ADAPTIVE CARRIER TRACKING FOR DIRECT-TO-EARTH MARS COMMUNICATIONS

C. G. Lopes, E. Satorius, and A. H. Sayed

ABSTRACT

We propose a robust and low complexity scheme to estimate and track carrier frequency from signals traveling under low SNR conditions in highly non-stationary 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 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, leading to a universal scheme for frequency estimation and tracking.

1. INTRODUCTION

In January 2004 the Mars Exploration Rovers (MER) Spirit and Opportunity successfully landed in Martian soil. Both missions were launched by NASA/JPL to reveal the water historical profile at lower latitudes in the red planet, along with other geophysical data. As part of the Mars continuous exploration effort, missions to come will encounter more severe conditions for communications, rendering tougher scenarios to send data back to Earth.

In the formidable logistic effort towards Mars, the entry, descent and landing (EDL) phase was the most crucial period of the mission [1, 2]. During this phase, a complex sequence of events take place, and health and status signals are sent back in real time through the direct-to-earth (DTE) channel. This information is fundamental to flag the mission status and consequently improve future designs in case of mission failure (*en passant*: during the Mars Pathfinder mission, the signal was temporarily lost).

In order to support spacecraft-to-earth communications during the EDL phase, we develop a robust and low complexity carrier frequency estimation and tracking technique that is able to operate under low SNR and highly non-stationary conditions, common to the adverse EDL scenario. The method combines the natural tracking abilities of adaptive filters with linear prediction techniques and universal prediction concepts, leading to encouraging results that can be applied to other standard frequency estimation problems as well.

2. DIRECT-TO-EARTH 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. At the spacecraft end, every 10 seconds, mission signals are sent 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 has been adopted by JPL [1], with a nominal carrier frequency of $f_c^0 = 8.4$ GHz (X-band) and employing a constellation of 256 possible symbols.

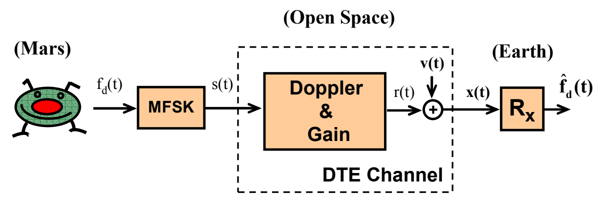

Fig. 1. Direct-to-Earth communications.

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 signal generation can be found in [2]. In order to recover the MFSK-data, we need a reliable carrier frequency estimate. The spacecraft high dynamics caused by the EDL procedures leads to severe Doppler shifts in the nominal carrier frequency [1]:

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

Due to the large Doppler component $f(t)$, it is assumed in this work that there is no embedded data and that the signal is down-converted and sampled upon reception, such that the only frequency content in the remaining signal is the Doppler component, i.e., $f_c(t) = f(t)$. In other words, the signal that we will be dealing with at the Earth end is of the form

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

where ω is the discrete time-varying Doppler component and $v(i)$ arises from an ergodic white process with variance σ_v^2. Our objective is to estimate and track ω from measurements $\{x(i)\}$.

3. THE ESTIMATION AND TRACKING SCHEME

3.1. Linear prediction

One approach for estimation and frequency tracking is to formulate a linear prediction problem. In linear prediction, the current value of $x(i)$ is predicted by linearly combining its observed past values:

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

This material was supported by the Jet Propulsion Laboratory under award 1276256 and by NSF award ECS-0601266. The first author was also supported by a fellowship from CAPES, Brazil, under award 1168/01-0. The authors Lopes and Sayed are with the Electrical Engineering Department, UCLA, Los Angeles, CA 90095. Email: {cassio,sayed}ee.ucla.edu. Dr. Satorius is with the Jet Propulsion Laboratory, Pasadena, CA 91109, Email: Edgar.H.Satorius@jpl.nasa.gov

The predictor's goal is to minimize the estimation error

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

in some sense. With the signal model (2), when the predictor coefficients $c(k)$ are optimally designed in the minimum mean-square error sense, the corresponding 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 that enables a simple frequency estimation procedure.

The optimum predictor vector c^o is obtained by solving

$$w^o = \arg\min_{c} E|x(i) - \hat{x}(i)|^2 \tag{6}$$

where

$$\hat{x}(i) = x_{i-1} c \tag{7}$$

and

$$x_{i-1} = \begin{bmatrix} x(i-1) & x(i-2) & \cdots & x(i-M) \end{bmatrix} \tag{8}$$

is the row observation vector. Using model (2), the optimal predictor can be shown to be [5]:

$$c^o = \frac{1}{\sigma_v^2 + M} \begin{pmatrix} e^{j\omega} \\ e^{j2\omega} \\ \vdots \\ e^{jM\omega} \end{pmatrix}. \tag{9}$$

As equation (9) reveals, the optimal prediction vector c^o contains information about the desired unknown frequency ω. Figure 2 depicts the root locus of the corresponding optimal error predictor filter $Q^o(z)$ as a function of signal-to-noise ratio (SNR), for $M = 10$ and $\omega = \pi/3$. One of the M roots, which we define as $r_o = \rho_o e^{j\theta_o}$, lies closer to the unit circle exactly at $\theta_o = \omega = \pi/3$. Note that as the SNR decreases, it becomes harder to find r_o, which is the root that ultimately delivers the desired frequency estimate.

3.2. Adaptive Linear Prediction (ALP)

Since the parameter of interest ω is non-stationary, we may employ an adaptive filter to equip the predictor with tracking abilities and perform the design in (6) automatically.

Among many possible adaptive algorithms [5], we have tested the LMS, the normalized LMS (ϵ-NLMS), affine projection and RLS variants, namely, the exponentially weighted RLS and the sliding window RLS. Among them, the ϵ-NLMS presents low complexity and has reported the best performance; its update equation is given by

$$c_i = c_{i-1} + \mu \frac{x_{i-1}^*}{\|x_{i-1}\|^2 + \epsilon} \big(x(i) - x_{i-1} c_{i-1} \big) \tag{10}$$

Figure 3 presents the ALP scheme. At each time i, the adaptive predictor presents the corresponding error prediction 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 frequency f at time i is then found from

$$\hat{f}(i) = \frac{\theta_o}{2\pi} \cdot F_s \quad \text{(Hz)} \tag{11}$$

where F_s is the sampling frequency.

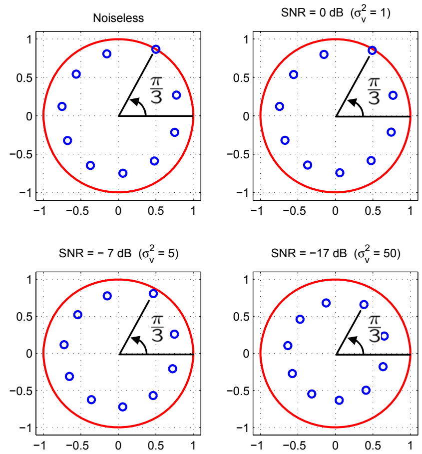

Fig. 2. Roots of the optimal error prediction polynomial $Q^o(z)$ for $M = 10$ and $\omega = \pi/3$.

4. A MODEL MIXTURE PREDICTION SCHEME

In general, choosing good predictor parameters (such as M and μ), is a difficult task. Not to mention that in a non-stationary environment the optimal parameter set may dynamically change over time. One can see the impact of slightly different designs in Figure 4. Two different - but similar - ϵ-NLMS predictors were designed in an environment with SNR $= 10$ dB ($\sigma_v^2 = 0.1$). The first predictor with $M_1 = 7$ and $\mu_1 = 0.7$, and the second predictor with $M_2 = 10$ and $\mu_2 = 0.5$. Observe how a slight change in the design leads to dramatically different performances. To get around the sensitivity issue, we pursue a combination approach. The goal is to employ a mixture of multiple individual predictors spanning a reasonable range of the unknown parameters. The individual predictors outputs are efficiently combined by a supervisor such that the global system is able to perform as well as the best individual predictor [3], [4].

4.1. Convex Combiners

Specifically, we extend the ideas presented in [3]. Before presenting the error prediction filter to the root solver (refer to Fig. 3), we improve the quality of the predictor by dynamically combining a span of L normalized LMS predictors with different orders. The individual predictors are independent and are combined according to their individual performance (see Fig. 5):

$$\hat{x}(i) = \sum_{k=1}^{L} \lambda_k \hat{x}_k(i) \quad , \qquad \sum_{k=1}^{L} \lambda_k = 1 \tag{12}$$

which is equivalent to using the prediction vector

$$c_{i-1} = \sum_{k=1}^{L} \lambda_k c_{k,i-1} \tag{13}$$

The k^{th} predictor has order M_k and step-size μ_k:

$$\hat{x}_k(i) = x_{i-1} c_{k,i-1} \tag{14}$$

The regressor x_{i-1} has the order M_L of the largest predictor and is presented to all individual predictors c_k, which are filled out with zeros to match the vector dimensions when necessary.

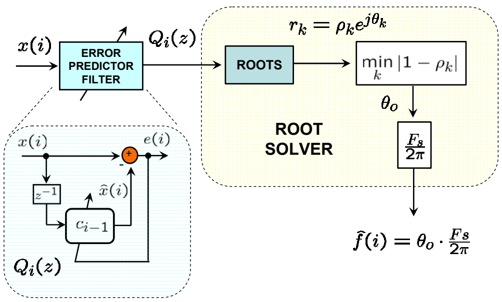

Fig. 3. An adaptive linear prediction implementation.

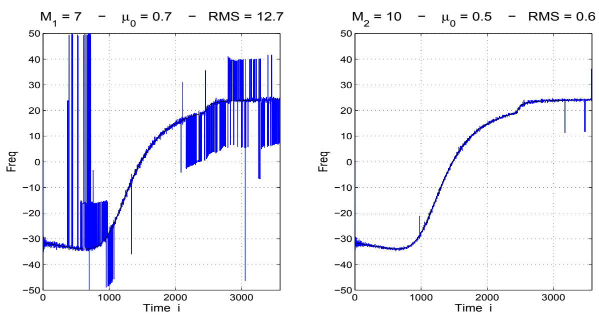

Fig. 4. Sensitivity of the ALP solution.

The key step is to design the combiners *efficiently*. To provide convexity to the predictors' aggregation we use combiners of the form:

$$\lambda_k = \frac{y_k}{\sum_{\ell=1}^{L} y_\ell} \quad , \qquad y_k = f(a_k) \tag{15}$$

where y_k is a generic *real* activation function of a *complex* argument a_k. It is a function at our choice and the *complex* coefficient a_k is what is truly adapted, say as [3]

$$a_k(i) = a_k(i-1) - \mu_a \left[\nabla_{a_k} |e(i)|^2\right]^*_{a_k = a_k(i-1)} \tag{16}$$

It can be shown that for a generic function $y_k = f(a_k)$ one gets

$$\begin{aligned} \nabla_{a_k} |e(i)|^2 &= -e^*(i) x_{i-1} \big(c_k - c\big) \frac{\partial y_k}{\partial a_k} \cdot \frac{1}{\sum_\ell y_\ell} \\ &= -e^*(i) \big(\hat{x}_k(i) - \hat{x}(i)\big) \frac{\partial y_k}{\partial a_k} \cdot \frac{1}{\sum_\ell y_\ell} \end{aligned} \tag{17}$$

Some possible choices for y_k that have been tested are

$$y_k = \left|e^{-\frac{a_k}{2}}\right|^2 \tag{18}$$

$$y_k = e^{-|a_k|^2} \tag{19}$$

$$y_k = |a_k|^2 \tag{20}$$

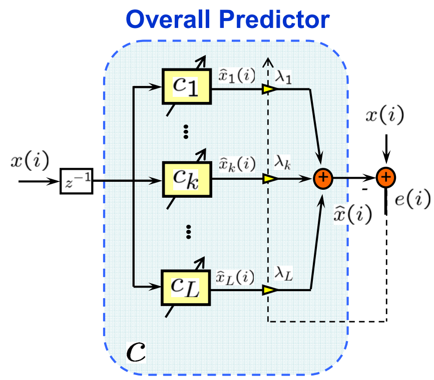

Fig. 5. A bank of combination filters.

leading to the following adaptive combiners:

$$\lambda_k(i) = \frac{\left|e^{-\frac{a_k}{2}}\right|^2}{\sum_{\ell=1}^{L} \left|e^{-\frac{a_\ell}{2}}\right|^2} \tag{21}$$

$$\lambda_k(i) = \frac{e^{-|a_k|^2}}{\sum_{\ell=1}^{L} e^{-|a_\ell|^2}} \tag{22}$$

$$\lambda_k(i) = \frac{|a_k|^2}{\sum_{\ell=1}^{L} |a_\ell|^2} \tag{23}$$

Using (16) and (18) we obtain the following adaptation rule for the first combiner (21):

$$a_k(i) = a_k(i-1) - \mu_a e(i) \big(\hat{x}_k(i) - \hat{x}(i)\big)^* \lambda_k(i) \tag{24}$$

where $\hat{x}(i)$ and $\hat{x}_k(i)$ are defined as in (12) and (14), respectively. Likewise, the two last combiners (22) and (23) lead to the adaption rule:

$$a_k(i) = a_k(i-1) - \mu_a a_k(i-1) e(i) \big(\hat{x}_k(i) - \hat{x}(i)\big)^* \lambda_k(i) \tag{25}$$

When complexity is an issue, we may limit the number of filters to $L = 2$, using, for instance, a low-order filter and a high order filter to capture signals with richer dynamics. With this assumption it is possible to derive a convex combination rule with only one adaptive coefficient a, by using $\lambda_1 = \lambda$, $\lambda_2 = 1 - \lambda$, and choosing λ as

$$\lambda(i) = \frac{1}{1 + \left|e^{-\frac{a(i-1)}{2}}\right|^2} \tag{26}$$

In this case, the adaptive rule for $a(i)$ becomes

$$a(i) = a(i-1) + \mu_a e(i) \big(\hat{x}_1(i) - \hat{x}_2(i)\big)^* \lambda(i) \big(1 - \lambda(i)\big) \tag{27}$$

5. ENHANCEMENT TECHNIQUES

Note that the noise effect already mentioned in Fig. 2 can be boosted by the stochastic gradient disturbances introduced by the predictors (10) as well as the adaptive combiners rule (16). This may cause spikes in the estimated frequency that are not correlated with the underlying true frequency ω, as Fig. 4 shows.

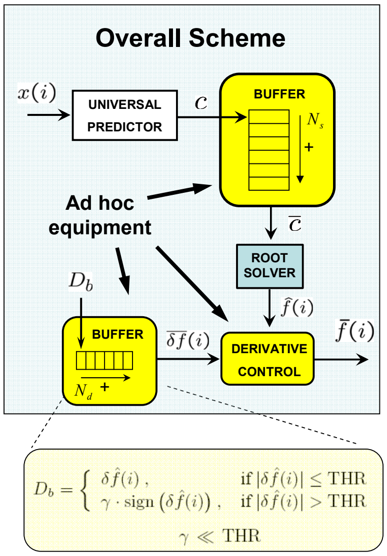

Fig. 6. The combination scheme with smoothing and derivative control.

In order to enhance the quality of the estimation we may use a smoothing procedure along with derivative control. The smoothing is basically an average performed over the last N_s global predictor vectors c in (13). The derivative control is a bit more complex, with a special buffer to smooth the spikes that may take place in noisy scenarios. The derivative buffer keeps track of the average $\overline{\delta f}(i)$ of the last N_d "*good*" derivative samples. Let

$$\delta\hat{f}(i) \triangleq \hat{f}(i) - \hat{f}(i-1) \tag{28}$$

and for a given threshold THR define a good derivative sample D_b as:

$$D_b = \begin{cases} \delta\hat{f}(i), & \text{if } |\delta\hat{f}(i)| \leq \text{THR} \\ \gamma \cdot \text{sign}\left(\delta\hat{f}(i)\right), & \text{if } |\delta\hat{f}(i)| > \text{THR} \end{cases} \tag{29}$$

where $\gamma \ll \text{THR}$. The derivative buffer is always fed with D_b. Whenever the derivative is bigger than THR, the sign information is kept but the magnitude is clamped, improving the smoothing process.

Finally, the enhanced estimated frequency $\bar{f}(i)$ is given by

$$\bar{f}(i) = \begin{cases} \hat{f}(i), & \text{if } |\delta\hat{f}(i)| \leq \text{THR} \\ \hat{f}(i-1) + \overline{\delta f}(i), & \text{if } |\delta\hat{f}(i)| > \text{THR} \end{cases} \tag{30}$$

6. SIMULATION RESULTS

Besides direct graphical comparison, we also use as figure of merit the total root mean-square error (RMS):

$$\text{RMS} \triangleq \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|f(i) - \bar{f}(i)\right|^2} \tag{31}$$

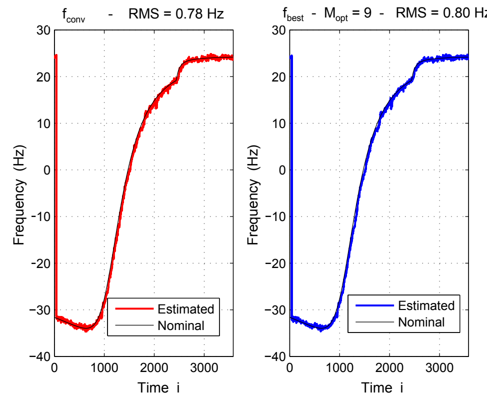

Fig. 7. Example 1: convex ALP (left) and ALP (right).

We present simulations corresponding to tough scenarios, going even further (i.e., worse) than the nominal SNR scenarios encountered by Spirit and Opportunity [1, 2]. Two curves are shown. The first curve compares our proposed convex ALP scheme (left plots) with the best individual ALP estimator (right plots), indicating the RMS error attained in each case. The second curve shows the time evolution of the convex combiners λ_k (left plots) and a comparison of the RMS error of the individual predictors as a function of filter order with our convex ALP scheme (fixed order M_L). In all simulations the convex combination attained universality with respect to the class of ϵ-NLMS predictors spanning orders up M_L, for a given step-size $\mu_k = \mu_0$. Note that in all cases we are confronting our convex scheme with the best individual predictor. The same sample frequency $F_s = 100$ Hz is used in all examples, and the Doppler profile is of the same nature as employed in [1].

In the first example we use the combiner (23) in a scenario with $\text{SNR} = 17\,\text{dB/Hz}$ ($\sigma_v^2 = 2$) and filter orders $\text{M} = [7:2:13]$. The ϵ-NLMS predictors are tuned to $\mu_0 = 0.1$ and the adaptive combining rule (25) is tuned with $\mu_a = 0.7$. Note in Fig. 7 that the convex ALP scheme attained better performance ($\text{RMS} = 0.78$ Hz) than the best individual ALP ($\text{RMS} = 0.80$ Hz).

In the second example we increase the noise level such that $\text{SNR} = 14$ dB/Hz ($\sigma_v^2 = 4$) with predictors spanning orders $\text{M} = [8:3:17]$ and tuned at $\mu_0 = 0.15$. The adaptive rule (24) was tuned with $\mu_a = 0.1$ using the activation function (21). In Fig. 8 we observe a dramatic improvement in performance when applying the proposed convex combination together with the simple enhancement techniques. Figure 9 attests the universality attained by the convex scheme.

In the third example we explore the two filters case, using (26) and (27) and decreasing the SNR even further, to $\text{SNR} = 12$ dB/Hz ($\sigma_v^2 = 6$). The two predictors have orders $M_1 = 7$ and $M_2 = 16$ with step-size $\mu_0 = 0.1$. The adaptation rule (27) employed $\mu_a = 0.1$. Figures 10 and 11 depict the results, implying that there is no need for complexity to improve considerably the performance when the combination of predictors is performed efficiently.

Finally, in the last example we compare our convex scheme with the original maximum likelihood scheme presented in [1]. As Fig.12 suggests, our scheme can be quite competitive, leading to similar performance albeit with much less computations than the maximum likelihood approach.

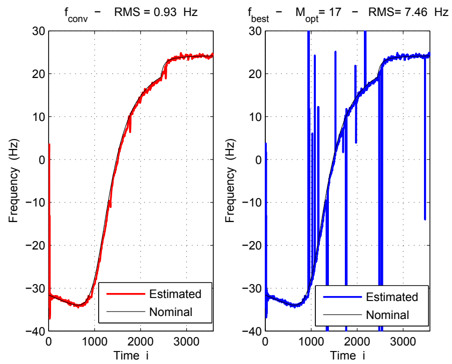

Fig. 8. Example 2: convex ALP (left) and ALP (right).

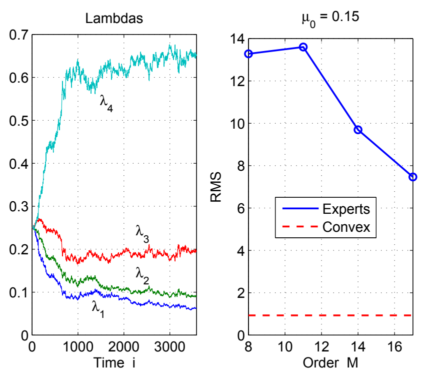

Fig. 9. Example 2: the combiners and the RMS error.

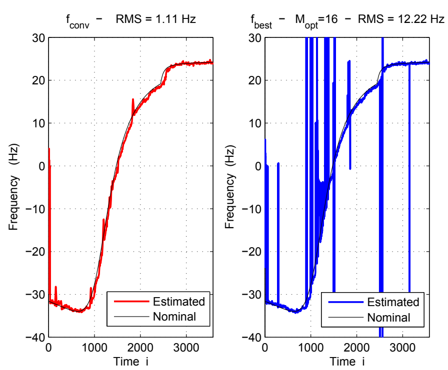

Fig. 10. Example 3: convex ALP (left) and ALP (right).

7. CONCLUSION

We proposed a simple yet robust scheme - a critical feature in deep space communications - for frequency estimation and tracking. It presents good performance for a wide range of SNR, although we focused in this work on low SNR, which is common in the Martian EDL scenario. The investigated convex combination of individual adaptive predictors is able to outperform the best individual predictor, achieving universality in the class of ϵ-NLMS predictors up to order M_L, operating with fixed and small step-sizes due to low SNR conditions.

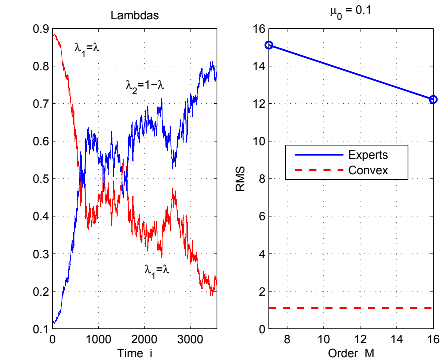

Fig. 11. Example 3: The combiners and the RMS error.

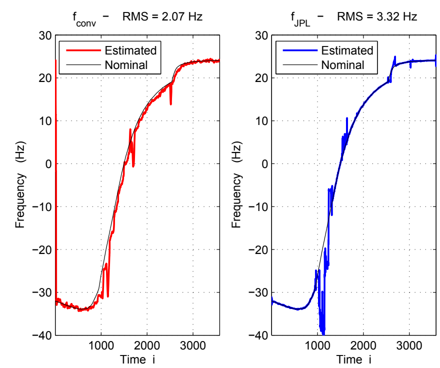

Fig. 12. Example 4: Performance comparison of the convex ALP (left plot) with the maximum likelihood method proposed in [1](right plot).

Acknowledgment

The authors wish to thank Mr. N. Khajehnouri for useful discussions on the topic of the paper.

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.

[3] 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.

[4] N. Merhav and M. Feder. “Universal prediction,” *IEEE Trans. on Inf. Theory*, vol. 44, pp. 2124-2147, 1998.

[5] A. H. Sayed. *Fundamentals of Adaptive Filtering*, Wiley, NJ, 2003.