EFFICIENT ADAPTIVE CARRIER TRACKING FOR MARS TO EARTH COMMUNICATIONS DURING ENTRY, DESCENT AND LANDING

C. G. Lopes[†], E. Satorius[‡], P. Estabrook[‡] and A. H. Sayed[†]

[†]Electrical Engineering Department, UCLA, Los Angeles, CA 90095
Email:{cassio,sayed}ee.ucla.edu

[‡] Jet Propulsion Laboratory, Pasadena, CA 91109
Email: {Edgar.H.Satorius,Polly.Estabrook}@jpl.nasa.gov

ABSTRACT

In the Mars rover missions the signals transmitted back to Earth travel under low SNR conditions in highly non-stationary channels [1, 2]. During the entry, descent and landing phase (EDL), the spacecraft high dynamics yields severe Doppler effects. We propose a robust and low complexity scheme to estimate and track carrier frequency from the received signals at the Earth end. The method employs a hierarchical arrangement of convex linear prediction cells that is dynamically adapted to respond to the channel conditions. The adaptive combination is able to outperform the best individual estimator in the set, leading to a universal scheme for frequency estimation and tracking. In order to compensate the lag error effect, we explore an efficient forward and backward aggregation scheme that improves considerably the frequency RMS error as compared to the original method [3].

Index Terms–Adaptive filters, frequency estimation, combination of filters, Mars exploration.

1. INTRODUCTION

In space missions to Mars, the most critical period for communications is the entry, descent and landing phase (EDL) [1, 2]. During this phase, a complex sequence of events takes place, and health and status signals are sent back in real time to Earth through the direct-to-earth (DTE) channel. In order to support spacecraft-to-earth communications, we recently developed a low complexity carrier frequency estimation and tracking technique [3] that is able to operate under low SNR and highly non-stationary conditions, common to the adverse EDL scenario. In this work we design more efficient convex estimators and explore a method to mitigate bias errors in the frequency estimates.

2. EDL COMMUNICATIONS

Due to the EDL events, the signals travel through the DTE channel experiencing a combination of severe Doppler shift, time-varying gain and noise. These effects make the recovery of the data from the received signal a challenging task.

The EDL events are flagged into the transmitted signal $s(t)$, which is a 256-MFSK signal whose nominal carrier frequency is $f_c^0 = 8.4$ GHz (X-band). At the Earth end, the received signal $x(t)$ is comprised of a distorted signal component $r(t)$ disturbed by noise $v(t)$, as illustrated in Fig. 1. A detailed description of the DTE channel and signals generation can be found in [2].
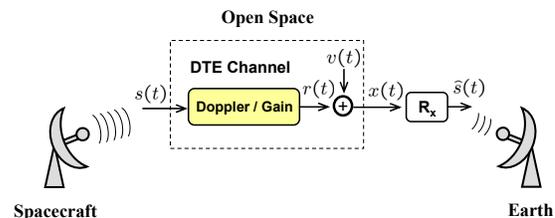
Fig. 1. Direct-to-Earth communications.

In order to recover the MFSK data, we need a reliable estimate of the carrier frequency, whose nominal component f_c^0 is shifted by a strong Doppler component $f(t)$ caused by the spacecraft high dynamics [1]:

$$f_c(t) = f_c^0 + f(t) \qquad (1)$$

Figure 2 illustrates a typical Doppler profile experienced by the landers of the Opportunity and Spirit missions. The acceleration peak and the parachute deployment represent the most challenging EDL epochs for frequency estimation.

Due to the nature of the problem, it is reasonable to assume that there is no embedded data and that the signal is down-converted and sampled upon reception, so that the received signal can be modeled as [2]:

$$x(i) = e^{j\omega i} + v(i) \qquad (2)$$

where ω is the discrete time-varying Doppler component and $v(i)$ arises from an ergodic white noise process with variance σ_v^2. Our objective is to estimate and track ω from measurements $\{x(i)\}$.
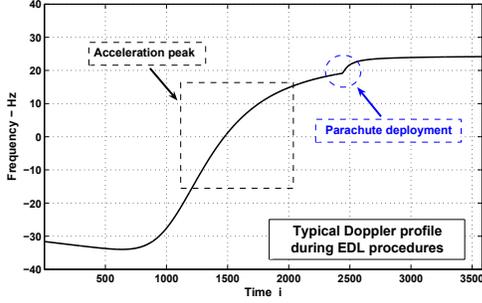
Fig. 2. Doppler profile during EDL for the Opportunity and Spirit Mars missions (reduced bandwidth).

3. FREQUENCY ESTIMATION VIA ADAPTIVE LINEAR PREDICTION

It is well known that for signal models of the form (2), the Maximum-Likelihood (ML) technique can be effective [4], [5]. This is the core technique proposed in [1, 2]. However, in the general non-stationary case, ML requires a multidimensional search that can be computationally prohibitive.

One alternative for estimation and frequency tracking is to formulate a linear prediction problem [5]–[12]. In this approach, a linear predictor, generally an FIR structure, described by

$$\hat{x}(i) = \sum_{k=1}^{M} c(k)x(i-k) \qquad (3)$$

is designed to minimize the estimation error

$$e(i) = x(i) - \hat{x}(i) \qquad (4)$$

according to some criterion, for example, in the least (mean) squares sense. With the signal model (2), the optimal least mean-squares predictor coefficients $c(k)$ are given by [13]:

$$c^o = \frac{1}{\sigma_v^2 + M} \mathrm{col}\{e^{j\omega}, e^{j2\omega}, \dots, e^{jM\omega}\} \qquad (5)$$

This formulation is related to the Maximum Entropy Method of spectral analysis, as well as AR process modeling [7, 12]. The prediction-error filter $Q(z)$, defined by

$$Q(z) = 1 - \sum_{k=1}^{M} c(k)z^{-k} \qquad (6)$$

is related to the input signal power spectral density $S_x(e^{j\omega})$ via

$$S_x(e^{j\omega}) = \frac{K}{|Q(e^{j\omega})|^2} \qquad (7)$$

for a scaling factor K. The model (2) implies that the spectrum of $x(i)$ will be comprised primarily of a single lobe peak, corresponding to the sinusoidal data component. From (7), the closest root of $Q(z)$ to the unit circle is responsible for the peak, and its phase is an estimate of the Doppler frequency $f(i)$ [9, 11, 12].

A useful approach is to employ adaptive filters to efficiently design $Q(z)$. In this work, predictors from the affine projection family [13] are tested, with a particular emphasis on NLMS, due to its robustness, simplicity and reported efficiency [3]:

$$c_i = c_{i-1} + \mu \frac{x_{i-1}^*}{\|x_{i-1}\|^2 + \varepsilon} \left(x(i) - x_{i-1}c_{i-1} \right) \qquad (8)$$

where

$$x_{i-1} = [x(i-1) \, x(i-2) \, \cdots \, x(i-M+1)] \qquad (9)$$

At each time i, the adaptive predictor presents the prediction-error filter $Q_i(z)$ to a root solver, which finds the closest root to the unit circle, $r_o = \rho_o e^{j\theta_o}$. An estimate of the unknown Doppler frequency $f(i)$ in Hertz is then found from

$$\hat{f}(i) = \frac{\theta_o}{2\pi} \cdot F_s \quad (\mathrm{Hz}) \qquad (10)$$

where F_s is the sampling frequency. However, the position of the roots of $Q_i(z)$ tends to be sensitive to perturbation in the filter coefficients, caused by low SNR and gradient noise.

One technique to reduce the noise effect is proposed in [5], where a reduced rank least-squares (RRLS) FIR tracker is presented, inspired by the work in [7]. For comparison purposes, we implemented a forward RRLS predictor, i.e., a predictor that uses past samples to predict the current state, with a forgetting factor $0 \ll \alpha < 1$ to cope with channel dynamics.

Figure 3 compares the NLMS and forward RRLS solution for different design parameters at SNR $= 14$dB $-$ Hz$(-6$dB$)$. In the top row, left plot, one can see that despite the small filter order $(M = 5)$, the NLMS presents good performance, maintaining the lock status throughout the EDL phase. However, a slight change in the design parameters leads to considerable degradation in performance, as shown in the right plot. The bottom row shows the forward RRLS. Following [5], we employed $M = 18$. The left plot shows the algorithm response for a typical choice of the forgetting factor, $\alpha = 0.996$ [13]. One can see that the lock is lost in the region of peak acceleration. In the right plot, the RRLS algorithm was equipped with a faster forgetting factor, $\lambda = 0.9$, improving the lock ability, but the presence of spikes due to low SNR became dominant. As the forgetting factor λ approaches unity, the performance severely degrades and the LS-based algorithm is unable to react to the changing channel.

Figure 3 also shows the robustness of the NLMS predictor: even for a small order $(M = 5)$, the NLMS algorithm is able to outperform the RRLS solution, and at an inexpensive computational cost.

As Fig. 4 shows, the RRLS solution can be improved further by fine tuning the forgetting factor, in this case $\alpha = 0.98$. Nevertheless, increasing the NLMS order slightly, i.e., $M = 8$, with $\mu = 0.075$, also improves further the NLMS response; thus outperforming the RRLS solution in both examples. Note that this was achieved with small predictor
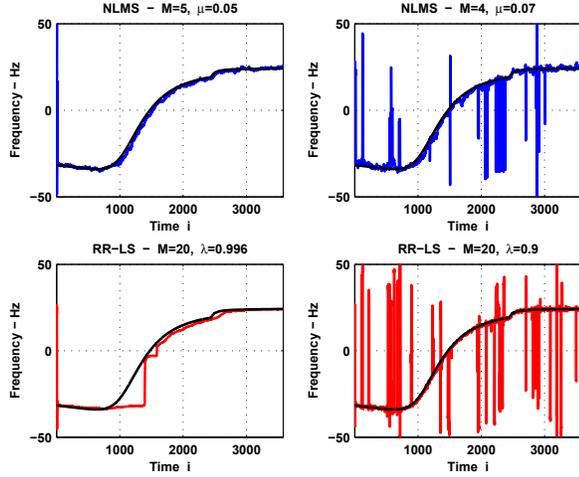
Fig. 3. Sensitivity of linear prediction methods.

orders, $M = 5$ and $M = 8$, representing an inexpensive computational cost, when compared to the RRLS approach.
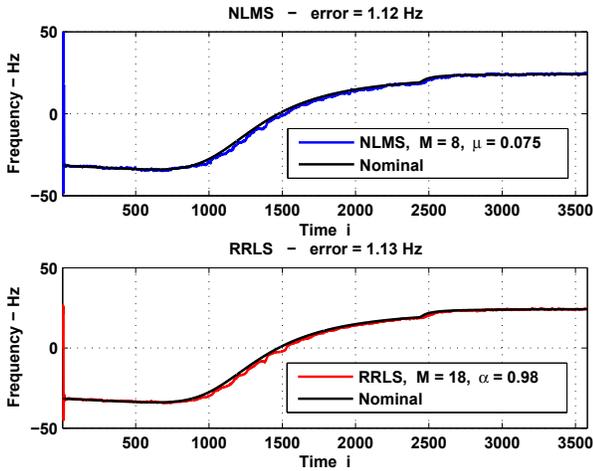
Fig. 4. Comparison: NLMS versus forward RRLS.

4. COMBINATION PREDICTION SCHEME

As indicated by Figs. 3 and 4, both NLMS and RRLS are sensitive to the design parameters. The procedure we present now applies to both NLMS and RRLS. However we focus on the NLMS algorithm.

In [3], we employed a convex mixture of multiple individual predictors to overcome the design sensitivity of the root configuration with respect to the predictor parameters. Normalized LMS predictors with orders M_k and step-sizes μ_k were organized into a single combination layer so that the individual predictors are independent. The single layer configurations were able to perform as well as the best individual predictor [3, 14]. However, when the number of filters

L is increased, the extra gradient noise introduced by the combiners may compromise the overall performance. To combat this effect, we explore here a hierarchical arrangement of simple $L = 2$ convex cells (L2-cell)– see Fig. 5, that operate using one combiner λ only [3]. The combiner coefficient at time i is computed as

$$\lambda(i) = \frac{1}{1 + \left|e^{-\frac{a(i-1)}{2}}\right|^2} \qquad (11)$$

where a is a complex quantity that is adapted as [14]

$$a(i) = a(i-1) - \mu_a\left[\nabla_a|e(i)|^2\right]^*_{a=a(i-1)} \qquad (12)$$

and $e(i) = x(i) - \hat{x}(i)$, with $\hat{x}(i) = x_{i-1}c_{i-1}$, is the global prediction error defined in terms of

$$c_{i-1} = \lambda(i)c_{u,i-1} + (1 - \lambda(i))c_{\ell,i-1} \qquad (13)$$

It follows that, for the output layer, [3]:

$$a(i) = a(i-1) + \mu_a e(i)\left(\hat{x}_u(i) - \hat{x}_\ell(i)\right)^*\lambda(i)\left(1 - \lambda(i)\right) \quad (14)$$

For the input layer, the upper and lower cell errors are given by $e_u(i) = x(i) - \hat{x}_u(i)$ and $e_\ell(i) = x(i) - \hat{x}_\ell(i)$, where $\hat{x}_u(i) = x_{u,i-1}$ and $\hat{x}_\ell(i) = x_{\ell,i-1}$, and they are defined in terms of

$$c_{u,i-1} = \lambda_u(i)c_{1,i-1} + (1 - \lambda_u(i))c_{2,i-1} \quad \text{(upper)} \quad (15)$$
$$c_{\ell,i-1} = \lambda_\ell(i)c_{3,i-1} + (1 - \lambda_\ell(i))c_{4,i-1} \quad \text{(lower)} \quad (16)$$

The two L2-cells in the input layer adopt a similar combining function:

$$\lambda_u(i) = \frac{1}{1 + \left|e^{-\frac{a_u(i-1)}{2}}\right|^2} \quad \text{and} \quad \lambda_\ell(i) = \frac{1}{1 + \left|e^{-\frac{a_\ell(i-1)}{2}}\right|^2}$$

with learning rules

$$a_u(i) = a_u(i-1) + \mu_a e_u(i)\left(\hat{x}_1(i) - \hat{x}_2(i)\right)^*\lambda_u(i)\left(1 - \lambda_u(i)\right)$$
$$a_\ell(i) = a_\ell(i-1) + \mu_a e_\ell(i)\left(\hat{x}_3(i) - \hat{x}_4(i)\right)^*\lambda_\ell(i)\left(1 - \lambda_\ell(i)\right)$$

Note that the regressors $x_{u,i-1}$ and $x_{\ell,i-1}$ have the same order as c_2 and c_4, respectively. Whenever necessary, the regressors are filled out with zeros to match the dimensions in the equations [3]. In addition, the combiners λ, λ_u and λ_ℓ are time-smoothed over their past N_{ham} values via a Hamming half-window, generating $\overline{\lambda}$, $\overline{\lambda_u}$ and $\overline{\lambda_\ell}$. More recent samples are emphasized, so that the window peak is at the current sample. The window coefficients are normalized to sum up to unity. This procedure helps combat the extra gradient noise introduced by the learning rules a, a_u and a_ℓ.

In Fig. 6 we run a simulation with low SNR ($12\text{dB} - \text{Hz}$, or $-8dB$) to illustrate the superior design of the hierarchical arrangement, as compared to the original L2-cells or the individual NLMS predictors. For this example, 10 experiments were performed with $\mu_1 = \mu_2 = 0.075$ and $\mu_3 = \mu_4 = 0.11$ for the individual predictors and $\mu_a = 0.4$ for the learning rules. The left plot shows the evolution of λ,
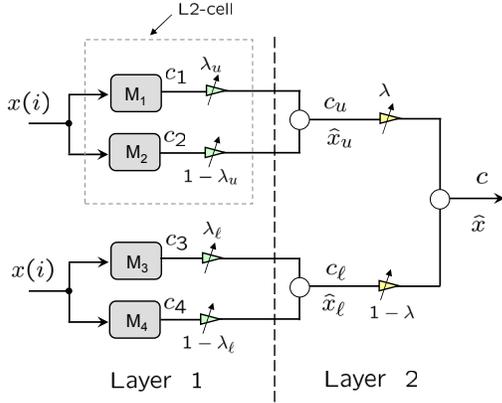
Fig. 5. Convex hierarchical scheme for $L = 4$ experts in the input layer. The order of the k-th NLMS predictor is denoted by M_k.
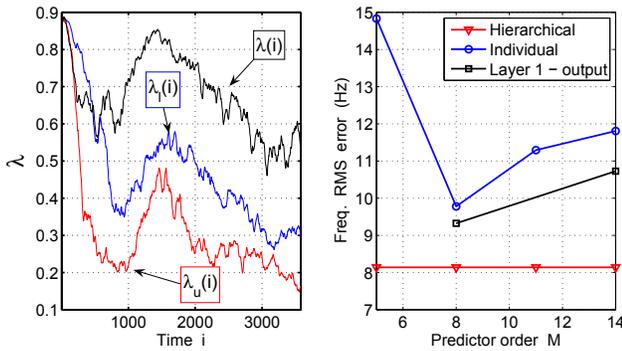
Fig. 6. Universality of the convex hierarchical scheme - SNR = 12 dB-Hz (-8dB).

λ_u and λ_ℓ. Note how they are adjusted to properly select the best predictor. In the region of peak acceleration, the combiners correctly tend to assign larger weights to the smaller order predictors, since they can react faster to the channel. In periods of lower dynamics, they commute to the larger order predictors in order to decrease the mean square error. This effect is emphasized for a larger step-size μ_a. The right plot depicts the frequency root mean squared (RMS) error of all the predictors in the system. All the L2-cells' outputs outperform their (input) individual predictors. The RMS error is defined in terms of the nominal Doppler profile $f(i)$ (refer to Fig. 2) and an estimate of interest $\hat{f}(i)$:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |f(i) - \hat{f}(i)|^2}$$

5. TRACKING LOCK CONTROL

The low SNR effects in the frequency estimates are worsened by the stochastic gradient disturbances introduced by the predictors (8) as well as the adaptive combiners rule
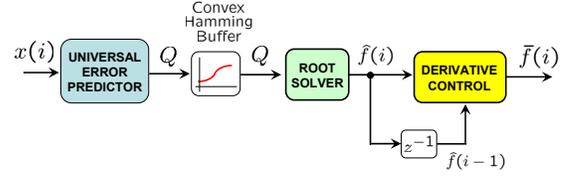
Fig. 7. The combination scheme with smoothing and derivative control.

(12). This leads to spikes in the estimated frequency that are not related to the actual Doppler frequency. The lock control attempts to enforce continuity of frequency, reflecting the natural behavior of the underlying physical process. The original scheme has been improved by employing a convex Hamming smoother over the derivatives buffer [3], which keeps track of the most recent "good" derivative samples, and also over consecutive predictors c_{i-1} ($Q_i(z)$). Whenever a jump is detected, continuity is enforced:

$$\bar{f}(i) = \begin{cases} \hat{f}(i), & \text{if } |\delta \hat{f}(i)| \leq \text{THR} \\ \hat{f}(i-1) + \overline{\delta f}(i), & \text{if } |\delta \hat{f}(i)| > \text{THR} \end{cases} \quad (17)$$

where $\delta f = \hat{f}(i) - \hat{f}(i-1)$, and $\overline{\delta f}$ is the spike-free derivative obtained from the temporal hamming smoother.

6. LAG-ERROR COMPENSATION

The estimates delivered by the ALPs will be generally biased due to the root solver operation [5]. The effects of the low SNR conditions are worsened if the NLMS step-size is large, since the gradient noise is directly amplified by that. As a result, the step-size is required to be small ($\mu \approx 0.1$), otherwise the spurious frequency spikes become severely dominant. On the other hand, a small step-size has the counter-effect of decreasing the ability of the filter to respond to the process dynamics, thus increasing the lag error (bias). In other words, in the non-stationary scenario, the estimates delivered by the ALPs will be biased not only due to the root solver operation, but also as a result of the learning latency experienced by the adaptive filters [13]. This effect can be compensated by exploring the learning latency of the adaptive predictors, as explained next.

When performing forward prediction, the adaptive filter suffers from a latency effect, which causes the forward frequency estimate $\bar{f}$ to always fall behind the true value, generating a causal lag error – solid (red) line in Fig. 8, which represents the forward operation, from point A to point B. Now, considering an off-line processing scenario, if a backward prediction is performed, from point B towards point A, the latency effect will again cause the backward estimate $\bar{f}_b$ to fall "behind" the true frequency value. However, due to the virtual "flip" in the time processing, the lag will now be anti-causal with respect to the original time axis, and also to the previous forward estimates, as illustrated by the dashed (blue) curve in Fig. 8. By simply performing a time average of both $\bar{f}$ and $\bar{f}_b$ estimates, the bias error can be greatly
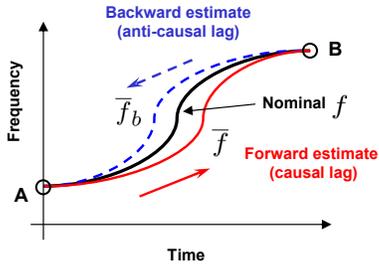
Fig. 8. Principle of the bias error compensation.

combatted. The idea can be extended to a general mapping

$$\overline{f}_{nl} = g\left(\overline{f}, \overline{f}_b\right) \tag{18}$$

for a properly chosen function $g(\cdot)$. For instance, $g(\cdot)$ could be a median filter to combat also the spurious frequency spikes.
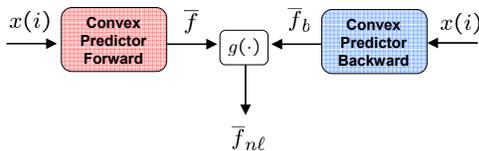
Fig. 9. The global scheme with lag-error compensation. The forward and backward predictors are implementations of the scheme in Fig. 7.

In Figure 10, we run a simulation for a wide range of SNR, and compare the performance of the original L2-cell scheme, with the entire system proposed (Figs. 9 and 7). The set of predictors employed is $M_k = [9, 13, 17, 21]$, with $\mu_k = 0.11$. The improvement in performance is expressive, especially for large SNR.

7. CONCLUDING REMARKS

We are currently developing methods that efficiently extract the frequency content from the predictor coefficients, leading to substantial improvement in performance, especially at low SNR. In addition, signal pre-conditioning techniques to enhance the signal and improve performance are also being tested. More sophisticated and more robust lock enforcement techniques are currently being developed.

8. REFERENCES

[1] W. J. Hurd, P. Estabrook, C. S. Racho, and E. Satorius. "Critical spacecraft-to-earth communications for Mars exploration rover (MER) entry, descent and landing,"*Proc. IEEE Aerospace Conference*, vol.3, pp. 1283-1292, MT, March 2002.

[2] E. Satorius, P. Estabrook, J. Wilson, D. Fort. "Direct-to-Earth communications and signal processing for Mars exploration rover entry, descent and landing," *The Interplanetary Network Progress Report*, IPN Progress Report 42-153, May 2003.
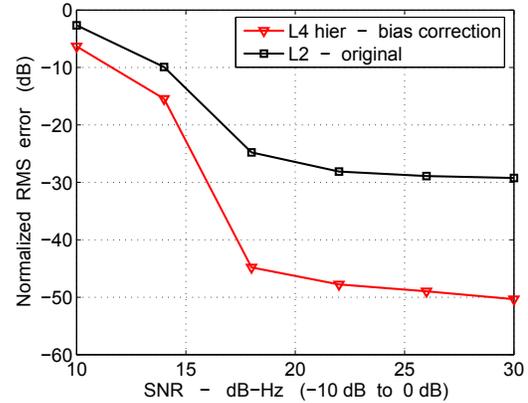
Fig. 10. RMSE profile of the $L = 4$ hierarchical convex system proposed (with bias correction) versus the original $L = 2$ scheme. The normalization is performed with respect to $\|f(i)\|^2$.

[3] C. G. Lopes, E. Satorius, A. H. Sayed. "Adaptive carrier tracking for Direct-to-Earth Mars communications," *Proc. of 40th Asilomar Conf. on Signals, Systems and Computers*, pp. 1042-1046, Pacific Grove, CA, Nov 2006.

[4] V. A. Vilnrotter, S. Hinedi, R. Kumar "Frequency estimation techniques for high dynamics trajectories," *IEEE Trans. on Aerospace and Eletronic Systems*, vol. 25, No.4, pp. 559-577, July 1989.

[5] D. W. Tufts and R. Kumaresan, "Estimating the parameters of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, no. 9, pp. 975-989, Sept. 1982.

[6] J. A. Chambers and A. G. Constantinides "Frequency tracking using constrained adaptive notch filters synthesized from allpass sections," *IEE Proceedings*, vol. 137, no.6, pp.475–481, December 1990.

[7] T. J. Ulrych and R. W. Clayton. "Time series modelling and maximum entropy," *Phys. Earth Planet. Interiors*, vol. 12, pp. 188-200, Aug. 1976

[8] B. Friedlander "A recursive maximum likelihood algorithm for ARMA line enhancement," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, No. 4, pp. 651-657, August 1982.

[9] L. J. Griffiths "Rapid Measurement of digital instantaneous frequency," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-23, No. 2, pp. 207-222, April 1975.

[10] K. L. Peacock and S. Treitel "Predictive deconvolution: theory and practice," *Geophysics*, vol. 34, pp. 155-169, April 1969.

[11] L. B. Jackson, D. W. Tufts, F. K. Soong, and R. M. Rao "Frequency estimation by linear prediction," *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 352-356, April 1978.

[12] E. H. Satorius and J. R. Zeidler "Maximum entropy spectral analysis of multiple sinusoids in noise," *Proc. of Geophysics*, vol. 43, no.6, pp. 1111–1118, October 1978.

[13] A. H. Sayed. *Fundamentals of Adaptive Filtering*, Wiley, NY, 2003.

[14] J. Arenas-Garcia, A. R. Figueiras-Vidal, and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters,"*IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1078–1090, March 2006.