A HOMOTOPY APPROACH FOR MULTIRATE SPECTRUM ESTIMATION

Ali Nasiri Amini, Mir Shahrouz Takyar, and Tryphon T. Georgiou, Fellow, IEEE

University of Minnesota
Department of Electrical and Computer Engineering
200 Union Street, Minneapolis, MN 55455, USA

ABSTRACT

We present a technique for spectral analysis in the context of multi-rate sampling by a collection of sensors. Correlation of the time-domain samples gives rise to moment constraints for the power spectrum. A homotopy-based technique is then used to identify consistent power spectra. The spectra we obtain are at a minimum distance in the Kullback-Leibler sense to a given "prior" and the "maximum entropy" power spectrum corresponds to the special case where the prior is white.

I. INTRODUCTION

There are numerous engineering applications that require estimation of the power spectrum of a stochastic process from a finite observation record. For instance, in communications, radar, sonar, and geophysical seismology, spectral analysis of a recorded signal/echo is essential for data compression, speaker recognition, target identification, or the identification of underlying geological morphology. To this end, spectral estimation and analysis techniques played a central role in signal processing research over the past several decades.

In this paper, we consider a problem which is typical when several independent measurements of a stochastic process are taken, possibly from different sensors, for the purpose of spectral analysis. This is quite typical in sensor networks, nonuniform phased-array antenna, and remote sensing. In fact, the setting we will consider is typified by non-uniform sampling of an underlying stochastic process. The low resolution samples from a variety of sensors are to be integrated and utilized for identifying the power spectrum of the process. In particular, we focus on the problem as put forth in Jahromi, Francis, and Kwong [1], and we adapt a mathematical framework in [2] for its solution.

II. MULTIRATE OBSERVATIONS

Consider the setup shown in Fig. 1 which represents a *multirate sensing system* having as outputs low-rate observations or measurements $v_i(n)$ of the signal of interest $x(n)$. The linear filters in Fig. 1 model bandwidth limitations of the sensing apparatus. These filters are followed by down-sampler units to model differing sampling rates and communication strength between the various sensors and a fusion center in the sensor network.

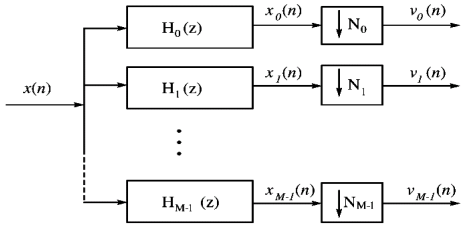


Fig. 1. M-channel multirate observer

If $x(n)$ is wide-sense stationary (WSS) then so are the observed processes $v_0(n)$ to $v_{M-1}(n)$. Their second-order statistics are related to the second-order statistics of $x(n)$ via $R_{v_i}(k) = R_{x_i}(N_i k)$, where $R_{x_i}(k) = (h_i(k) \star h_i(-k)) \star R_x(k)$. Here $h_i(k)$ denotes the impulse response of $H_i(z)$ and "$\star$" denotes convolution. Thus, the second-order statistics of the observation signals and the power spectral density (PSD) of $x(k)$ are related as follows:

$$R_{v_i}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega}) \|H_i(e^{j\omega})\|^2 e^{jN_i k\omega} d\omega.$$

Throughout, for simplicity, we assume $x(n)$ a real valued process and hence, it has a symmetric power spectrum which satisfies:

$$R_{v_i}(k) = \frac{1}{\pi} \int_{0}^{\pi} P_x(e^{j\omega}) \|H_i(e^{j\omega})\|^2 \cos(N_i k\omega) d\omega. \quad (1)$$

The autocorrelation samples $R_{v_i}(k)$ of the observable low-rate signals can be estimated from the corresponding sensors measurements $\{v_i(0), \ldots, v_i(K_i)\}$. The variance of the error in the estimated autocorrelation samples, denoted by $\hat{R}_{v_i}(k)$, increases as k increases. Thus, $R_{v_i}(k)$ can be estimated reliably only up to some order determined by the length of the

This work was supported in part by NSF and AFSOR.

time domain data. Here, we assume that autocorrelation coefficients of $v_i(n)$ have been estimated up to $k = L_i - 1$, for $i = 1, 2, \ldots$ respectively. Henceforth, these estimates are denoted by

$$\rho_i(k) := \hat{R}_{v_i}(k) \text{ for } 0 \leq i \leq M-1, 0 \leq k \leq L_i - 1 \quad (2)$$

and the problem we address is to *determine a power spectral density $P_x(e^{j\omega}) \geq 0$, which is consistent with the moment constraints in (1) and (2).*

This is a typical inverse problem since, in general, there are many solutions which are consistent with the moment constraints. It is a standard practice, in such inverse problems to seek a maximum entropy solution (see [3, 4]). The basic idea is to use entropy as a regularizing functional in order to identify a solution. The principle of maximum entropy is typically justified as seeking the power spectrum of the most "unpredictable stochastic process" which is consistent with moment constraints. Yet, a slight reformulation, where one replaces the entropy functional with the Kullback-Leibler distance to another spectrum—a "prior," leads to a versatile tool for spectral analysis (see e.g., [4, 5]).

There have been several different approaches for solving such optimization problems. The classical case of searching for the maximum entropy given a number of ordinary autocorrelation lags, leads to the set of linear Levinson's equations. This is of course not the case in general (see e.g., [2, 6–9] and the references therein). Jahromi et al. [1] followed a similar formalism and developed an algorithm to approximate the maximum entropy solution of multi-rate sampling spectra. We follow instead the formalism in [2] and show that in fact the approximate maximum entropy solution generated by algorithm in [1] can be vastly different from the exact maximum entropy solution. We then show, that a fairly good fit to simulated power spectra can be obtained from multi-rate covariance data and a reasonable choice of "priors," in the spirit of [5].

III. POWER SPECTRUM FROM MOMENTS

We seek power spectral distributions consistent with (1) and a given set of estimated values for the second-order statistics in (2). The values in (2) for such integral constraints are also known as *moments* of the unknown distribution corresponding to the integration kernels

$$g_{i,k}(\omega) := \frac{1}{\pi} \|H_i(e^{j\omega})\|^2 \cos(kN_i\omega),$$

with $0 \leq i \leq M-1, 0 \leq k \leq L_i - 1$. In case the kernel functions possess a "shift structure," and they can be obtained from one another by application of a suitable shift operator (see [10]), then existence of solutions to (1) can be easily tested by checking positivity of an associated quadratic form, and solutions can be obtained via solving linear equations which are analogous to the Levinson equations. However, when the integration kernels have no apparent structure, as is the case here, solutions can be sought as extrema of regularizing (entropy) functionals. Below we explain one such approach which is based on deforming the values for the moments (see [2] for details).

Let $G(\omega) := [g_{0,0}(\omega),\ g_{0,1}(\omega),\ \ldots]^T$ denote the column vector containing all the integration kernels, and similarly let $R = [\rho_0(0),\ \rho_0(1),\ \ldots]^T$ denote the corresponding moments. Here "T" denotes transposition. The moment constraints are now expressed compactly as

$$R := \int_0^\pi P(\omega) G(\omega) d\omega \quad (3)$$

where $P(\omega)$ is a simplified notation for the sought non-negative distribution function $P_x(e^{j\omega})$.

The basic idea is to seek minimizers of suitable weighted entropy functionals which can be expressed in closed form (see [11] and the references therein). Indeed, the minimizer of the Kullback-Leibler (KL) distance

$$\int_0^\pi \Psi(\omega) \log\left(\frac{\Psi(\omega)}{P(\omega)}\right),$$

between a "prior" $\Psi(\omega)$ and $P(\omega)$ subject to (3), when it exists, it is of the form $\Psi(\omega)/\langle \lambda, G(\omega)\rangle$ where λ is a vector of Lagrange multipliers for the optimization problem and $\langle \cdot, \cdot \rangle$ denotes inner product. Knowing the form of minimizers of the KL-distance, we seek values for the vector λ which render it consistent with the moments, i.e., we seek a solution λ to the nonlinear vectorial equation

$$R := \int_0^\pi \frac{\Psi(\omega)}{\langle \lambda, G(\omega)\rangle} G(\omega) d\omega. \quad (4)$$

This can be computed in the following way.

First note that (4) defines a mapping between the linear spaces where λ and R reside. These have the same dimension. Moment vectors R that are generated by a non-negative $P(\omega)$ form a cone which we denote by $\mathcal{K}$. Similarly, in the dual space, Lagrange multipliers which give rise to a non-negative distribution satisfy $\langle \lambda, G(\omega)\rangle \geq 0$ for all ω in the support set for the frequency variable. They form a dual cone, which we denote by $\mathcal{K}_+^*$. Finally, the Jacobian ∇h of the map $h : \lambda \mapsto R$ specified by (4), is represented by the matrix

$$W(\lambda) := \int_0^\pi \left(G(\omega) \frac{\Psi(\omega)}{\langle \lambda, G(\omega)\rangle^2} G^T(\omega) \right) d\omega$$

which turns out to be bounded and invertible for all λ in the interior of $\mathcal{K}_+^*$. Thus, in seeking a solution λ to (4) for any given vector of moments R, we can trace a path from an initial choice $\lambda_0 \in \mathcal{K}$ to the sought solution λ. The path can be

readily constructed to correspond to the linear segment connecting R_0 to R in the primal space. Both R_0 (obtained from a suitable starting value λ_0) and R belong to $\mathcal{K}$ and hence the path $R_\alpha = (1-\alpha)R_0 + \alpha R$, for $\alpha \in [0,1]$ lies in $\mathcal{K}$ as well. Since $dR_\alpha/d\alpha = R - R_0$, the corresponding values in the dual space satisfy

$$\frac{d\lambda_\alpha}{d\alpha} = -W(\lambda_\alpha)^{-1}(R_0 - R)$$

for $\alpha \in [0,1]$, and after replacing the path variable α by t where $\alpha = 1 - e^{-t}$,

$$\frac{d\lambda(t)}{dt} = -W(\lambda)^{-1}\left(R - \int_0^\pi G(\omega)\frac{\Psi(\omega)}{\langle \lambda, G(\omega)\rangle}d\omega\right) \quad (5)$$

for $t \geq 0$. A schematic is shown in Fig. 2. Such a path

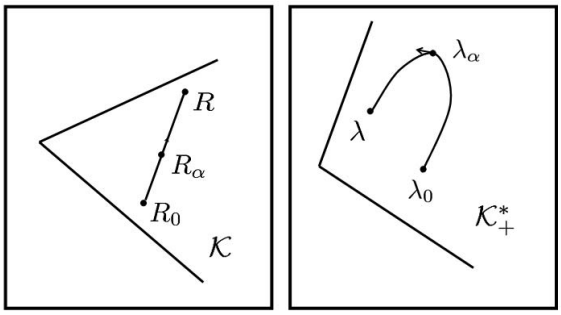


Fig. 2. Schematic of the homotopy path and the corresponding trajectory in the space of λ

can be followed numerically if the Jacobian remains bounded and nonsingular throughout. This is indeed the case provided $R \in \mathcal{K}$. If however, the data vector R does not allow for a consistent power density function, the above differential equation diverges in finite time (see [2, Theorem 2]).

Without loss of generality we can assume that the first entry of $G(\omega)$ is positive throughout the range of frequencies.[1] Then a convenient starting point for integrating (5) is $\lambda_0 = [1\ 0\ \ldots\ 0]^T$. Divergence suggests that the data vector R is not admissible and that no power spectrum consistent with R is possible. On the other hand, if R is a legitimate set of covariance samples then the solution $\lambda(t)$ to (5) will tend to a limit point, exponentially fast, as $t \to \infty$ (see [2] for details).

This theory can be readily applied for identifying a spectrum which is consistent with multirate observations. All that is needed is a differential equation solver to integrate (5) starting from λ_0 until $\lambda(t_F)$ so that R_t matches R to any given accuracy. The steps involved for identifying a power spectral density are summarized below.

[1]If it is not the case, equivalently a suitable linear combination of the kernel functions can be placed as the first entry. Such linear combination exits if the union of the pass band of the filters $H_i(\omega)$ covers the entire frequency band.

Algorithm
Inputs:
$\rho_i(k)$, for $i = 0, \ldots, M-1$ and $k = 0, \ldots, L_i - 1$,
transfer functions of $H_i(z)$, down-sampling ratios N_i,
and a positive function $\Psi(\omega)$.
Output:
An estimate $P_x(e^{j\omega})$ of the input PSD.
Procedure:
1) Integrate $\frac{d\lambda(t)}{dt}$ as in (5) from $t = 0$ and $\lambda(0)$ to $t = t_F$.
2) Return $P_x(e^{j\omega}) = \Psi(\omega)/\langle \lambda(t_F), G(\omega)\rangle$.

A nice feature of this algorithm is that given a "prior" $\Psi(\omega)$ for the unknown density, it generates the closest density in a KL-like sense to $\Psi(\omega)$ which is consistent with the measurements. Note that when $\Psi(\omega) = 1$ the obtained solution has maximal possible entropy rate in the sense of Burg.

IV. NUMERICAL STUDIES

We demonstrate the performance of the proposed algorithm by reworking Example 4 in [1] and compare our results with the results presented there.

Example: Consider a M-channel multirate observer system as in Fig. 1 with $M = 3$. We assume that the filters are FIR filters with impulse responses
$h_0 = [0.075\ \ 0.167\ \ 0.205\ \ 0.166\ \ 0.075]$,
$h_1 = [0.465\ \ 0.125\ \ -0.315\ \ 0.097\ \ -0.026]$,
$h_2 = [0.193\ \ 0.423\ \ 0.367\ \ -0.097\ \ -0.040]$,
and downsampling-rates are $N_0 = 2$, $N_1 = 2$, and $N_2 = 4$. The frequency responses of these filters were designed simply to show lowpass, bandpass, and high-pass characteristics. The unobservable input signal $x(n)$ is chosen to be a lowpass Gaussian ARMA process. Coefficients of this ARMA process were calculated using MATLAB command [a,b]= YULEWALK (10, [0 .5 .8 1], [1 1 0 0]), and the PSD of $x(n)$ is taken as $P_x(e^{j\omega}) = \|\frac{a(e^{-j\omega})}{b(e^{-j\omega})}\|^2$ for the calculated a and b.

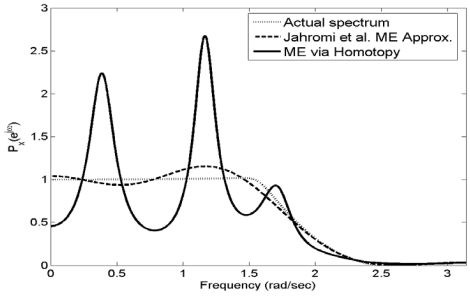
The correlation coefficients associated with the low-rate observable signals $v_i(n)$ are shown in Table I. Then Fig. 3 compares the approximate maximum entropy solution computed in [1] with the exact solution obtained as in previous section. In particular, Fig. 3a corresponds to the case when only the data in the first three columns of the Table I is used, i.e. $L_0 = L_1 = L_2 = 3$. Fig. 3b shows the case when all the data in the Table I, i.e. $L_0 = L_1 = L_2 = 6$. Note that the exact maximum entropy solution computed via the algorithm proposed in this paper and the approximate one computed via Jahromi *et al.* algorithm are vastly different.

It was explained that the output of the proposed algorithm can be viewed as the best approximation of the "prior" $\Psi(\omega)$ among all the PSD functions that are consistent with the given statistics. In order to show the potential of such a tool we apply a lowpass prior template instead of $\Psi(\omega) = 1$. In particular, $\Psi(\omega)$ is chosen as a lowpass Gaussian ARMA process

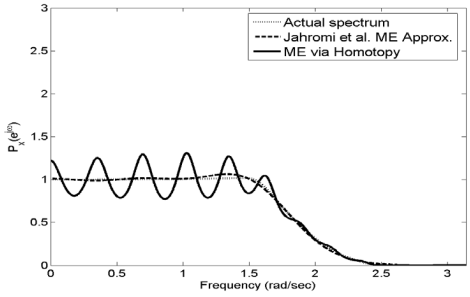
again calculated via MATLAB filter design command [a,b]= YULEWALK(8, [0 .54 .81 1], [1 1 0.01 0.01]). As the input data we use only the first three columns of Table I similar to Fig. 3a. The estimated spectrum is shown in Fig. 4. Comparison of Fig. 4 and Fig. 3a shows that using a reasonable "prior" leads to a significantly better match.

Table 1. CORRELATION COEFFICIENTS

$\rho_i(k)$	k=0	k=1	k=2	k=3	k=4	k=5
i=0	.1084	.0583	.0056	−.0001	.0000	.0000
i=1	.1974	−.1260	.0423	−.0169	.0063	−.0038
i=2	.0438	.0141	−.0002	.0008	−.0002	.0001



(a)



(b)

Fig. 3. PSD estimates using the data given in Table I.

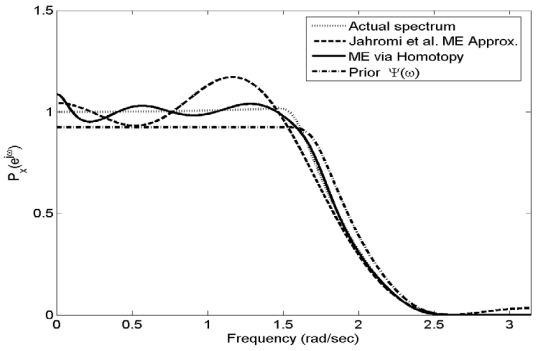


Fig. 4. PSD estimation using prior template

V. REFERENCES

[1] O. Jahromi, B. Francis, and R. Kwong, "Spectrum estimation using multirate observations," *IEEE Trans. Signal Processing*, vol. 52, July 2004.

[2] T. Georgiou, "Solution of the general moment problem via a one-parameter imbedding," *IEEE Trans. Automat. Contr.*, vol. 50, pp. 811–826, June 2005.

[3] J. Burg, "Maximum entropy spectral analysis," Ph.D. dissertation, Stanford University, Stanford, 1975.

[4] I. Csiszár, "Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems," *The Annals of Probability*, vol. 30, pp. 880–887, Dec. 1982.

[5] T. Georgiou and A. Lindquiest, "Kullback-leibler approximation of spectral density functions," *IEEE Trans. Info. Theory*, vol. 49, Nov. 2003.

[6] A. Lewis, "Consistency of moment systems," *Can. J. Math.*, vol. 47, pp. 995–1066, 1995.

[7] M. Miller and D. Snyder, "The role of likelihood and entropy in incomplete-data problems: applications to estimating poin-process intensities and toeplitz constrained covariances," *Proceedings of IEEE*, vol. 75, pp. 892–907, 1987.

[8] S. Lang and J. McClellan, "Multidimensional mem spectral estimatiion," *IEEE Trans. on ASSP*, vol. 30, pp. 880–887, Dec. 1982.

[9] C. I. Byrnes and A. Lindquist, *Control and Modeling of Complex Systems: Cybernetics in the 21st Century*. Boston: Birhäuser, 2003, ch. A convex optimization approach to generalized moment problems, pp. 3–21.

[10] T. Georgiou and A. Lindquiest, "Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parameterization," *IEEE Trans. Automat. Contr.*, vol. 47, Nov. 2002.

[11] C. I. Byrnes, S. V. Gusev, and A. Lindquist, "From finite covariance windows to modeling filters: A convex optimization approach," *SIAM J. Control and Opt.*, vol. 37, pp. 211–229, 1999.