Adaptive Carrier Tracking for Mars to Earth Communications During Entry, Descent, and Landing

C. G. LOPES, Member, IEEE
University of São Paulo

E. H. SATORIUS
P. ESTABROOK
Jet Propulsion Laboratory

A. H. SAYED
UCLA

We propose a robust and low complexity scheme to estimate and track carrier frequency from signals traveling under low signal-to-noise ratio (SNR) conditions in highly nonstationary channels. These scenarios arise in planetary exploration missions subject to high dynamics, such as the Mars exploration rover missions. The method comprises a bank of adaptive linear predictors (ALP) supervised by a convex combiner that dynamically aggregates the individual predictors. The adaptive combination is able to outperform the best individual estimator in the set, which leads to a universal scheme for frequency estimation and tracking. A simple technique for bias compensation considerably improves the ALP performance. It is also shown that retrieval of frequency content by a fast Fourier transform (FFT)-search method, instead of only inspecting the angle of a particular root of the error predictor filter, enhances performance, particularly at very low SNR levels. Simple techniques that enforce frequency continuity improve further the overall performance. In summary we illustrate by extensive simulations that adaptive linear prediction methods render a robust and competitive frequency tracking technique.

Manuscript received August 9, 2008; revised April 7 and 24, 2009; released for publication May 29, 2009.

IEEE Log No. T-AES/46/4/938803.

The refereeing of this contribution was handled by M. Rice.

This material was supported by the Jet Propulsion Laboratory under Award 1276256 and by the National Science Foundation Award ECS-0601266.

The work of C. G. Lopes was also supported by a fellowship from CAPES, Brazil, under Award 1168/01-0.

Brief versions of this work appear in *Proceedings of the 40th Asilomar Conference on Signals, System, and Computers*, Nov. 2006, and in *Proceedings of the IEEE Workshop on Statistical Signal Processing*, Aug. 2007.

Authors' current addresses: C. Lopes, Dept. of Electronic Systems, University of São Paulo, Escola Politécnica, São Paulo, Brazil; A. Sayed, Dept. of Electrical Engineering, UCLA, Los Angeles, CA 90095, E-mail: (sayed@ee.ucla.edu); E. Satorius and P. Estabrook, Jet Propulsion Laboratory, Pasadena, CA 91109.

I. INTRODUCTION

As part of the NASA exploration effort of the red planet, the new Mars Science Laboratory (MSL) rover is larger, heavier, and embedded with more sophisticated instruments than its predecessors, the Spirit and the Opportunity rovers. MSL assesses whether Mars has ever had, or still has, environmental conditions to support microbial life.

The most critical phase of the mission is the entry, descent, and landing (EDL). The lander enters the Mars atmosphere at hypersonic speed and undergoes a myriad of events until it finally rests on Martian soil. These events are registered by signaling flags that are embedded into an M-ary frequency-shift keying (MFSK)-modulated carrier and sent back to Earth in real time via the direct-to-Earth (DTE) channel, which is employed as a backup system in case of failure of the main link [1, 2]. These signals reflect the health and status of the mission and are crucial to improve future designs in case of mission failure. Due to the EDL events, such signals travel through the DTE channel and experience a combination of severe Doppler effect, time-varying gain, and noise. These effects make the recovery of the data from the received signal a challenging task.

In order to retrieve the mission data, the severe Doppler shifts in the FSK carrier caused by the EDL dynamics must be estimated, and this is the purpose of our work. The original estimation procedure employed in the Spirit and the Opportunity missions was based on maximum likelihood (ML) techniques [1, 2]. We follow a different approach, aiming at low complexity and robustness. For this purpose we revisit adaptive linear prediction (ALP) techniques, a topic that has been available in the literature of frequency tracking for quite some time [3–13].

Despite the limitations of ALP methods in typical high signal-to-noise ratio (SNR) and (moderately) stationary scenarios, the unusual conditions of the EDL dynamics unveil that they can be competitive if combined with recent developments in adaptive filtering [14–18], together with some enhancement techniques presented in the next sections [19, 20]. Specifically, we develop a robust and low complexity carrier frequency tracking scheme that is able to operate under low SNR and highly nonstationary conditions during EDL. The method combines the natural tracking abilities of adaptive filters with universal prediction techniques, allied with lag error compensation and efficient frequency retrieval, order-adaptive smoothing, as well as robust frequency lock control.

The paper is organized as follows. Section II introduces the EDL particularities and the adopted models for the channel and signals. Sections III and IV revisit linear prediction (LP) and ALP techniques, introducing the optimal frequency profile estimate

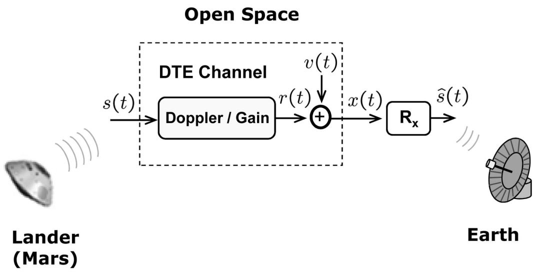


Fig. 1. DTE communications.

for the time-varying case, and the normalized least-mean squares (NLMS) predictor as a robust tracker. Section V proposes convex combination methods to mitigate the sensitivity of the ALPs to predictor parameters. A simple lag-error compensation procedure is presented in Section VI, followed by efficient fast Fourier transform (FFT)-based frequency retrieval presented in Section VII, altogether yielding a favorable improvement margin. Section VIII introduces an order-adaptive smoothing method and a frequency derivative control routine, both improving frequency tracking lock; the subsystems are then grouped, and performance is compared with the original ML-based solution [1, 2] and with the reduced-rank least-squares (RR-LS) method [9].

II. ENTRY, DESCENT, AND LANDING COMMUNICATIONS

At the spacecraft end, every 10 s signals that register the lander status are sent back to Earth as the EDL events take place. Due to the critical channel conditions, phase-coherent communication is not viable. A modified MFSK modulation technique is adopted by JPL [1], with a nominal carrier frequency of $f_c^0 = 8.4$ GHz (X-band) and employing a constellation of 256 symbols.

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 the signal generation can be found in [2]. The spacecraft high dynamics leads to time-varying Doppler shifts in the nominal carrier frequency [1]:

$$f_c(t) = f_c^0 + f(t) + f_d(t) \tag{1}$$

where $f_c^0 = 8.4$ GHz, $f(t)$ denotes the Doppler shifts, and $f_d(t)$ represents the data component. Therefore, in order to recover the MFSK data, we need a reliable estimate of the Doppler component. Due to the large values of $f(t)$, we may assume that $f(t) + f_d(t) \approx f(t)$ so that the embedded data $f_d(t)$ may be disregarded for $f(t)$ estimation purposes.[1] In other words after

[1]When the relative motion between source and receiver is slow, this assumption is invalid.

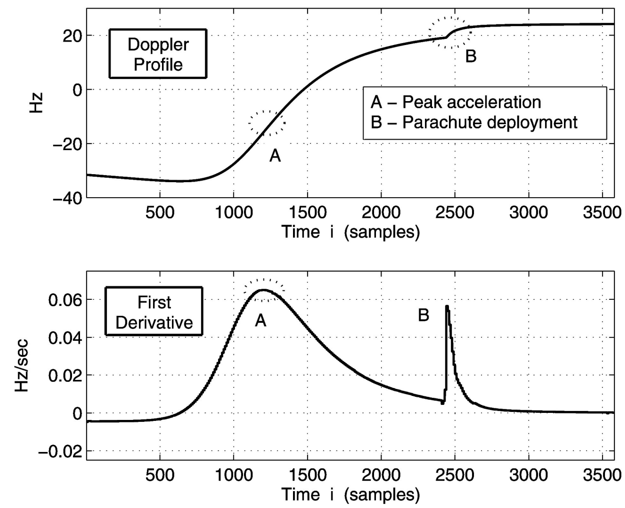


Fig. 2. Typical Doppler frequency profile $f(i)$ encountered during EDL on Mars (reduced band) and its first derivative.

down-conversion, the only frequency content in the received signal is the Doppler component, i.e.,

$$f_c(t) = f(t). \tag{2}$$

After sampling, the signal that we are dealing with at the Earth end is of the form

$$x(i) = e^{j\phi(i)} + v(i) \tag{3}$$

where $v(i)$ models the channel noise and where it is assumed to arise from an ergodic white process with variance σ_v^2, and

$$\phi(i) = 2\pi \sum_{k=0}^{i} f(k) \tag{4}$$

with $f(i)$ representing the discrete time-varying Doppler component, depicted in Fig. 2.

Our objective is to estimate and track $f(i)$ from the measurements $\{x(i)\}$. After the Doppler component has been estimated and removed, the EDL data can be properly recovered. We consider signals with reduced bandwidth so that the sampling frequency is chosen as $F_s = 100$ Hz, which is representative of the EDL dynamics. (See Fig. 2.)

III. LINEAR PREDICTION REVISITED: A NEW PROBLEM FOR AN OLD SOLUTION

For signals comprised of a narrowband component embedded in a wideband background noise, the direct frequency estimation via ML is quite effective; however it leads to a (nonlinear) least-squares problem that requires a multidimensional search in general, particularly for nonstationary processes [3, 4]. In addition the frequency resolution of such methods tends to be limited by the length of the observation window [5].

A. Linear Prediction

One alternative approach for estimation and frequency tracking is to formulate an LP problem [6–9]. The LP-based methods have been originally proposed to overcome the frequency resolution limitations of periodogram-based (which includes ML) and autocorrelation methods [5]. As a byproduct, these methods are, sometimes, orders of magnitude simpler than direct ML approaches.

Linear prediction methods are usually avoided in (quasi) stationary and high SNR applications due to their bias and suboptimality. However in the critical EDL scenario, with very low SNR and highly nonstationary DTE channel conditions, the robustness and learning ability of linear predictors driven by adaptive algorithms are more relevant than their bias and suboptimality, which is explored in the next sections.

The principle of LP frequency estimation is simple. When an M-order predictor with coefficients $c(k)$ is optimally designed in the mean-square error (MSE) sense, for the signal model (3), the resulting error predictor filter $Q(z)$, defined by

$$Q(z) = 1 - \sum_{k=1}^{M} c(k) z^{-k} \tag{5}$$

presents a particular root configuration: the phase θ_0 of the root that is closest to the unit circle yields an estimate of the Doppler component f via

$$\hat{f} = \frac{\theta_o}{2\pi} \cdot F_s \quad \text{(Hz)} \tag{6}$$

where F_s is the sampling frequency.

B. Predictor Design in the Time-Varying Case

The time-varying optimum (forward) predictor vector c_i^o, corresponding to a dynamic Doppler component $f(i)$, can be determined by solving

$$c_i^o = \arg\min_c E|x(i) - \hat{x}(i)|^2 \tag{7}$$

with

$$\hat{x}(i) = x_{i-1} c \tag{8}$$

with

$$\begin{aligned} x_{i-1} &= [x(i-1) \quad x(i-2) \cdots x(i-M)] \\ c &= [c(1) \quad c(2) \cdots c(M-1) \quad c(M)]^T \end{aligned} \tag{9}$$

where x_{i-1} is the row observation vector and c collects the predictor coefficients. The solution to (7) is the solution to the time-varying normal equations [21]:

$$R_{u,i} c_i^o = R_{du,i} \tag{10}$$

where $R_{u,i}$ and $R_{du,i}$ are, respectively, the covariance and cross-covariance matrices of the regressor u and the desired signal d, with

$$d \leftarrow x(i) \qquad \text{and} \qquad u \leftarrow x_{i-1}. \tag{11}$$

Since we have that

$$x(i) = r(i) + v(i) \tag{12}$$

with $r(i) = e^{j\phi(i)}$, then we may express the regressor x_{i-1} as

$$x_{i-1} = r_{i-1} + v_{i-1} \tag{13}$$

where

$$\begin{aligned} r_{i-1} &= [r(i-1) \quad r(i-2) \cdots r(i-M)] \\ v_{i-1} &= [v(i-1) \quad v(i-2) \cdots v(i-M)]. \end{aligned}$$

We proceed by calculating the data covariance matrices using the fact that $v(i)$ is white noise and that $r(i)$ is deterministic. From model (3) and (13), we have

$$\begin{aligned} R_{du,i} &= Ex(i)x_{i-1}^* = E(r(i) + v(i)) \cdot (r_{i-1}^* + v_{i-1}^*) \\ &= r(i) r_{i-1}^* \end{aligned} \tag{14}$$

and

$$\begin{aligned} R_{u,i} &= Ex_{i-1}^* x_{i-1} \\ &= E(r_{i-1}^* + v_{i-1}^*) \cdot (r_{i-1} + v_{i-1}) \\ &= r_{i-1}^* r_{i-1} + \sigma_v^2 I_M. \end{aligned} \tag{15}$$

Applying the matrix inversion lemma [21] to (15) leads to

$$\begin{aligned} R_{u,i}^{-1} &= \frac{1}{\sigma_v^2} I_M - \frac{1}{\sigma_v^2} r_{i-1}^* \left(1 + \frac{1}{\sigma_v^2} r_{i-1} r_{i-1}^*\right)^{-1} r_{i-1} \frac{1}{\sigma_v^2} \\ &= \frac{1}{\sigma_v^2} \left(I_M - \frac{1}{\sigma_v^2 + M} r_{i-1}^* r_{i-1}\right) \end{aligned} \tag{16}$$

which, postmultiplied by (14), yields

$$c_i^o = \frac{r(i)}{\sigma_v^2 + M} r_{i-1}^*. \tag{17}$$

The optimal sequence $\{c_i^o\}$ results in the mean-square optimal Doppler profile estimate $\hat{f}^o(i)$. Fig. 3 depicts the frequency profile obtained from $\{c_i^o\}$ and (6), with $M = 60$ and $\sigma_v^2 = 1$ (SNR = 0 dB). As expected, note how the resulting curve is biased [7, 9, 10, 11]. Most of this effect, however, can be canceled via a technique presented in Section VI.

IV. ADAPTIVE LINEAR PREDICTION (ALP)

The solution (17) obviously depends on information that is not available, so that we need to resort to adaptive solutions. Structurally, infinite-duration impulse response (IIR) and finite-duration impulse response (FIR) solutions are

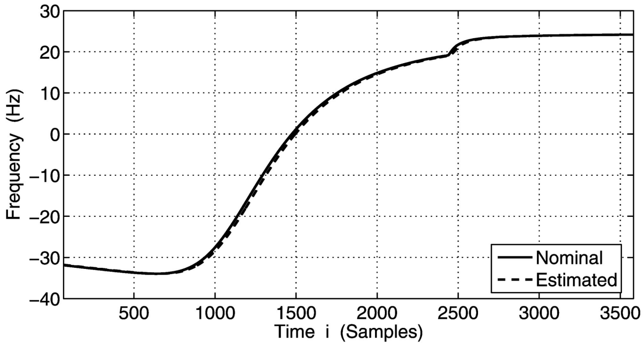


Fig. 3. Frequency estimate $\hat{f}^o(i)$ obtained from (17) and (6).

available, as well as LP-equivalent designs, such as the adaptive line enhancer (ALE) and the adaptive notch filter (ANF), covering AR, MA, and ARMA formulations [12, 13]. We rely on adaptive FIR predictors due to their inherent robustness. Another fundamental reason is the easiness with which FIR predictors can be combined to improve tracking performance, as shown later in Section V.

Adaptive FIR predictors can be obtained recursively as follows:

$$c_i = c_{i-1} + \mu \hat{R}_{u,i}^{-1}(\hat{R}_{du,i} - \hat{R}_{u,i}c_{i-1}) \tag{18}$$

where μ is a step-size parameter. Different approximations for the data-dependent covariance matrices $\hat{R}_u$ and $\hat{R}_{du}$ lead to different adaptive predictors with different abilities [21]. Once an adaptive predictor is chosen, a root-solver can be employed to retrieve the Doppler component estimates $\hat{f}(i)$ by using (5) and (6).

A set of adaptive algorithms that is quite relevant for the frequency tracking problem is the affine projection family. These algorithms use the following approximations for the covariance and cross-covariance matrices in (18) (refer to (14), (15) and (11)):

$$\hat{R}_{u,i} = U_i^* U_i + \epsilon I \qquad \text{and} \qquad \hat{R}_{du,i} = U_i^* d_i \tag{19}$$

where

$$U_i = \mathrm{col}\{x_{i-1}, x_{i-2}, \ldots, x_{i-K}\}$$

$$d_i = \mathrm{col}\{x(i), x(i-1), \ldots, x(i-K+1)\}$$

where K is referred to as the order of the algorithm and ϵ is a small positive number. The resulting algorithm (18) is quite robust, allied with a good tracking ability; besides, the computational complexity can be balanced by properly choosing the algorithm order K. For $K = 1$ we obtain the NLMS algorithm, which is employed due to its low complexity and good performance exhibited in the tests carried out. The NLMS update equation takes the form

$$c_i = c_{i-1} + \frac{\mu x_{i-1}^*}{\|x_{i-1}\|^2 + \epsilon}(x(i) - x_{i-1}c_{i-1}). \tag{20}$$

Another related algorithm that has been be tested in this context in [5] is the least-mean squares (LMS) filter:

$$c_i = c_{i-1} + \mu x_{i-1}^*(x(i) - x_{i-1}c_{i-1}). \tag{21}$$

Due to its performance we also consider for comparison the method proposed in [9], which employs a finite sliding window to find an RR-LS optimum predictor. The effect of reducing the rank is to mitigate the sensitivity of least-squares solutions to low SNR conditions (as happens in the EDL scenario). We employ a forward only least-squares predictor, obtained as follows:

$$H_{i-1}c_i = x_i^T \tag{22}$$

where the matrix H_{i-1} collects the regressors:[2]

$$H_{i-1} = \mathrm{col}\{x_{i-1}, x_{i-2}, \ldots, x_{i-N}\} \tag{23}$$

and N is the RR-LS order. The vector c_i is calculated from

$$c_i = [H_{i-1}^{(m)}]^\dagger x_i^T \tag{24}$$

where $H_{i-1}^{(m)}$ is an m-rank, $m \le M$, approximation for H_{i-1} and † denotes its pseudo-inverse. Forward-backward formulations may present superior performance [7, 9] but at a higher computational cost.

We extensively simulated the three main algorithms mentioned above, namely the NLMS, LMS, and RR-LS. Figs. 4–5 present two examples that compare the three algorithms using two representative sets of parameters. In Fig. 4 the LMS-based solution confirms its robustness and achieves roughly the same performance as the RR-LS algorithm, albeit at a much lower computational cost. Fig. 5 shows a great deal of improvement for the RR-LS with the new set of parameters and a dramatic performance decay for the LMS algorithm. In the two examples the NLMS outperforms both algorithms.

Fig. 6 shows a comparison of the RR-LS, LMS, and NLMS mean-square performance covering the 10 dB-Hz to 30 dB-Hz range.[3] The same set of parameters as Figs. 4 and 5 is employed. Note how the NLMS predictor outperforms both the LMS and the RR-LS algorithms. In summary, the LMS algorithm is quite sensitive to parameter design and to low SNR conditions and is not fast enough to cope with the channel effects. When properly tuned the RR-LS can be improved, but it may still lose frequency lock, especially in periods of high dynamics.

[2]The $\mathrm{col}\{\cdot\}$ operator stacks scalars or row vectors on top of each other [21].

[3]Considering the noise over the full spectrum and since we employ $F_s = 100$ Hz in this work, this is equivalent to the −10 dB to 10 dB range. In addition, unless otherwise stated, in this work the root mean-square frequency error (RMSE) plots are normalized by $\|f\|^2$, the norm of the nominal Doppler profile vector (refer to Fig. 2).

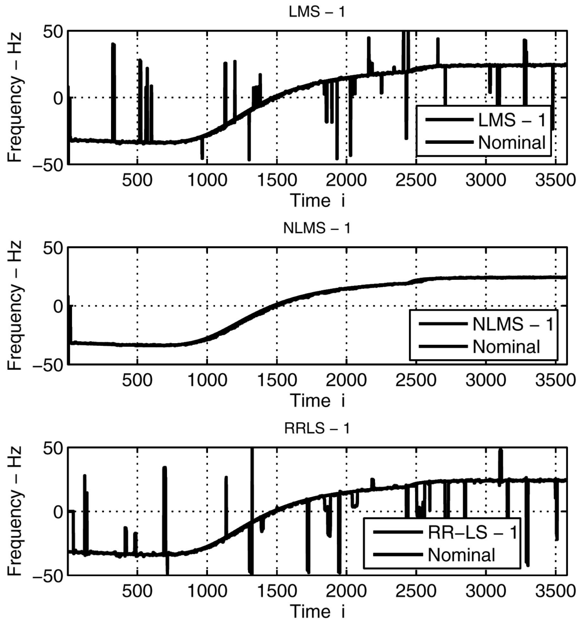


Fig. 4. Example 1: Comparison of typical algorithms for ALP frequency estimation with SNR = −6 dB. For LMS: $M_{\mathrm{LMS1}} = 15$, and $\mu_{\mathrm{LMS1}} = 0.04/M_{\mathrm{LMS1}}$. For RR-LS: $M_{\mathrm{LS1}} = 15$, and $N_{\mathrm{LS1}} = 40$. For NLMS: $M_{\mathrm{NLMS}} = 15$, and $\mu_{\mathrm{NLMS}} = 0.11$.

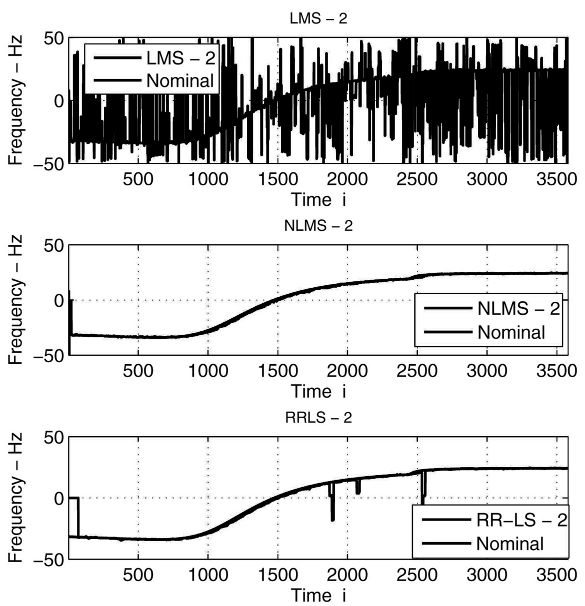


Fig. 5. Example 2: Comparison of typical algorithms for ALP frequency estimation with SNR = −6 dB. For LMS: $M_{\mathrm{LMS2}} = 15$, and $\mu_{\mathrm{LMS2}} = 0.09/M_{\mathrm{LMS2}}$. For RR-LS: $M_{\mathrm{LS2}} = 20$, and $N_{\mathrm{LS2}} = 70$. The NLMS predictor is run with the parameters from example 1.

V. ADAPTIVE COMBINATION SCHEMES

In general, choosing good predictor parameters (such as M and μ) is a difficult task and may depend on information not available, such as the frequency profile itself, not to mention that in a nonstationary environment, the optimal parameter set changes over time. As seen in the previous section, the NLMS is more robust than the RR-LS and the LMS, however it is also sensitive to changes in filter parameters. The impact of slightly different designs can be seen in Fig. 7. Two different, but similar, NLMS predictors are designed in an environment with SNR = 10 dB ($\sigma_v^2 = 0.1$). The first predictor employs $M_1 = 7$ and $\mu_1 = 0.7$, and the second predictor employs $M_2 = 10$ and $\mu_2 = 0.5$. Observe how a slight change in the design parameters leads to dramatically different performances. As indicated by the bottom plot of Fig. 7, a good design enables the adaptive filter to deliver a good prediction, thus leading to good frequency estimates.

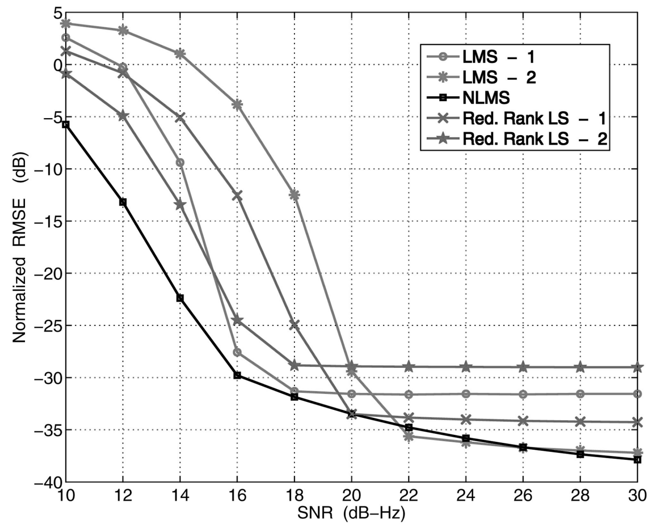


Fig. 6. Comparison of typical algorithms for ALP frequency estimation. Same set of parameters as in Figs. 4–5.

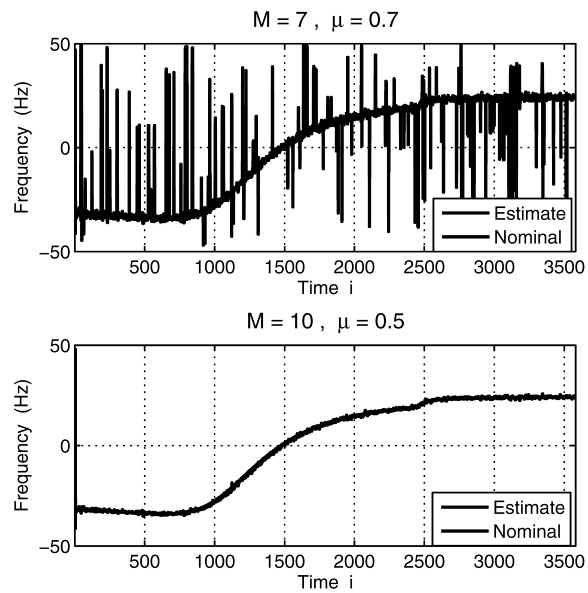


Fig. 7. Sensitivity of ALP solution (20).

To get around the sensitivity issue, instead of attempting to design the system parameters under little or no information, we change the design paradigm by adopting a combination approach

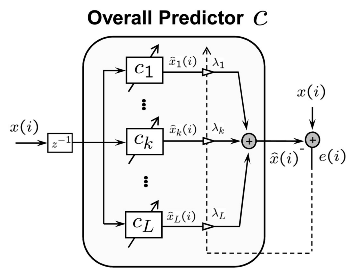


Fig. 8. Adaptive combination scheme (25).

[14, 19, 20]. The goal is to employ a mixture of multiple individual predictors that span a reasonable range of the unknown parameters. The individual predictors' outputs are then efficiently combined by a convex supervisor such that the global system is able to perform at least as well as the best individual predictor [22].

A. Convex Combiners

Before presenting the error prediction filter $Q_i(z)$ to the root solver, the quality of prediction can be improved by combining a span of L different normalized LMS predictors, with different orders and step-sizes [19, 20]. The individual predictors are combined according to their individual performance as [14]

$$c_{i-1} = \sum_{k=1}^{L} \lambda_k c_{k,i-1} \tag{25}$$

and

$$\hat{x}(i) = \sum_{k=1}^{L} \lambda_k \hat{x}_k(i), \qquad \sum_{k=1}^{L} \lambda_k = 1 \tag{26}$$

as depicted in Fig. 8, where the kth predictor has order M_k and step-size μ_k:

$$\hat{x}_k(i) = x_{i-1} c_{k,i-1}. \tag{27}$$

The regressor x_{i-1} in (27) has the order M_L of the largest predictor and is employed by all individual predictors c_k (extended with zeros in (27) to match the vector dimensions whenever necessary). The key step is to select the combiners $\{\lambda_k\}$. To ensure convexity of the predictors' aggregation, we use combination coefficients of the form:

$$\lambda_k = \frac{y_k}{\sum_{\ell=1}^{L} y_\ell}, \qquad y_k = f(a_k) \tag{28}$$

where y_k is a real activation function of some complex argument a_k. This formulation is an extension of the original work [14] to the arbitrary order and arbitrary activation function case and to handle complex signals. The function $y_k = f(a_k)$ is at our choice and the coefficient a_k can be dynamically adapted as [14]

$$a_k(i) = a_k(i-1) - \mu_a [\nabla_{a_k} |e(i)|^2]^*_{a_k = a_k(i-1)} \tag{29}$$

where

$$e(i) = x(i) - \hat{x}(i) \tag{30}$$

is the overall prediction error, with $\hat{x}(i)$ given by (26). It can be seen that, for a generic function $y_k = f(a_k)$, we obtain

$$\begin{aligned} \nabla_{a_k} |e(i)|^2 &= -e^*(i) x_{i-1} (c_k - c) \frac{\partial y_k}{\partial a_k} \cdot \frac{1}{\sum_\ell y_\ell} \\ &= -e^*(i)(\hat{x}_k(i) - \hat{x}(i)) \frac{\partial y_k}{\partial a_k} \cdot \frac{1}{\sum_\ell y_\ell}. \end{aligned} \tag{31}$$

A number of different combining functions can be used to aggregate the L experts [19]:

$$y_k = |e^{-a_k/2}|^2 \tag{32}$$

$$y_k = e^{-|a_k|^2} \tag{33}$$

$$y_k = |a_k|^2. \tag{34}$$

Alternative strategies may also be applied, for instance, by combining different adaptive algorithms [16, 17].

B. Hierarchical Convex Combiners

In [19] we employed a convex mixture of individual predictors to overcome the design sensitivity of the root configuration with respect to the predictor parameters. NLMS predictors, with orders M_k and step-sizes μ_k, are organized into a single combination layer. The single layer configurations are able to perform as well as the best individual predictor [19]. However, when the number of filters L is increased, the extra gradient noise introduced by the convex combiners may compromise the overall performance. Results in [19] indicate that the $L = 2$ convex combination performs similar to structures containing more predictors; in part that is due to the presence of one single combiner λ per pair of predictors, as opposed to one combiner per predictor for $L > 2$. This approach decreases the overall gradient noise. However only two filters may not cover the necessary range of parameters to capture the rich dynamics of the EDL events. To overcome the tradeoff we may explore a hierarchical arrangement of $L = 2$ convex prediction cells [20], or L2 cells for short (see Fig. 9). By doing so we still operate using one combiner λ only per pair of predictors, while embedding more predictors into the system. As a design example we consider $L = 4$ predictors: one low-order L2 cell to respond quickly during periods of higher dynamics and a higher order L2 cell to improve the system prediction capabilities in periods of lower dynamics. The outputs of both input cells are aggregated by an output L2 cell that automatically balances the high and low order input cells. The topology can be generalized to an arbitrary number of predictors by organizing them into cells and layers.

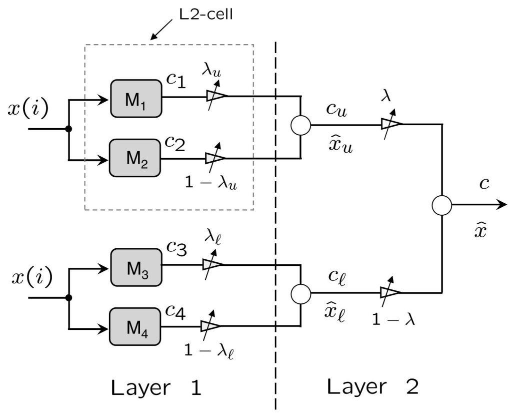


Fig. 9. Convex hierarchical scheme for $L = 4$ experts in the input layer. The order of kth NLMS predictor is denoted by M_k.

The structure in Fig. 9 is designed as follows. We first design the L2 cell at the output. For $L = 2$, only one combiner λ is needed, and the overall predictor is a convex combination of its input predictors:

$$c_{i-1} = \lambda(i)c_{u,i-1} + (1-\lambda(i))c_{\ell,i-1} \tag{35}$$

where $c_{u,i-1}$ and $c_{\ell,i-1}$ are the convex predictors for the upper and lower L2 cells, respectively, and where they are defined in terms of the individual predictors c_k in the input layer:

$$c_{u,i-1} = \lambda_u(i)c_{1,i-1} + (1-\lambda_u(i))c_{2,i-1} \quad \text{(upper)} \tag{36}$$

$$c_{\ell,i-1} = \lambda_\ell(i)c_{3,i-1} + (1-\lambda_\ell(i))c_{4,i-1} \quad \text{(lower).} \tag{37}$$

One adaptive rule to train the output L2 cell is obtained from (29) considering one combiner only [19, 20]:

$$a(i) = a(i-1) - \mu_a[\nabla_a|e(i)|^2]^*_{a=a(i-1)}. \tag{38}$$

The gradient in (38) is given by

$$\begin{aligned}\nabla_a|e|^2 &= \frac{\partial e^* e}{\partial e}\cdot\frac{\partial e}{\partial a} = e^*(-x)\cdot\frac{\partial c}{\partial a}\\ &= -e^* x\left(c_u\frac{\partial\lambda}{\partial a} - c_\ell\frac{\partial\lambda}{\partial a}\right)\\ &= -e^*(i)(\hat{x}_u(i)-\hat{x}_\ell(i))\cdot\frac{\partial\lambda}{\partial a}\end{aligned} \tag{39}$$

with

$$\hat{x}_u(i) = x_{u,i-1}c_{u,i-1}, \qquad \hat{x}_\ell(i) = x_{\ell,i-1}c_{\ell,i-1} \tag{40}$$

which leads to

$$a(i) = a(i-1) + \mu_a e(i)(\hat{x}^*_u(i) - \hat{x}^*_\ell(i))\cdot\frac{\partial\lambda}{\partial a}. \tag{41}$$

We are free to choose the activation function $\lambda(a)$, as previously mentioned. A function that presents good performance is [19, 20]:

$$\lambda(i) = \frac{1}{1+|e^{-a(i-1)/2}|^2}. \tag{42}$$

The adaptive rule for the L2 output cell then becomes

$$a(i) = a(i-1) + \mu_a e(i)(\hat{x}_u(i)-\hat{x}_\ell(i))^*\lambda(i)(1-\lambda(i)). \tag{43}$$

The global error $e(i)$ is given by (30), only now it employs the hierarchical predictor (35). The regressors of the upper and lower cells $x_{u,i-1}$ and $x_{\ell,i-1}$ have the same order as the predictors c_2 and c_4, respectively.

The upper and lower L2 cells can be designed in the same manner as the output layer. Therefore their combiners are given by

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

and the corresponding adaptive rules are

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

where now

$$e_u(i) = x(i) - \hat{x}_u(i) \quad \text{and} \quad e_\ell(i) = x(i) - \hat{x}_\ell(i). \tag{44}$$

The last step is to design the individual predictors, which are trained using the NLMS algorithm:

$$c_{k,i} = c_{k,i-1} + \mu_k\frac{x^*_{k,i-1}}{\|x_{k,i-1}\|^2+\epsilon}(x(i) - x_{k,i-1}c_{k,i-1}) \quad \text{for} \quad k = 1,\ldots,L. \tag{45}$$

In addition, the combiners $\lambda(i)$, $\lambda_u(i)$, and $\lambda_\ell(i)$ are time-smoothed over their past N_{ham} values via a Hamming half-window w_{ham}, thus generating $\bar{\lambda}(i)$, $\bar{\lambda}_u(i)$, and $\bar{\lambda}_\ell(i)$. For instance for $\bar{\lambda}(i)$ we have

$$\bar{\lambda}(i) = \lambda_i w_{\text{ham}} \tag{46}$$

where

$$\lambda_i = [\lambda(i) \quad \lambda(i-1) \quad \lambda(i-2)\cdots\lambda(i-N_{\text{ham}}+1)]$$
$$w_{\text{ham}} = [w(0) \quad w(1) \quad w(2)\cdots w(N_{\text{ham}}-1)]^T.$$

More recent samples are emphasized so that the window peak is at $w(0)$, which corresponds to the current sample. The window coefficients $\{w(\cdot)\}$ are normalized to sum up to unity. This procedure helps combat the extra gradient noise introduced by the learning rules $a(i)$, $a_u(i)$, and $a_\ell(i)$.

Fig. 10 presents a time snapshot of the combiners evolution for SNR = 14 dB-Hz and for $\mu_a = \mu_u = \mu_\ell = 5$. Note how the combiners are correctly assigned: in periods of relative low activity, the higher order predictors are selected since they lead to smaller errors. For periods with high dynamics, the lower order predictors are selected once they are able to react faster to changes in the channel conditions.

We ran a simulation example covering the whole SNR range of interest, namely 10 dB-Hz to 30 dB-Hz (-10 dB to 10 dB), with $M_k = \{9, 13, 17, 21\}$ and

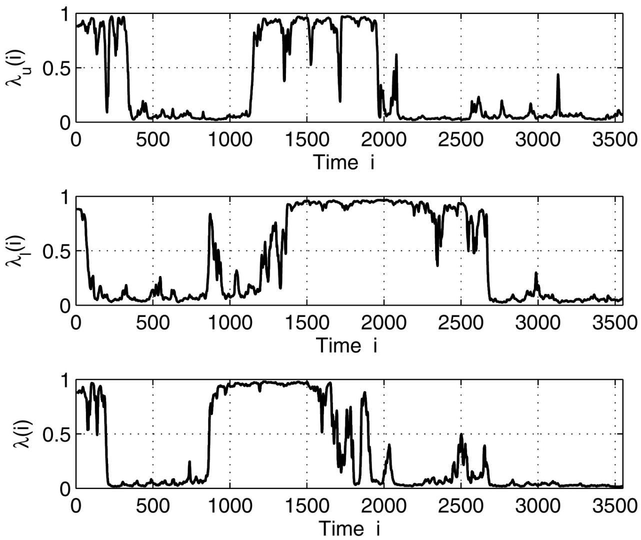


Fig. 10. Time snapshot of convex coefficients for SNR = 14 dB-Hz (−6 dB).

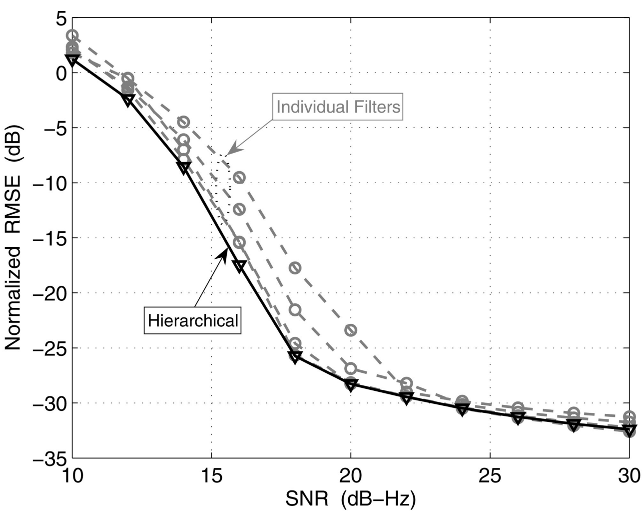


Fig. 11. Universality of combination scheme for $L = 4$.

$\mu_k = 0.11$. As Fig. 11 depicts the hierarchical scheme achieves universality over the full SNR range.

VI. LAG-ERROR COMPENSATION

The estimation error in nonstationary scenarios is comprised of two components: variance and bias. For LMS-like algorithms the variance decreases as the step-size decreases. However the lag error (bias) decreases with larger step-sizes since the filter becomes faster and since it is able to keep up with the time-varying nature of the underlying process. As a result there is a trade-off in terms of the step-size. However, due to the low SNR levels of the received signals at the Earth end, we cannot employ large step-sizes. This is because the loss of frequency lock, represented by the large spikes in the frequency estimates, as well as the variance component, would worsen considerably. One needs to compensate the lag

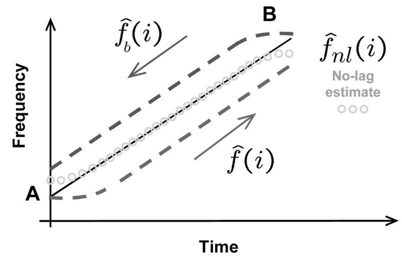


Fig. 12. Principle of lag-error compensation.

error in some other way. One way to achieve this goal is by exploring the learning latency of the adaptive filters [20].

For simplicity, consider a linear frequency profile $f(i)$ to illustrate the procedure (see Fig. 12). An adaptive filter will constantly try to follow the reference signal, but it falls behind due to the latency of the learning process. As a consequence, in the forward direction the resulting frequency estimates will suffer from a casual bias error, i.e., the predicted frequency profile $\hat{f}(i)$, resulting from point A to point B in Fig. 12, will lie underneath the nominal frequency profile $f(i)$. Now, if we perform a backward prediction, i.e., from point B to point A, a similar effect takes place: due to the latency of the learning process, the frequency profile estimate $\hat{f}_b(i)$ also suffers from a bias, which is causal with respect to the flipped time axis, since the prediction is backwards, but anticausal with respect to the original time axis, as one can see in Fig. 12. Therefore the anticausal bias from the backward estimates $\hat{f}_b(i)$ can be employed to compensate the causal bias from the forward prediction. A compensated frequency profile $\hat{f}_{nl}(i)$, where "nl" is a mnemonic for "no-lag," can be computed by a simple point-to-point average:

$$\hat{f}_{nl} = \tfrac{1}{2}(\hat{f} + \hat{f}_b). \tag{47}$$

Fig. 13 shows the efficiency of the procedure for the EDL frequency profile. Visibly the anticausal estimate compensates the causal bias error. In Fig. 14 we observe the benefits of the lag compensation scheme over the full SNR range. For low SNR the error due to loss of lock (i.e., spikes) is predominant, and therefore the gain is noticeable but not so expressive. In the higher SNR range, when most of the error arises from the bias, the compensation is expressive.

Other mappings, possibly nonlinear, may be applied in order to further enhance the estimates:

$$\hat{f}_{nl}(i) = g(\hat{f}, \hat{f}_b). \tag{48}$$

One such operator is the median filter, which could be explored to remove the spikes that arise by screening the forward and backward estimates simultaneously inside a short observation window since it is unlikely to encounter spikes in both forward and backward estimates at the same time instant.

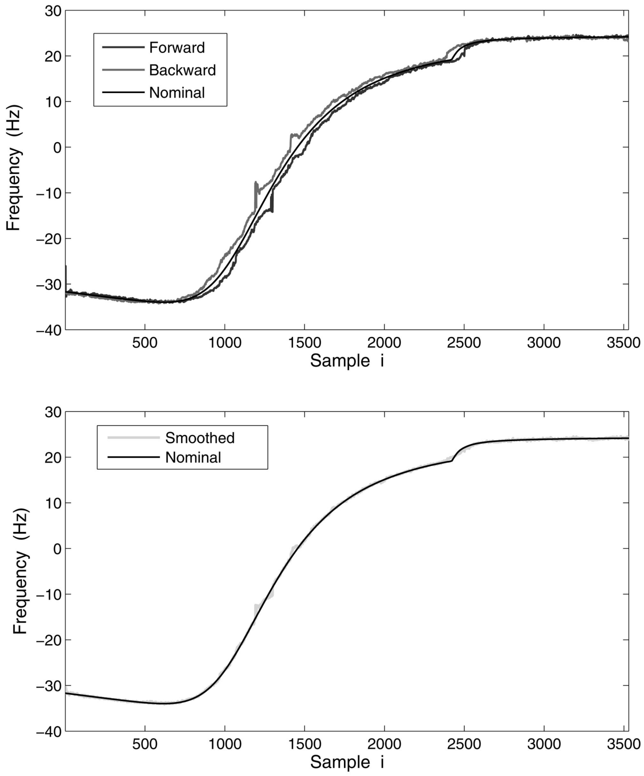


Fig. 13. Example of lag error compensation for typical EDL signals.

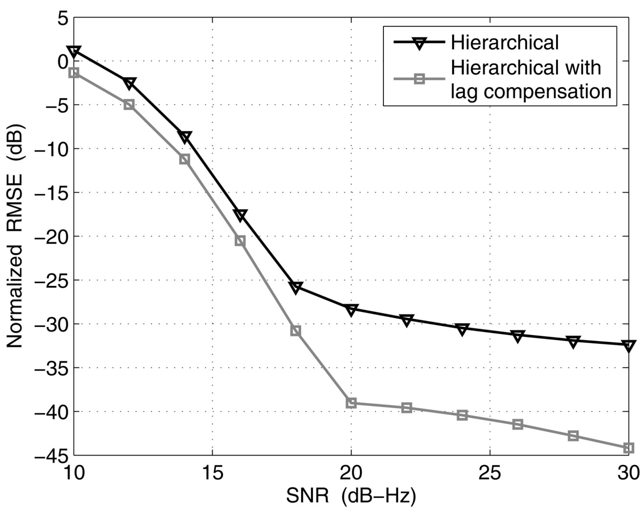


Fig. 14. Comparison: Hierarchical scheme with lag compensation and plain hierarchical scheme.

In Mars mission applications, due to the off-line nature of the estimation problem, this method can be properly employed. It could also be applied to block processing applications, where the partial frequency estimates are obtained within each data block.

VII. EFFICIENT FFT-BASED ESTIMATION

We have thus far employed a simple root solver to retrieve the frequency content from the coefficients of $Q_i(z)$. However the roots of that polynomial, when the digital filter is implemented in the direct form, are quite sensitive to disturbances in the filter coefficients, especially for high order filters, and this affects the performance of the frequency tracker based on the root solver, resulting in loss of lock.

It can be shown that, when the predictor c is optimally designed (LMS sense), the maximum entropy estimate of the input signal power spectrum is related to the optimal predictor error filter (5) via [5, 9, 10, 11]

$$\hat{S}_x(\omega) = \frac{B_0}{|Q(\omega)|^2} \tag{49}$$

where B_0 is a constant. As a consequence, to retrieve the narrow band Doppler component, we may alternatively search for a peak in $\hat{S}_x(\omega)$, or a notch in $|Q(\omega)|$, instead of seeking the closest root to the unit circle in $Q(z)$.

It turns out that, in the EDL channel, with very low SNR, searching for the Doppler component directly in the FFT-based instantaneous input power spectrum estimate $\hat{S}_{x,i}(\omega)$ results in a noticeable improvement in performance as compared with the root-solver approach. Thus, at each time instant, we look for a notch in the magnitude-squared FFT of the instantaneous predictor error filter $Q_i(\omega)$. Intuitively, the FFT tends to concentrate the narrow band component (present in the predictor coefficients) around the true frequency component and to evenly distribute the noise component along the spectrum. That is, besides the fact that the adaptive filter captures the underlying power spectrum of the input signal in its coefficients, the FFT further enhances the peak location. In addition, as the filter order increases, the FFT computational efficiency has advantages over the operation of extracting roots of a high order polynomial.

A time snapshot of both root solver and FFT methods is presented in Fig. 15, where the robustness of the FFT extractor is clear. Fig. 16 compares the root solver and the FFT extractor for the 10 dB-Hz to 30 dB-Hz SNR range (-10 dB to 10 dB). The improvement in frequency MSE at the low SNR range is considerable (from 5 dB to roughly 10 dB). At higher SNRs the effect is still noticeable (with advantage to the FFT extractor), but is not so expressive, which may explain why, in the literature, the two methods seem to be tacitly assumed as similar: the SNR in usual applications is (much) higher than in the EDL context, and the two methods are roughly similar in such a range.

Fig. 17 presents the improvement achieved with joint FFT extraction and lag compensation. Note the substantial improvement in MSE over the full SNR range, above 10 dB in most of the range, with a peak of 15 dB.

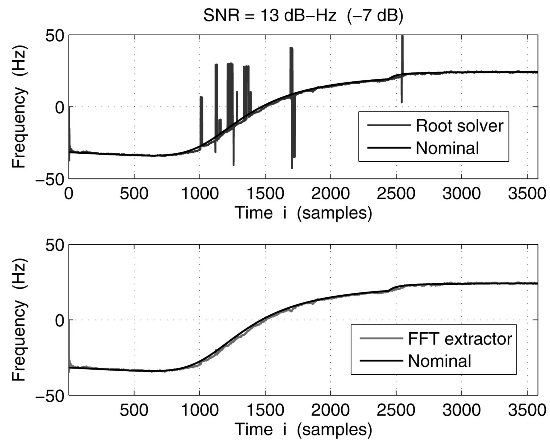


Fig. 15. Comparison: Time snapshots of root solver (top) and FFT extractor (bottom).

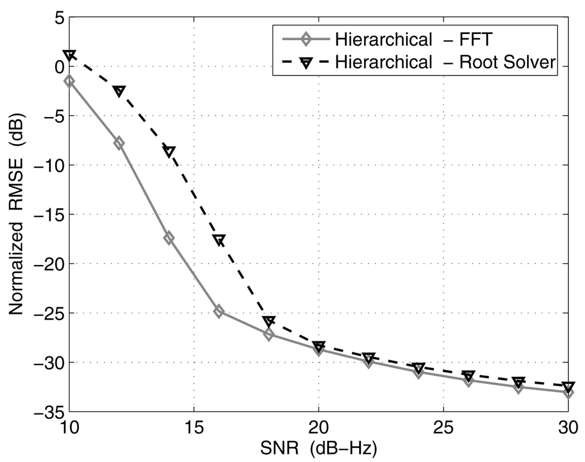


Fig. 16. Comparison between root solver and FFT extractor.

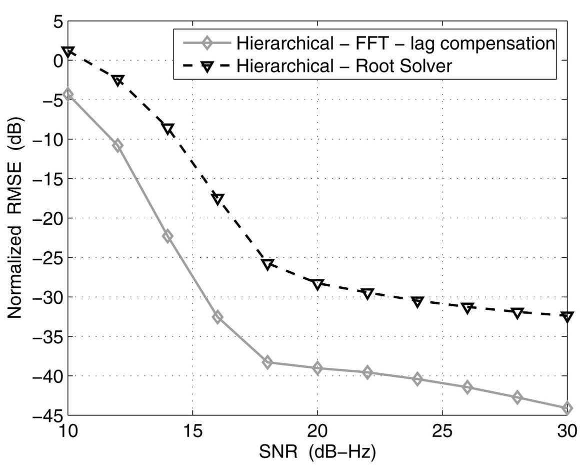


Fig. 17. Improved ALP performance with joint FFT smoothing and lag compensation.

VIII. ENHANCEMENT TECHNIQUES

The channel noise impact in the ALP sensitivity can be increased by the stochastic gradient disturbances introduced by the predictors (20) as well as by the adaptive combiners (38). This may cause spikes in the estimated frequency that are not correlated with the underlying true frequency $f(i)$, as Fig. 7 shows. This phenomenon is usually referred to as loss of tracking lock, or as a loss of frequency continuity, and simple methods that enforce frequency continuity can be incorporated to improve the tracking ability of the overall scheme. The underlying assumption is that the dynamics of the lander is continuous in time, which implies a continuous Doppler profile as well.

A. Order-Adaptive Q-Zeroing Smoothing

Due to the channel conditions, at every time instant, the performance of filters with different orders may vary considerably. If we perform an experiment with multiple ALPs running independently in a low SNR scenario, filters with different orders eventually lose the tracking lock at different time instants (as observed in simulations). In other words at any time instant, it is usually possible to find a filter that has not lost lock. In order to explore this fact, we devise a simple strategy that, at every iteration, searches for a filter that maintains frequency continuity relative to the previous iteration. We do so by sampling the p leading entries of the convex combination vector c (refer to (35)), which emulates a vector c_p in a lower dimension subspace:

$$c_p = [c]_{1:p}, \qquad p \in [p_{\min}, M] \tag{50}$$

where M is the order of the predictor c. The corresponding "order-reduced" error predictor vector[4] Q_p (see (5)) is then presented to the FFT extractor, and frequency continuity is checked. If discontinuity is detected[5] the order p is decreased, and the continuity of the new Q_p is rechecked. This process is repeated until the continuity criterion is met, or until the predefined lower bound $p_{\min}$ is achieved. In case either condition is met and the spike persists, the error is treated by the next control routine (refer to Subsection VIIIB). This strategy is a simple way to emulate a search for the filter order that preserves frequency continuity without the requirement of implementing several filters simultaneously.

Fig. 18 presents two instantaneous runs for $p_{\min} = 5$, $M = 25$ and $\mu = 0.11$ at SNR = 10 dB-Hz (-10 dB). The improvement in frequency tracking lock is expressive. The scheme presented in this section can combat the majority of the frequency spikes, which implies an improvement in the probability of frequency error. As a result it quite reasonably improves the MSE (roughly 7 dB) in the

[4]This is implemented by zeroing the trailing $M - p$ entries of Q.

[5]Loss of frequency lock is declared whenever the frequency difference from previous iteration is beyond a predefined threshold THR, usually chosen as a multiple of the maximum possible frequency difference (an estimate). Another possibility is to set THR as a fraction of the sampling frequency F_s.

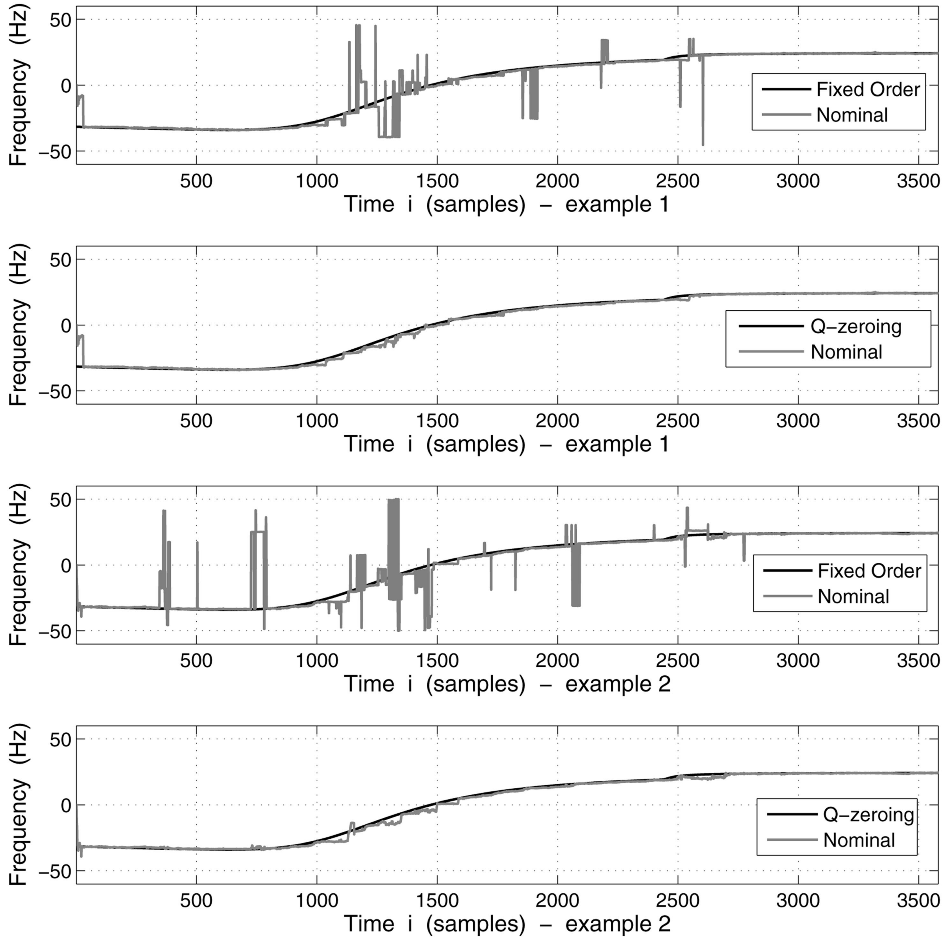


Fig. 18. Comparison: Time snapshots of fixed order versus Q-zeroing smoothing for SNR = 10 dB-Hz (−10 dB), $p_{\min} = 5$, $M = 25$, and $\mu = 0.11$.

very low to low SNR range (10 dB-Hz to 20 dB-Hz, or −10 dB to 0 dB), as depicted in Fig. 19. Fig. 19 depicts the frequency error by comparing both fixed order ($M = 25$) and Q-zeroing smoothing (same parameters as Fig. 18).

B. Derivative Control

If, after the Q-zeroing reprocessing routine, frequency spikes are still found, we may resort to a derivative control strategy [19], once more assuming that the underlying lander dynamics (e.g., velocity) are continuous in time. In other words, Doppler frequency jumps result from a tracking failure since the lander does not change its momentum instantaneously.

The control strategy works as follows. A derivative buffer keeps track of the average $\overline{\delta f}(i)$ of the last N_d "good" derivative samples (defined ahead). Whenever a spike cannot be addressed by the Q-zeroing routine, the derivative control forces the time continuity of the frequency estimates by updating the previous (good) frequency estimate $\hat{f}(i-1)$ with the average of the last N_d "good" derivative samples, which captures the tendency of the underlying physical process. Let

$$\delta\hat{f}(i) \triangleq \hat{f}(i) - \hat{f}(i-1) \tag{51}$$

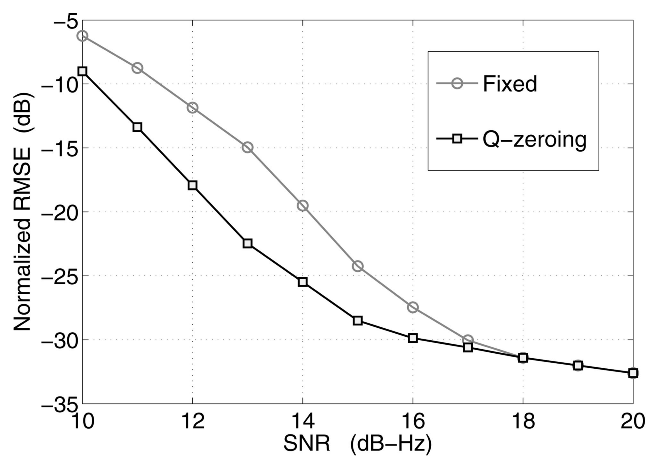


Fig. 19. Comparison: Fixed order versus Q-zeroing smoothing for $p_{\min} = 5$, $M = 25$, and $\mu = 0.11$.

and for a given THR, define a good derivative sample D_b as

$$D_b = \begin{cases} \delta\hat{f}(i), & \text{if} \quad |\delta\hat{f}(i)| \leq \text{THR} \\ \gamma \cdot \text{sign}(\delta\hat{f}(i)), & \text{if} \quad |\delta\hat{f}(i)| > \text{THR} \end{cases} \tag{52}$$

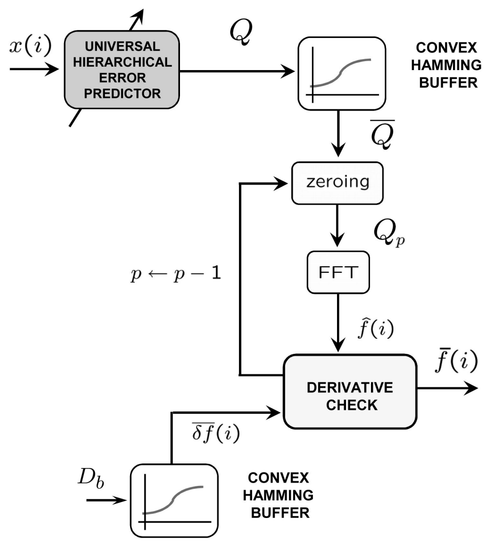


Fig. 20. Convex combination scheme with order-adaptive smoothing and derivative control.

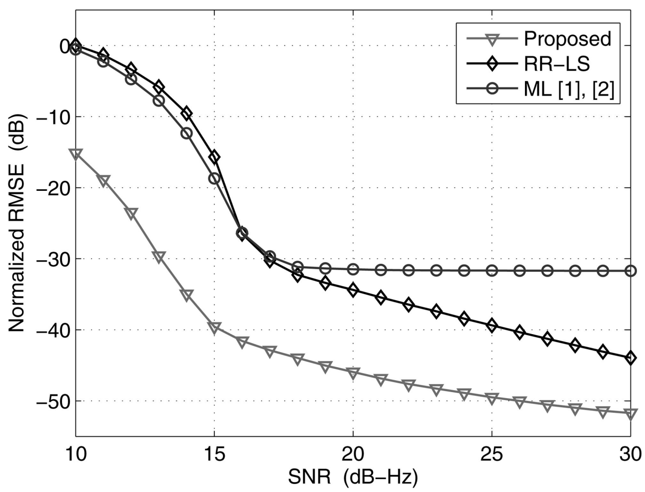


Fig. 21. Comparison of all algorithms. $L = 4$, $M_1 = 9$, $M_2 = 13$, $M_3 = 17$, $M_4 = 21$, $\mu_k = 0.11$, and $\mu_a = 0.25$. 100 experiments.

where $\gamma \ll$ THR. The derivative buffer is always fed with D_b. Whenever the derivative is bigger than the THR, the sign information is kept, but the magnitude is clamped, which improves the average of the derivatives. The enhanced frequency estimate $\bar{f}(i)$ is then given by

$$\bar{f}(i) = \begin{cases} \hat{f}(i), & \text{if} \quad |\delta\hat{f}(i)| \leq \text{THR} \\ \hat{f}(i-1) + \overline{\delta f}(i), & \text{if} \quad |\delta\hat{f}(i)| > \text{THR} \end{cases}. \tag{53}$$

IX. CONCLUSION

Fig. 20 shows the block diagram of the proposed system and gathers all the subsystems developed throughout the sections.

This work devises an adaptive hierarchical combination of the NLMS adaptive prediction structure that is assisted by frequency lock routines and with built-in FFT-based frequency retrieval and added lag-error compensation. Fig. 21 compares the approximate ML-like approach of [1, 2], the forward RR-LS implementation of [9], and the proposed scheme. It is seen that the ALP solution shows improved performance. The new system exhibits a frequency root MSE improvement over a wide SNR range.

ACKNOWLEDGMENT

The authors would like to acknowledge the useful discussions and references suggested by Dr. Andre Tkacenko from JPL.

REFERENCES

[1] Hurd, W. J., Estabrook, P., Racho, C. S., and Satorius, E. Critical spacecraft-to-earth communications for Mars exploration rover (MER) entry, descent and landing. In *Proceedings of the IEEE Aerospace Conference*, vol. 3, Mar. 2002, 1283–1292.

[2] Satorius, E., Estabrook, P., Wilson, J., and Fort, D. 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.

[3] Vilnrotter, V. A., Hinedi, S., and Kumar, R. Frequency estimation techniques for high dynamics trajectories. *IEEE Transactions on Aerospace and Electronic Systems*, **25**, 4 (July 1989), 559–577.

[4] Rife, C. D. and Boorstyn, R. R. Single-tone parameter estimation from discrete time observations. *IEEE Transactions on Information Theory*, **IT-20** (Sept. 1974).

[5] Griffiths, L. J. Rapid measurement of digital instantaneous frequency. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-23**, 2 (Apr. 1975), 207–222.

[6] Nuttall, A. H. Spectral analysis of a univariate process with bad data points via maximum entropy and linear predictive techniques. In *Naval Underwater System Scientific and Engineering Studies* (NUSC) Spectral Estimation, NUSC, New London, CT, Mar. 1976.

[7] Ulrych, T. J. and Clayton, R. W. Time series modeling and maximum entropy. *Physics of the Earth and Planetary Interiors*, **12** (Aug. 1976), 188–200.

[8] Rife, C. D. and Boorstyn, R. R. Multiple tone parameter estimation from discrete time observations. *Bell System Technical Journal*, (Nov. 1976), 1389–1410.

[9] Tufts, D. W. and Kumaresan, R. Estimating the parameters of multiple sinusoids: Making linear prediction perform like maximum likelihood. *Proceedings of the IEEE*, **70**, 9 (Sept. 1982), 975–989.

[10] Peacock, K. L. and Treitel, S. Predictive deconvolution: Theory and practice. *Geophysics*, **34** (Apr. 1969), 155–169.

[11] Jackson, L. B., Tufts, D. W., Soong, F. K., and Rao, R. M. Frequency estimation by linear prediction. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, (Apr. 1978), 352–356.

[12] Chambers, J. A. and Constantinides, A. G. Frequency tracking using constrained adaptive notch filters synthesized from allpass sections. *Proceedings of the IEE*, **137**, 6 (Dec. 1990), 475–481.

[13] Friedlander, B. A recursive maximum likelihood algorithm for ARMA line enhancement. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-30**, 4 (Aug. 1982), 651–657.

[14] Arenas-Garcia, J., Figueiras-Vidal, A. R., and Sayed, A. H. Mean-square performance of a convex combination of two adaptive filters. *IEEE Transactions on Signal Processing*, **54**, 3 (Mar. 2006), 1078–1090.

[15] Zhang, Y. and Chambers, J. A. Convex combination of adaptive filters for a variable tap-length LMS algorithm. *IEEE Signal Processing Letters*, **13**, 10 (Oct. 2006), 628–631.

[16] Silva, M. T. M. and Nascimento, V. H. Convex combination of adaptive filters with different tracking abilities. In *Proceedings of the IEEE ICASSP*, vol. 3, Honolulu, HI, Apr. 2007, 925–928.

[17] Silva, M. T. M. and Nascimento, V. H. Improving the tracking capability of adaptive filters via convex combination. *IEEE Transactions on Signal Processing*, **56**, 7 (July 2008), 3137–3149.

[18] Mandic, D., Vayanos, P., Boukis, C., Jelfs, B., Goh, S. L., Gautama, T., and Rutkowski, T. Collaborative adaptive learning using hybrid filters. In *Proceedings of the IEEE ICASSP*, vol. 3, Honolulu, HI, Apr. 2007, 921–924.

[19] Lopes, C. G., Satorius, E., and Sayed, A. H. Adaptive carrier tracking for Direct-to-Earth Mars communications. In *Proceedings of the 40th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2006, 1042–1046.

[20] Lopes, C. G., Satorius, E. H., Eastbrook, P., and Sayed, A. H. Efficient adaptive carrier tracking for Mars to Earth communications during entry, descent, and landing. In *Proceedings of the IEEE Workshop on Statistical Signal Processing* (SSP), Madison, WI, Aug. 2007, 517–521.

[21] Sayed, A. H. *Fundamentals of Adaptive Filtering*. Hoboken, NJ: Wiley, 2003.

[22] Merhav, N. and Feder, M. Universal prediction. *IEEE Transactions on Information Theory*, **44** (1998), 2124–2147.

Cassio G. Lopes (S'06—M'08) received his B.S. degree and M.S. degrees in electrical engineering from the Federal University of Santa Catarina, Brazil in 1997 and 1999, respectively, and M.S. and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA) in 2004 and 2008, respectively.

From 1999 to 2002 he worked with hardware design for biomedical signals acquisition and transmission over the Internet. He also taught undergraduate-level courses in computer science and electrical engineering. From 2005 to 2007 he worked in cooperation with the NASA Jet Propulsion Laboratory to develop frequency tracking algorithms to support direct-to-Earth Mars communications during entry, descent, and landing. In 2008 he worked at the Instituto Tecnologico de Aeronautica (ITA) Sao Jose dos Campos, Brazil as a postdoctoral researcher, developing distributed inertial navigation systems for aerospace applications. He is currently with the Department of Electronic Systems of the University of Sao Paulo (USP), Escola Politecnica, where he is an assistant professor of electrical engineering. His current research interests are theory and methods for adaptive and statistical signal processing, distributed adaptive estimation and data fusion, as well as frequency tracking algorithms.

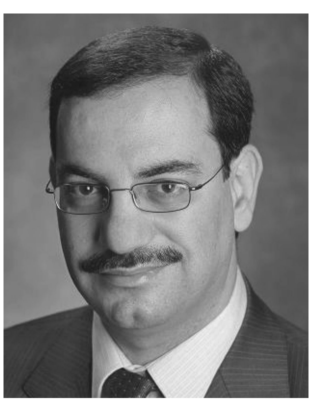
Edgar Satorius received his B.Sc. in engineering from the University of California, Los Angeles and M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA.

He is a principal member of the technical staff, Flight Communications Systems Section of the Jet Propulsion Laboratory, Pasadena, CA. He performs systems analysis in the development of digital signal processing and communications systems with specific applications to blind demodulation, digital direction finding, and digital receivers. He has published over 90 articles and holds two patents in the field of digital signal processing and its applications. In addition he is an adjunct associate professor at the University of Southern California, where he teaches digital signal processing courses.

Polly Estabrook received her B.A. in engineering physics from the University of California, Berkeley, and M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA.

She is the deputy manager of the Communication Architectures and Research Section at JPL. Since 2005 she has lead various design teams to define the modifications necessary for NASA's Space Communication and Navigation infrastructure to support the planned manned missions to the Moon and Mars. From 2000 to 2004 she was the lead telecom system engineer for the Mars Exploration Project, responsible for the performance of the entry descent and landing telecommunications system and for the overall design and performance of the Direct to Earth and relay communications systems. Her area of expertise lies in the design of deep space in-situ (lander to orbiter, sensor web) communications systems, near-Earth TDRS-based communications systems, and mobile satellite personal communication systems.

In 2004, Dr. Estabrook received the NASA Exceptional Achievement Medal for her work on the Mars Exploration Rover Telecom System. She has written over 35 technical papers and chaired numerous IEEE and AIAA conference sessions.



Ali H. Sayed received his Ph.D. degree from Stanford University, Stanford, CA, in 1992.

He is a professor of and the Chairman of Electrical Engineering at the University of California, Los Angeles (UCLA) and principal investigator of the Adaptive Systems Laboratory. He has published widely, with over 300 articles and five books in the areas of statistical signal processing, estimation theory, adaptive filtering, signal processing for communications and wireless networking, and fast algorithms for large structured problems. He is a coauthor of the textbook *Linear Estimation* (Prentice-Hall, 2000), coauthor of the research monograph *Indefinite Quadratic Estimation and Control* (SIAM, 1999), and coeditor of *Fast Algorithms for Matrices with Structure* (SIAM, 1999). He is also the author of the textbooks *Fundamentals of Adaptive Filtering* (Wiley, 2003) and *Adaptive Filters* (Wiley, 2008). He has contributed several encyclopedia and handbook articles.

Dr. Sayed has served on the editorial boards of the *IEEE Signal Processing Magazine*, the *European Signal Processing Journal*, the *International Journal on Adaptive Control and Signal Processing*, and the *SIAM Journal on Matrix Analysis and Applications*. He also served as the Editor-in-Chief of the *IEEE Transactions on Signal Processing* from 2003 to 2005 and the *EURASIP Journal on Advances in Signal Processing* from 2006 to 2007. He is a member of the Signal Processing for Communications and the Signal Processing Theory and Methods technical committees of the IEEE Signal Processing Society. He has served on the Publications (2003–2005), Awards (2005), and Conference (2007–present) Boards of the IEEE Signal Processing Society. He served on the Board of Governors (2007–2008) of the same society and is now serving as Vice-President of Publications (2009–present). His work has received several recognitions, including the 1996 IEEE Donald G. Fink Award, the 2002 Best Paper Award from the IEEE Signal Processing Society, the 2003 Kuwait Prize in Basic Sciences, the 2005 Terman Award, the 2005 Young Author Best Paper Award from the IEEE Signal Processing Society, and two Best Student Paper Awards at international meetings (1999 and 2001). He has served as a 2005 distinguished lecturer of the IEEE Signal Processing Society and as general chairman of ICASSP 2008.