{-n} + \ldots,$$

or, equivalently, such that

$$\frac{f^k(\infty)}{k!} = r_k, \quad k = 0, 1, \ldots, n. \tag{20}$$

Hence covariance matching is equivalent to finding a positive real interpolant taking prescribed values at ∞. Consequently we have a special case of Nevanlinna-Pick interpolation. Recall that the stable zeros of $f(z) + f(z^{-1})$ are called the *spectral zeros* of f.

In this context, LPC filtering amounts finding an interpolant with all spectral zeros at the origin and all interpolation points at infinity. The points $z = 0$ and $z = \infty$ are the points, inside and outside the unit circle respectively, which are furthest away from the unit circle, where the spectrum is evaluated. This explains the shortcomings of the LPC filter.

The better resolution of the LNN filter comes from the fact that the spectral zeros have been moved from the origin to points in the open unit disc closer to the unit circle at frequencies where the notches in the spectrum occur. Still, however, the interpolation points are all at $z = \infty$.

Now, suppose we are interested in some particular frequency band, where we would like to have higher resolution. It would then make sense to replace the interpolation condition (20) by a condition

$$f(z_k) = w_k, \quad k = 0, 1, \ldots, n, \tag{21}$$

where $z_0 = \infty$, and where the distinct points $z_1, \ldots, z_n$ are chosen close to the unit circle in the selected frequency band, but still outside the closed unit disc, where f is analytic. We assume that the sets $\{z_0, z_1, \ldots, z_n\}$ and $\{w_0, w_1, \ldots, w_n\}$ are *self-conjugate*, i.e., the point are either real or occur in complex conjugate pairs. For notational reasons, to become clear later, we introduce the points

$$p_k := z_k^{-1}, \quad k = 0, 1, \ldots, n, \tag{22}$$

which then are all located in the open unit disc. In particular, $p_0 = 0$.

Nevanlinna-Pick interpolation with degree constraint

This leads to the following problem [11–13, 22, 23]. Given $n + 1$ points in the open right half-plane, $w_0, w_1, \ldots, w_n$, with the property that the *Pick matrix*

$$P_n = \left[\frac{w_k + \bar{w}_\ell}{1 - p_k \bar{p}_\ell}\right]_{k,\ell=0}^{n} \tag{23}$$

is positive definite, find a positive real, rational function f of degree at most n which satisfies the interpolation condition (21).

Define the polynomial

$$\tau(z) := \prod_{k=1}^{n} (z - p_k),$$

and let $\mathcal{S}$ be the class of all real, rational functions of the type

$$S(z) = \frac{\rho(z)\rho(z^{-1})}{\tau(z)\tau(z^{-1})},$$

where $\rho(z)$ is a stable polynomial of degree at most n. Moreover, let $\mathcal{S}_+$ be the subclass of all $S \in \mathcal{S}$ which are positive on the unit circle. Now, for any $P \in \mathcal{S}_+$, define the generalized entropy gain

$$\mathbb{I}_P(f) = \int_{-\pi}^{\pi} P(e^{i\theta}) \log(\text{Re}\{f(e^{i\theta})\}) d\theta \tag{24}$$

on the space $\mathcal{C}$ of positive real functions.

Then [12, Theorem 4.1] provides us with the following result.

Theorem 5 *Given any stable polynomial*

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \ldots + \sigma_{n-1} z + \sigma_n,$$

let $P \in \mathcal{S}_+$ be given by

$$P(z) = \frac{\sigma(z)\sigma(z^{-1})}{\tau(z)\tau(z^{-1})}. \tag{25}$$

Then, the constrained optimization problem to maximize $\mathbb{I}_P(f)$ over all $f \in \mathcal{C}$ subject to the interpolation condition

$$f(p_k^{-1}) = w_k, \quad k = 0, 1, \ldots, n, \tag{26}$$

has a unique solution. This solution is of the form

$$f(z) = \frac{b(z)}{a(z)}, \tag{27}$$

where $a(z)$ and $b(z)$ are polynomials of degree n such that

$$a(z)b(z^{-1}) + b(z)a(z^{-1}) = \sigma(z)\sigma(z^{-1}). \tag{28}$$

Conversely, if f is a positive-real function which satisfies the interpolation conditions as well as (27) and (28), then it is the unique solution to the optimization problem.

In particular, $\sigma(z) \equiv \tau(z)$ yields $P = 1$ so that $\mathbb{I}_P(f)$ becomes the usual entropy gain. There are fast algorithms to compute the corresponding solution, the *central solution*; see, e.g., [13].

In general, (28) yields

$$\Phi(e^{i\theta}) = 2\text{Re}\{f(e^{i\theta})\} = \frac{P(e^{i\theta})}{Q(e^{i\theta})},$$

where

$$Q(z) = \frac{a(z)a(z^{-1})}{\tau(z)\tau(z^{-1})}.$$

Consequently, the optimal shaping filter is

$$w(z) = \frac{\sigma(z)}{a(z)}.$$

To determine it, we need to find $a(z)$, or, equivalently, Q. However, this optimization problem is infinite-dimensional and therefore not easy to solve. As before, it has a dual with finitely many variables. Next, we turn to this problem.

To this end, let w be any real rational function which is analytic on and outside the unit circle and satisfies the interpolation condition

$$f(p_k^{-1}) = w_k, \quad k = 0, 1, \ldots, n. \tag{29}$$

Note that w need not be positive real, so it is simple to determine. For example, let w be a rational function of degree n with an arbitrary stable polynomial as denominator. Then the numerator polynomial satisfies a linear (Vandermonde) system of equations.

Now, for any $Q \in \mathcal{S}_+$ define the functional

$$\begin{aligned}\mathbb{J}_P(Q) \quad := \quad & \frac{1}{2\pi}\int_{-\pi}^{\pi} Q(e^{i\theta})\mathrm{Re}\{w(e^{i\theta})\}d\theta - \\ & \frac{1}{2\pi}\int_{-\pi}^{\pi} P(e^{i\theta})\log\{Q(e^{i\theta})\}d\theta.\end{aligned}$$

It can be shown that $\mathbb{J}_P$ does not depend on the particular choice of w but only on the values of w at the interpolation points.

Using duality theory, the maximization problem of Theorem 5 can be seen to be equivalent to the following convex optimization problem [12, Theorem 4.5].

Theorem 6 *The convex optimization problem*

$$\min_{Q \in \mathcal{S}} \mathbb{J}_P(Q) \tag{30}$$

has a unique solution, which belongs to $\mathcal{S}_+$.

To solve this convex optimization problem, one can use a Newton-type search algorithm. Expressions for the gradient and Hessian of $\mathbb{J}_P$, in suitable bases, are given in [13], where a numerical algorithm is described in detail.

Just as for the LNN filter, the bijection between the zero polynomial $\sigma(z)$ and the pole polynomial $a(z)$ is a diffeomorphism. In fact, as shown in [11], there is a foliation of the space $\mathcal{P}_n$ with one leaf for each choice of interpolation values $w_0, w_1, \ldots, w_n$, and the leaves of this *interpolation foliation* intersect the leaves of the filtering foliation, introduced above, transversely (under nonzero angle) so that each leaf of one foliation intersects each leaf of the other in one, and only one, point in $\mathcal{P}_n$. Figure 13 describes this situation in the case $n = 1$.

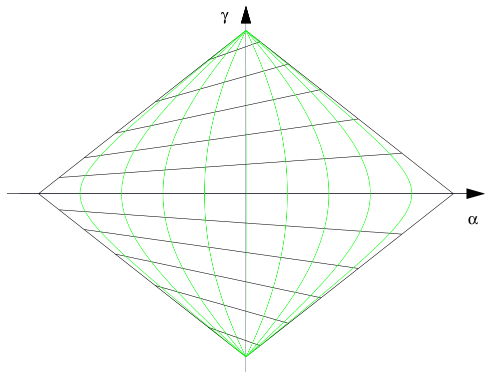

Figure 13: Interpolation and filtering foliations of $\mathcal{P}_1$.

A filter bank for estimating the interpolation values

It remains to describe out how to estimate the interpolation values $w_0, w_1, \ldots, w_n$ from the data

$$y_0, y_1, \ldots, y_N.$$

To this end, we follow the procedure in [13].

The key observation in [13], relating interpolating values to covariance statistics, can be explained as follows. Let

$$G(z) = \frac{z}{z-p}, \quad |p| < 1, \tag{31}$$

be a first-order filter

$$\xrightarrow{y} \boxed{G(z)} \xrightarrow{v}$$

driven by the process y. Then it can be shown that the output process v has the variance

$$\mathrm{E}\{v(t)^2\} = \frac{2}{1-p^2} f(p^{-1}).$$

Now, the variance $\mathrm{E}\{v(t)^2\}$ can be estimated from the output data

$$v_0, v_1, \ldots, v_N$$

by an ergodic sum of type (6), and hence we obtain the interpolation condition $f(p^{-1}) = w$, where

$$w := \frac{1-p^2}{2(N+1)} \sum_{t=0}^{N} v_k^2. \tag{32}$$

Next, consider a bank of first-order filters with filter poles $p_0, p_1, \ldots, p_n$, all driven by y, as depicted in Figure 14. More precisely, the filters are

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

Two first-order filters with complex conjugate poles could be combined to one real, second-order filter, but complex arithmetic could also be used.

The output data from each filter in the filter bank is processed as in (32) to yield the interpolation values $w_0, w_1, \ldots, w_n$.

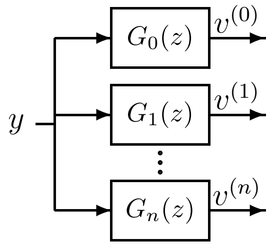

Figure 14: Filter bank.

Examples

We conclude with two examples from [13] where a suitable choice of filter bank poles results in significantly higher resolution than traditional techniques over the arc which is adjacent to the poles of the filter bank.

Example 1: Let us first consider a stochastic process y consisting of two sinusoids in colored noise. Figure 15 shows the spectrum, plotted on the unit circle, with the two spectral lines marked as arrows. A choice of filter bank poles in the vicinity of the spectral lines are marked by $\times$ in the same plot.

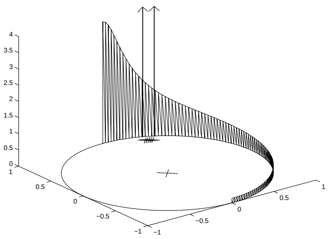

Figure 15: Spectrum plotted on unit circle.

Figure 16 compares the resolutions obtained by a periodogram based methods using state-of-the-art windowing technology (first row) and an LPC filter (second row) with that of our method [13] (third row), for which we use the simple default setting of the central solution (i.e., $P = 1$) and data obtained via a filter bank with poles at $0, \pm.85, 0.9e^{\pm.42i}, 0.9e^{\pm.44i}, 0.9e^{\pm.48i}, 0.9e^{\pm.50i}$. In these plots the exact location of the spectral lines is shown by vertical lines. In each case 300 data points were used and the order of the AR model (using LPC filtering) was chosen to be the same as that of the model resulting from our technique. In the left column, where five runs are superimposed to show the variability, the separation between the two spectral lines is 0.11, while in the second column the lines have been moved closer to a distance of 0.02, which is below the theoretical limit for the fast Fourier transform (periodogram). Thus the periodogram method, which does OK in the first case, fails in the second. LPC filtering fails in both cases. Finally, our method resolves the two peaks also in the case when they are close. Hence, we obtain a significant improvement in resolution by using a filter bank and analytic interpolation techniques.

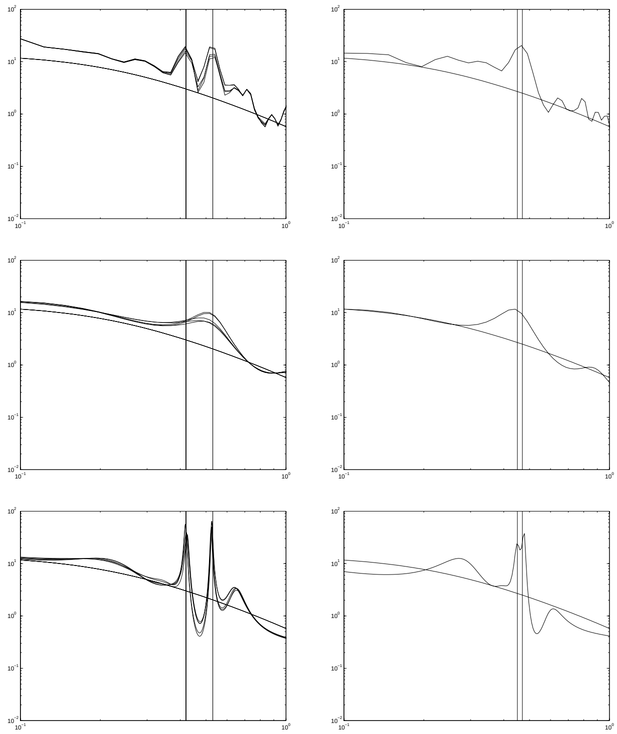

Figure 16: Estimation of spectral lines in colored noise.

Example 2: To demonstrate the power of the method when using nontrivial spectral zeros ($P \neq 1$) , let us next consider a spectrum with sharp transitions, as in Figure 17. This corresponds to a process obtained by passing white noise through a filter with a transfer function having poles and zeros close to each other. Such estimation problems are difficult, and we do better if we use the full power of our theory.

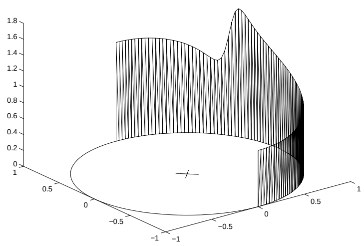

Figure 17: Spectrum plotted on unit circle.

In Figure 18 spectral estimates are depicted for two different processes of the type illustrated in Figure 17, the true spectra shown with dotted line. The top row shows periodogram estimates, the second row fourth order maximum entropy filter estimates, and the bottom row estimates obtained using our method with filter-bank poles set at $0, .8e^{\pm.8i}, .8e^{\pm1.3i}$ and spectral zeros at $0, -.8, .8e^{\pm i\pi/3.3}$, as depicted in Figure 19.

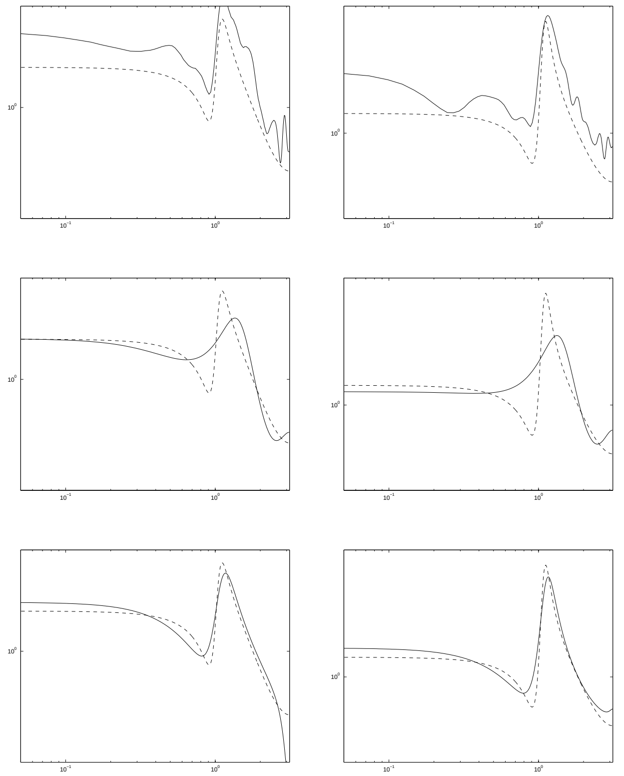

Figure 18: Estimating a spectrum with sharp transitions.

In this set of experiments we have used a data record long enough to eliminate fluctuations of the estimated spectra, namely 2000 samples. Noting that *the order is only four*, our estimates are remarkably good. For more discussion as well as further examples requiring the full power of the theory (i.e., specifying nontrivial spectral zeros) we refer to [13].

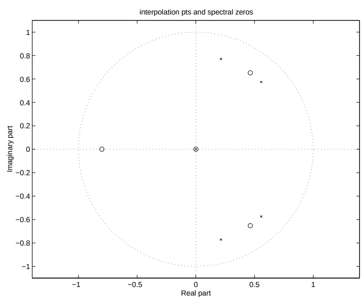

Figure 19: Spectral zeros and filter-bank poles.

6 Conclusions

One of the most widely used methods of spectral estimation in signal and speech processing is linear predictive coding (LPC). LPC has some attractive features, which account for its popularity, including the properties that the resulting shaping filter (i) matches a finite window of $n+1$ covariance lags, (ii) is rational of degree at most n, and (iii) has stable zeros and poles. A major disadvantage is that the modeling filter is "all-pole", i.e., an autoregressive (AR) model.

In this talk we reviewed some recent results in pole-zero modeling, in which the nice properties (i), (ii), (iii) of LPC filtering have been retained. The key observation is that covariance matching imposes interpolation conditions on the positive real part of the spectral density. This leads to the application of very recent mathematical results of analytic interpolation with degree constraint. In this context, LPC filtering amounts to choosing the spectral zeros at $z=0$ and the interpolation points at $z=\infty$. The points $z=0$ and $z=\infty$ are the points, inside and outside the unit circle respectively, which are furthest away from the unit circle, where the spectrum is evaluated, explaining the shortcomings of the LPC filter.

We started by moving the zeros closer to the unit circle at frequencies where notches occur, thus obtaining better resolution in these parts of the spectrum. We continued by also moving the interpolation points closer to the unit circle in a selected frequency band where higher resolution is required. This leads to a radically radically different approach to spectral estimation which is based on nontraditional covariance measurements. The basic idea is to determine covariance estimates after filtering the observed time series through a bank of filters with different frequency responses, which can be tuned as desired.

References

[1] C.I. Byrnes and A. Lindquist, *On the geometry of the Kimura-Georgiou parameterization of modeling filter*, Int. J. Contr. **(50)**: 2301–2312, 1989.

[2] C.I. Byrnes, A. Lindquist and T. McGregor, *Predictability and Unpredictability in Kalman filtering*, IEEE Trans. on Aut. Contr., **AC-36**: 563-, 1991.

[3] 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.

[4] 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, **AC-40:** 1841-1857, 1995.

[5] C. I. Byrnes and A. Lindquist, *On the Partial Stochastic Realization Problem*, IEEE Transactions on Automatic Control, **AC-42:** 1049-1070, 1997.

[6] 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**, eds. M. Alber, B. Hu, J. Rosenthal, Birkhäuser, pp. 83-106, 1997.

[7] 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, Birhäuser, 1997, pp. 101–136.

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

[9] C. I. Byrnes, P. Enqvist, and A. Lindquist, *Cepstral coefficients, covariance lags and pole-zero models for finite data strings*, submitted for publication, IEEE Trans. on Signal Processing.

[10] C. I. Byrnes, P. Enqvist, and A. Lindquist, in preparation.

[11] C. I. Byrnes and A. Lindquist, *On the duality between filtering and Nevanlinna-Pick interpolation*, SIAM J. Control and Optimization, to be published.

[12] C.I. 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*, submitted for publication, *IEEE Trans. on Aut. Control*.

[13] C. I. Byrnes, T. T. Georgiou, and A. Lindquist, *A new approach to Spectral Estimation: A tunable high-resolution spectral estimator*, *IEEE Trans. on Signal Processing*, to be published.

[14] C.I. Byrnes, T.T. Georgiou, and A. Lindquist, *Generalized interpolation in H_∞: Solutions of bounded complexity*, in preparation.

[15] 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.

[16] C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen*, Math. Ann. **64** (1907), 95–115.

[17] C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Functionen*, Rend. di Palermo **32** (1911), 193–217.

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

[19] P. Enqvist, PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2000.

[20] T.T. Georgiou, **Partial Realization of Covariance Sequences**, Ph.D. thesis, CMST, University of Florida, Gainesville 1983.

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

[22] T.T. Georgiou, *A Topological approach to Nevanlinna-Pick Interpolation*, SIAM J. on Math. Analysis, **18(5)**: 1248-1260, September 1987.

[23] T.T. Georgiou, *The interpolation problem with a degree constraint*, IEEE Trans. Aut. Control, **44(3)**: 631-635, 1999.

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

[25] S. Haykin, ed., **Nonlinear Methods of Spectral Analysis**, Springer Verlag, 1983.

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

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

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

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

[30] B. Porat, *Digital Processing of Random Signals*, Prentice Hall, 1994.

[31] L.R. Rabiner and R.W. Schafer, **Digital Processing of Speech Signals**, Prentice Hall, Englewood Cliffs, N.J., 1978.

[32] D. Sarason, *Generalized interpolation in H_∞*, Trans. American Math. Soc., **127** (1967), 179–203.

[33] I. Schur, *On power series which are bounded in the interior of the unit circle I and II*, Journal fur die reine und angewandte Mathematik **148** (1918), 122–145.

[34] O. Toeplitz, *Über die Fouriersche Entwicklung positiver Funktionen*, Rendiconti del Circolo Matematico di Palermo **32** (1911), 191–192.

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