

## COMBINATION OF ADAPTIVE FILTERS FOR RELATIVE NAVIGATION

*Luiz F. O. Chamon and Cássio G. Lopes*

Electrical Engineering Dept., University of São Paulo, São Paulo, Brazil  
email: chamon@usp.br ; cassio@lps.usp.br

### ABSTRACT

Relative navigation of spacecrafts may be accomplished via time delay estimates. In this work an adaptive filtering approach is employed, which involves an estimation and a detection step. By formally posing the detection problem, a more meaningful detector, that embeds a reliability measure into the delay estimates, is proposed. The estimation step is enhanced via convex combination schemes, that address the Poisson distributed signals, sparse channel and low signal-to-noise ratio. To evaluate time delay estimation techniques, different criteria based on probability of detection are studied, leading to a new figure of merit. The resulting solution outperforms the existing adaptive filters techniques under the new criterion, as shown by simulations.

### 1. INTRODUCTION

The use of signals delay for navigation is a well established method, used from the obsolete LORAN (LONG RANGE Navigation) to GPS (Global Positioning System) [1]. In space, however, it is not always easy to generate the beacons needed for such methods; yet formation flying is growing a key technology in these missions, e.g., the NASA EOS (Earth Observing System) program, of which Brazil participates [2]. Though beyond GPS range, deep space probes need the same accurate navigation information, relying mostly on ground stations, such as NASA's Deep Space Network (DSN) [3]. Many applications (e.g., *interferometric imaging* [4]), however, only require relative positioning, where the use of celestial X-ray sources as bearing signals have often been proposed.

Regardless of the source's nature [3,5], the navigation problem can be reduced to that of time delay estimation (TDE). Referring to Figure 1,  $\Delta\vec{x}$  is the relative position vector and  $\vec{n}$  is the normal vector, assumed to be identical for both ships since they are closer to each other than to the source. The delay ( $t_d$ ) between their received signals will depend on  $\Delta d = \vec{n} \cdot \Delta\vec{x} = ct_d$ , the relative distance in direction  $\vec{n}$ , where  $c$  is the speed of light [5]. Using more sources allows three dimensional positions to be calculated.

This work studies the TDE problem employing an adaptive filtering (AF) solution [6] and derives a reliability measure based on detection theory, which leads to a new detector and a more significant figure of merit to evaluate

TDE systems. A convex combination of AFs is proposed to improve delay estimates under the new criterion over a wide signal-to-noise ratio (SNR) range.

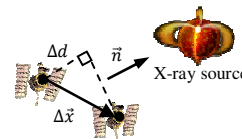


Figure 1. Relative position of two ships observing an X-ray source

### 2. THE TIME DELAY ESTIMATION PROBLEM

The signals are modeled observing that, due to the nature of X-ray sensors, both source signal and measurement noise at the spacecrafts can be represented as the realizations of Poisson processes. Their distributions are then given by  $f(k, \theta) = e^{-\theta} \frac{\theta^k}{k!}$ , where  $\theta$  is the expected number of arrivals in a time bin and  $f$  describes the probability of exactly  $k$  photons arriving in that bin. Also, since both vehicles are very far from the source, relativistic effects are considered to affect equally their detectors, so that the discrete measured signals can be summarized as

$$x_1(i) = s(i) + v_1(i), \quad (1)$$

$$x_2(i) = \alpha s(i - n_d) + v_2(i), \quad (2)$$

where  $s(i)$  is the Poisson distributed measurement of the celestial source's X-ray signal,  $0 < \alpha < 1$  is an attenuation factor,  $v_1(i)$  and  $v_2(i)$  are independent noise signals,  $n_d = \lfloor t_d/t_s \rfloor$ , and  $t_s$  is the sampling period [3,5]. To extract the delay from the above signals, the ships must share observational data, which is not an issue given that communication also plays central role in attitude control [4].

A variety of solutions exists for the TDE problem, though the most commonly used is probably the generalized cross-correlation (GCC). This set of methods is an improvement on classical correlation where information on signal and/or noise is used to enhance its performance. One of the so called processors, the Hannan-Thomson (HT) processor, is the maximum likelihood (ML) time delay estimator [7].

The approach undertaken in this work, however, sees (1) and (2) in terms of a discrete channel model (Figure 2). This pseudo-channel can be approximated by an FIR filter,

whose coefficients are captured by an  $M \times 1$  vector  $w^o$ ,  $M > n_d$ . AFs are used to identify the channel, whose optimal estimate is of the form  $w^o = [0 \cdots 0 \alpha 0 \cdots 0]^T$ , with  $\alpha$  at  $w^o[n_d]$  [6]. This method takes on two distinct steps (Figure 3): (i) the estimation of the pseudo-channel and (ii) the detection of the delay in this estimate.

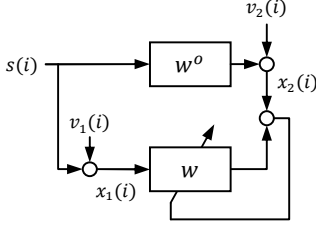


Figure 2. Channel model of the TDE problem

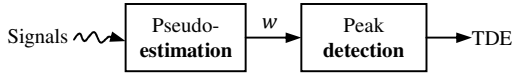


Figure 3. TDE through parameter estimation scheme

This pseudo-estimation approach to TDE may be asymptotically efficient [6,8] like ML, though the Cramer-Rao lower bound is known to be inadequate for low SNRs [9]; comparative studies have also revealed that under uncertain *a priori* knowledge and large noise LMS (Least Mean Squares) may perform as well as ML approximations [10] and other LMS-based algorithms [11], being less sensitive to changes in signals spectra than GCC [6]. In addition, previous work [5] shows that GCC may be outperformed the AF approach in low SNR. AFs robustness and simplicity further motivate their use in aerospace applications which impose hard constraints on power consumption and thus computational complexity. Finally, AFs are model-free, adapting even to non-stationary characteristics of systems [12]. These are the main reasons behind the choice of this technique over GCC [7], ML approximations [10] or sparse Bayesian methods [13].

### 3. THE DETECTION PROBLEM

The original detection step is usually restricted to finding a peak in the pseudo-estimate [5-7,10,11], namely

$$\hat{t}_d = \operatorname{argmax}(w) \cdot t_s, \quad (3)$$

failing to address the reliability of the detection, a measure of foremost importance in iterative estimations (AFs) [6], which undergo a transient stage, and low SNRs, where anomalous effects become more significant [14].

The reliability of (3) is intimately related to the probability of detecting  $w^o$  in  $w$ , which should manifest itself as a prominent peak over the corruption due to input noises. This problem, posed under a detection formalism, is equivalent to a hypothesis test that determines whether  $w$

has a portion due to a delay in the input signals ( $H^1$ ) or if it is statistically equivalent to noise ( $H^0$ ). Explicitly

$$\begin{aligned} H^1: w &= w^o + \tilde{w}, \\ H^0: w &= \tilde{w} \end{aligned}, \quad (4)$$

where  $\tilde{w}$  represents the corruption in the pseudo-estimate due to input noise. It is shown in [12] that for long AFs,  $w$  behaves as if normally distributed even if the input signals are not. Thus, the weights noise vector  $\tilde{w}$  is modeled as a Gaussian process  $\mathcal{N}(\Psi, R_{\tilde{w}})$ , where  $\Psi = \psi q$  is the mean vector, with  $\psi$  a scalar and  $q$  an  $M \times 1$  vector of ones, and  $R_{\tilde{w}} = E\tilde{w}\tilde{w}^*$  is the covariance matrix. The probability of detecting  $w^o$  then becomes the conditional probability of choosing  $H^1$  given  $H^1$  and describes how reliable (3) is at finding the correct  $t_d$ .

The optimum test for the decision (4) is [15]

$$\Lambda = \frac{f(\hat{w}|H^1)}{f(\hat{w}|H^0)} \geq \eta, \quad (5)$$

choosing  $H^1$  when  $\Lambda \geq \eta$  and  $H^0$  otherwise. Applying the hypotheses (4) leads to

$$\Lambda = \frac{\exp\left[-\frac{1}{2}(\hat{w}-w^o-\Psi)^T R_{\tilde{w}}^{-1}(\hat{w}-w^o-\Psi)\right]}{\exp\left[-\frac{1}{2}(\hat{w}-\Psi)^T R_{\tilde{w}}^{-1}(\hat{w}-\Psi)\right]}.$$

Since the natural logarithm is a monotonic function and that both sides of (5) are positive, the test is equivalent to

$$w^{oT} R_{\tilde{w}}^{-1}(\hat{w} - \Psi) \geq \ln \eta + \frac{1}{2} w^{oT} R_{\tilde{w}}^{-1} w^o \triangleq v. \quad (6)$$

#### 3.1. Detection and false alarm

Given that (3) can always deliver a peak detection,  $H^1$  is initially assumed, then the detection reliability is evaluated. Thus, only two outcomes of the experiment in (4) are of interest: (i) choosing  $H^1$  when  $H^0$  is true (false alarm) and (ii) choosing  $H^1$  when  $H^1$  true (detection). Since (6) is an affine transformation of  $w$ , the conditional probabilities (i) and (ii) are Gaussian [16], yielding

$$\begin{aligned} P_F &= P[H^1|H^0] = P[w^{oT} R_{\tilde{w}}^{-1}(\tilde{w} - \Psi) > v] = \\ &= Q\left(\frac{v}{(w^{oT} R_{\tilde{w}}^{-1} w^o)^{1/2}}\right), \end{aligned} \quad (7)$$

$$\begin{aligned} P_D &= P[H^1|H^1] = P[w^{oT} R_{\tilde{w}}^{-1}(w^o - \Psi) > v] = \\ &= Q\left(\frac{v - w^{oT} R_{\tilde{w}}^{-1}(w^o - \Psi)}{(w^{oT} R_{\tilde{w}}^{-1} w^o)^{1/2}}\right), \end{aligned} \quad (8)$$

where  $P_F$  and  $P_D$  are the probability of false alarm and detection, respectively, and  $Q$  is the complement of the standard normal cumulative distribution [16].

Ideally,  $P_F$  should be as small as possible while  $P_D$  is as large as possible. However, writing  $P_D$  as a function of  $P_F$ , namely

$$P_D = Q\left(Q^{-1}(P_F) - \frac{w^{oT} R_{\tilde{w}}^{-1}(w^o - \Psi)}{(w^{oT} R_{\tilde{w}}^{-1} w^o)^{1/2}}\right), \quad (9)$$

it is clear that those are conflicting objectives, since  $P_D$  is composed of a decreasing function of  $P_F$ . Even so, for any given  $P_F$ ,  $P_D$  can be enhanced increasing

$$\frac{w^{oT} R_{\tilde{w}}^{-1} (w^o - \Psi)}{(w^{oT} R_{\tilde{w}}^{-1} w^o)^{1/2}} = \frac{\alpha - \psi}{r}, \quad (10)$$

where  $r = [R_{\tilde{w}}]_{n_d, n_d}^{1/2}$  and  $\alpha$  can now be interpreted as the magnitude of the delay peak. Equation (10) induces a measure of reliability for the peak detection defined in (3), given that  $P_D$  depends directly on it.

### 3.2. The new time delay detector

Using (10), a tractable measure of reliability can be built in the definition of the detector (3) by finding estimates for  $\alpha$ ,  $\psi$  and  $r$ . The natural choice would be to use their ML estimates  $\hat{\alpha}$ ,  $\hat{\psi}$  and  $\hat{r}$ . However, ML estimates are known to present large standard errors under low SNRs unless  $w$  is very long [16]. As a result, an alternative measure of reliability is proposed in terms of available data.

First, define  $w'$  as the  $(M-1) \times 1$  vector containing all the samples in  $w$  but the correct delay tap  $w[n_d]$  (Figure 4). Under the  $H^1$  assumption taken earlier,  $w'$  is composed uniquely of noise (from  $\tilde{w}$ ) and  $n_d = \text{argmax}(w)$ . Now, roughly assuming  $\alpha \cong \max(w)$  and replacing the ML estimates by upper bounds, i.e.  $\hat{\psi} \leq \max(w')$  and  $\hat{r} \leq \max(w')$ , (10) becomes a practical and rather intuitive measure of reliability for (3), namely

$$\frac{\max(w) - \max(w')}{\max(w')}. \quad (11)$$

Note that (11) captures the tendency of  $P_D$  (see (9) and (10)). Besides, simulations have shown that, for low SNRs and considering the filter lengths employed in this work, it leads to better estimates for (10).

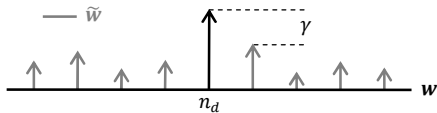


Figure 4. Illustration of definitions based on  $w$

The new detector is finally endowed with a confidence level  $\gamma$ , which expresses how well defined a peak must be before declaring the outcome of (3) a reliable delay estimate (Figure 4). Otherwise, the delay is declared *undefined*. Gathering (3), (11) and  $\gamma$ , the new detector is given by:

$$\hat{t}_d = \begin{cases} \text{argmax}(w) \cdot t_s, & \frac{\max(w) - \max(w')}{\max(w')} > \gamma, \\ \text{undefined}, & \text{otherwise} \end{cases} \quad (12)$$

Note that for  $\gamma = 0$  the *classical* detector (3) is recovered.

## 4. THE ESTIMATION PROBLEM

For (12) to be used, an estimate of the pseudo-channel must be provided, which may be obtained by minimizing the Mean Square Error (MSE) [12], defined as

$$\text{MSE} = E \|d(i) - u_i w\|_2^2, \quad (13)$$

where  $u_i \triangleq [x_1(i) \cdots x_1(i-M+1)]$  is a row regressor vector and  $d(i) \triangleq x_2(i)$  is the desired signal. In the system identification setup,  $d(i) = u_i w^o + n(i)$ , where  $n(i)$  models the plant noise and depends on  $v_1$  and  $v_2$  (Figure 2).

It is straightforward to show that minimizing (13) also minimizes  $\text{Tr}(R_{\tilde{w}})$  [12], and, therefore, from (9) and (10), increases  $P_D$ . AFs attempt to minimizing (13), hence their use is justified to increase  $P_D$ . Moreover, the optimal solution to (13) is equivalent to GCC ROTH processor [6].

### 4.1. Convex combination

The studied application presents rather unusual characteristics: Poisson-distributed signals, low SNRs and sparse channel, which makes difficult the design of one single AF able to address such scenario. In the literature there are techniques that automatically combine AFs via a supervisor, so that the global filter is able to outperform any of the individual AFs [17-20]. In convex combination schemes, the overall filter is obtained as

$$w_{i-1} = \lambda(i) w_{1,i-1} + (1 - \lambda(i)) w_{2,i-1}, \quad (16)$$

where  $\lambda \in [0, 1]$  to guarantee convexity and the AFs  $w_1$  and  $w_2$  evolve independently. A typical choice for  $\lambda$  is  $\lambda(i) = \frac{1}{1 + e^{-a(i-1)}}$ , where  $a$  is a support variable adapted to minimize the global estimation error  $e(i) = d(i) - u_i w_{i-1}$  following the steepest descent rule [20]

$$a(i) = a(i-1) - \mu_a [\nabla_a |e(i)|^2]_{a=a(i-1)}^*. \quad (17)$$

The resulting recursion can be shown to be

$$a(i) = a(i-1) + \mu_a e(i) u_i^* (w_{1,i-1} - w_{2,i-1}) \lambda(i) (1 - \lambda(i))$$

Candidate filters for  $w_1$  and  $w_2$  arise from the nature of the application, which support the use of the Least Mean Fourth (LMF) algorithm to address the low SNR and non-Gaussian signals features [21], and the Improved Proportionate Normalized LMS (IPNLMS) algorithm, which explicitly accounts for the sparse pseudo-channel [22]. Their recursions are shown below:

$$w_{1,i} = w_{1,i-1} + \mu u_i^* e_1^3(i) \quad (\text{LMF}) \quad (14)$$

$$w_{2,i} = w_{2,i-1} + \frac{\mu G}{\epsilon + u_i G u_i^*} u_i^* e_2(i) \quad (\text{IPNLMS}) \quad (15)$$

where for  $k = 1, 2$ ,  $e_k(i) = d(i) - u_i w_{k,i-1}$ ,  $w_{k,i}$  is an estimate for  $w$  at iteration  $i$ ,  $\mu > 0$  is the step-size,  $0 < \epsilon \ll 1$  is a regularization factor,  $G = \text{diag} \left\{ \frac{1-\beta}{2L} q + (1+\beta) \frac{|w_{2,i-1}|}{2 \|w_{2,i-1}\|_1} \right\}$ ,  $\beta$  is a constant (usually  $\beta = -0.5$ ) and  $\text{diag}\{\cdot\}$  is the diagonal operator.

## 5. A NEW FIGURE OF MERIT: DISCRIMINATION

Although the previous section showed that minimizing the MSE improves the detection capability, it is known that,

along with other common AFs performance assessments, it may be a misleading metric when it comes to TDE [6]. Since the delay is ultimately found by the detection stage, it is more suitable to use  $P_D$  as a figure of merit to evaluate TDE solutions. However, calculating  $P_D$  analytically is difficult, motivating different, sometimes fallacious, estimates of  $P_D$  via Monte Carlo simulations.

### 5.1. Existing detection criteria

The use of the mean weight vector as a criterion for detection has sometimes been suggested [5,6], with corresponding probability of detection given by

$$\hat{P}_D = \Pr [\operatorname{argmax}(Ew) = n_d]. \quad (18)$$

Although accurate at higher SNRs, it becomes misleading as the noise level grows, due to an increase in the probability of anomalies, a phenomenon intrinsic to TDE problems [14]. Figure 5 compares  $Ew$  to two  $w$  realizations after 1000 iterations of an LMS filter in conditions similar to [5]. Clearly, though the ensemble average presents a peak at the true delay, the individual realizations do not, suggesting that (18) is not reliable in a practical situation, where only one realization is available.

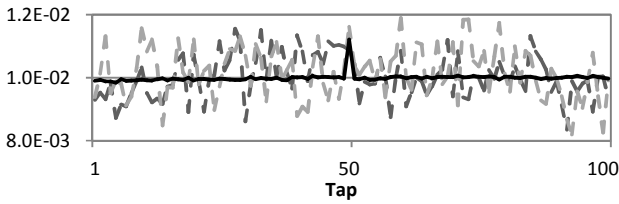


Figure 5. Comparison between the ensemble-averaged weight (solid line) and individual realizations (dashed lines)

Under the new proposed detector (12), the probability of detection is

$$P_D = \Pr \left( \frac{\max(w) - \max(w')}{\max(w')} > \gamma \mid \operatorname{argmax}(w) = n_d \right), \quad (19)$$

which represents the probability of detecting the peak at  $n_d$  with a margin of  $\gamma$ . The ML  $P_D$  estimate ( $\hat{P}_D^{ML}$ ) can be easily determined evaluating the ratio of correct detection to number of trials [16]. For  $\gamma = 0$ , the *accuracy percentage* (AP) [23] figure of merit is recovered.

### 5.2. A more convenient figure of merit

From (19), when  $\operatorname{argmax}(w) = n_d \Leftrightarrow \max(w) = \hat{w}[n_d]$ . Therefore, the *discrimination*  $\mathcal{D}$  can be defined from (11) as

$$\mathcal{D} = \frac{\hat{w}[n_d]}{\max(w')}, \quad (20)$$

so that (19) becomes  $P_D = \Pr[\mathcal{D} > \gamma + 1]$ . This quantity ( $\mathcal{D}$ ) allows, with a single Monte Carlo simulation, to determine which  $\gamma$  is required to meet a predefined  $P_D$ , which is very convenient for design purposes. To do so,

recall that for long enough filters  $w$  is approximately Gaussian [12]; then, from (20), so is  $\mathcal{D}$  (Figure 6). Thus,  $P_D$  can be estimated using

$$\hat{P}_D^{\mathcal{D}} = Q \left( \frac{\gamma + 1 - \mu_{\mathcal{D}}}{\sigma_{\mathcal{D}}} \right), \quad (21)$$

where  $\mu_{\mathcal{D}}$  and  $\sigma_{\mathcal{D}}$  are the mean and standard deviation of  $\mathcal{D}$ . Different  $\gamma$  can be tested via (21) using a single ensemble average evaluation.

Figure 7 compares  $\hat{P}_D$ ,  $\hat{P}_D^{ML}$  and  $\hat{P}_D^{\mathcal{D}}$  for SNRs ranging from -15 dB to 0 dB, with  $\gamma = 0$  and the same parameters as in Figure 5. As expected, (18) overestimates the ability of detecting the correct delay at low SNRs, while  $\hat{P}_D^{\mathcal{D}}$  follows closely  $\hat{P}_D^{ML}$ .

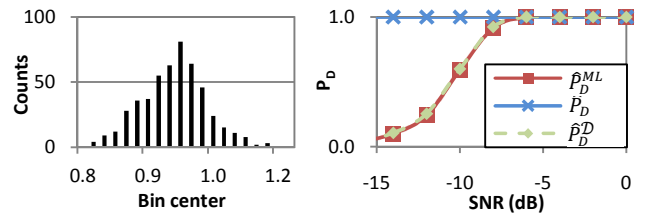


Figure 6. Histogram of  $\mathcal{D}$  (SNR = -12 dB)

Figure 7. Probability of detection estimates

## 6. SIMULATIONS

Using (1) and (2) as the model for the signals acquired by the spacecrafts, the AFs applied by [5] and the specific combination proposed in Section 4.1 are compared using (21) for different SNRs ( $-15 \leq \text{SNR} \leq 10$  in dB) in the scenario of relative navigation. The experiment is carried out in a stationary scenario with  $\alpha = 0.9$ ,  $M = 100$ ,  $n_d = 50$ , the source signal variance  $\theta_s = 1$ , and the measurement noise variances  $\theta_{v_1} = \theta_{v_2} = \theta_n$ . The AFs tuning lead to  $\mu_{LMS} = 1.2 \cdot 10^{-4}/M$ ,  $\mu_{NLMS} = 10^{-2}$ ,  $\mu_{LMF} = 9 \cdot 10^{-6}/M$ ,  $\mu_{IPNLMS} = 10^{-2}$ ,  $\beta = -0.5$ ,  $\epsilon = 10^{-6}$  and  $\mu_a = 0.5$ .

Figures 8 and 9 show the average  $\mathcal{D}$  after 5000 iterations, highlighting the region of interest, since for  $\text{SNR} > 0$  dB the tendencies do not change. The plots show the superiority of NLMS and IPNLMS for higher SNRs, the latter being slightly better, whilst LMS and LMF excel for SNRs below -8 dB. They also reveal the convex combination is an improvement over the individual AFs.

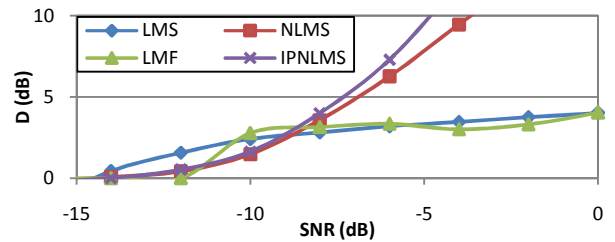


Figure 8.  $\mathcal{D} \times \text{SNR}$  for individual adaptive filters

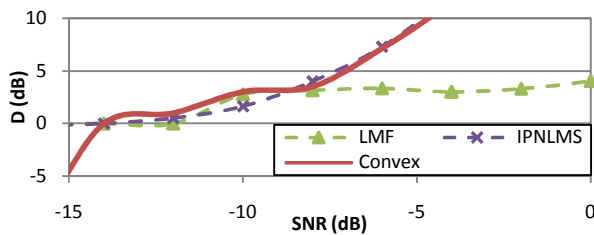


Figure 9.  $\mathcal{D}$  x SNR for convex LMF/IPNLMS

Although practical, Figures 8 and 9 can be misleading since they do not account for the variance of  $\mathcal{D}$ . To address that,  $P_D$  is calculated based on (21) adopting a threshold commonly used in the radar and wireless milieu [24]:  $\mathcal{D} = 3$  dB ( $\gamma = 0,41$ ). Figure 10 now clearly reveals the LMF superiority in terms of  $P_D$ , justifying the convex combination, which highly improved the detection at lower SNRs, albeit at the cost of performing slightly worse than the IPNLMS over a small region (Figure 11). Overall, it evidently enhanced the TDE over the SNR range.

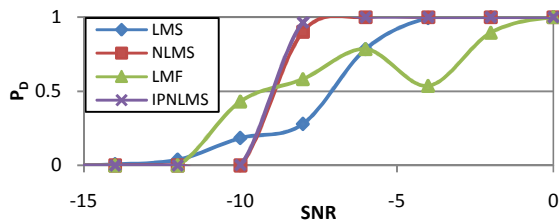


Figure 10. Probability of detection for  $\mathcal{D} = 3$  dB

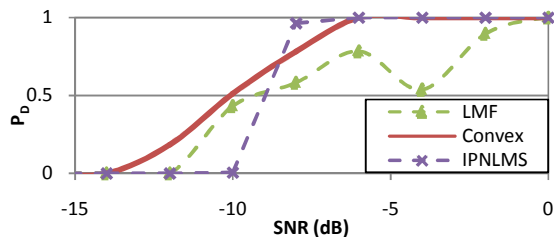


Figure 11.  $P_D$  x SNR for  $\mathcal{D} = 3$  dB for convex LMF/IPNLMS

## 7. CONCLUSION

TDE problems based on AFs involve an estimation stage followed by a detection stage. Both steps were addressed in this work by proposing a new detector that embeds a practical reliability measure and a convex combination scheme that improves the probability of detection. The latter was argued as a more meaningful metric to evaluate TDE solutions. Future works will include the use of hierarchical combinations [25] to employ other PNLMS-based AFs [26] and the study of non-stationary delays.

## REFERENCES

[1] U.S. Coast Guard Navigation Center [Online]. Available at <http://www.navcen.uscg.gov/>  
 [2] National Institute for Space Research website [Online]. Available: <http://satelite.cptec.inpe.br/informacao/orbitas.jsp>

[3] S. I. Sheikh et al., "Relative Navigation of Spacecraft Utilizing Bright, Aperiodic Celestial Sources," in *ION 63rd Ann. Meet.*, Cambridge, MA, 2007.  
 [4] R. S. Smith and F. Y. Hadaegh, "Distributed estimation, communication and control for deep space formations," *IET Con. Th. & App.*, vol. 1[2], Mar. 2007.  
 [5] A. A. Emadzadeh et al., "Online Time Delay Estimation of Pulsar Signals for Relative Navigation using Adaptive Filters," in *ION Pos., Loc. and Nav. Symp.*, Monterey, CA, 2008.  
 [6] F. A. Reed et al., "Time Delay Estimation using the LMS Adaptive Filter – Static Behavior," *IEEE Trans. ASSP*, v.29[3], 1981.  
 [7] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. ASSP*, v.24[4], 1976.  
 [8] L. Z. Qu and N. J. Bershard, "Comments on Time Delay Estimation using the LMS Adaptive Filter – Static Behavior," *IEEE Trans. ASSP*, v.33[6], 1985.  
 [9] D. Chazan et al., "Improved Lower Bounds on Signal Parameter Estimation," *IEEE Trans. Inf. Theory*, v.21[1], 1975.  
 [10] J. Krolik et al., "Time Delay Estimation of Signals with Uncertain Spectra," *IEEE Trans. ASSP*, v.36[12], 1988.  
 [11] H. C. So and P. C. Ching, "Comparative Study of five LMS-based adaptive time delay estimators," *IEE Proc. Radar, Sonar and Nav.*, v.148[1], 2001.  
 [12] A. H. Sayed, *Adaptive Filters*. Hoboken: Wiley-IEEE, 2008.  
 [13] M. Seeger, "Bayesian inference and optimal design for the sparse linear model," *J. Mach. Learn. Res.*, v.9, 2008.  
 [14] J. P. Ianniello, "Time Delay Estimation via Cross-Correlation in the Presence of Large Estimation Errors," *IEEE Trans. ASSP*, v.30[6], 1982.  
 [15] H. L. van Trees, *Detection, Estimation, and Modulation Theory – Part I*. Hoboken: Wiley-Interscience, 2001.  
 [16] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge: Cambridge University, 2006.  
 [17] C. G. Lopes et al., "Adaptive carrier tracking for Mars to Earth communications during entry, descent, and landing," *IEEE Trans. Aerosp. Electron. Syst.*, v.46[4], 2010.  
 [18] Y. Zhang and J. A. Chambers, "Convex Combination of Adaptive Filters for a Variable Tap-Length LMS Algorithm," *IEEE Signal Process. Letters*, v.13[10], 2006.  
 [19] M. T. M. Silva and V. H. Nascimento, "Improving the tracking capability of adaptive filters via convex combination," *IEEE Trans. Signal Process.*, v.56[7], 2008.  
 [20] J. Arenas-Garcia et al., "Mean-Square Performance of a Convex Combination of Two Adaptive Filters," *IEEE Trans. Signal Process.*, v.54[3], 2006.  
 [21] E. Walach and B. Widrow, "The Least Mean Fourth Adaptive Algorithm and its Family," *IEEE Trans. Inf. Theory*, v.30[2], 1984.  
 [22] J. Benesty and S. L. Gay, "An Improved PNLMS algorithm," in *Proc. ICASSP*, Orlando, FL, 2002.  
 [23] T. P. Bhardwaj and R. Nath, "Maximum Likelihood Estimation of Time-Delay in Multipath Acoustic Channel," *Signal Process.*, v.90, 2010.  
 [24] K. Chang, *RF and Microwave Wireless Systems*. Hoboken: Wiley-Interscience, 2000.  
 [25] C. G. Lopes et al., "Efficient Adaptive Carrier Tracking for Mars to Earth Communications During Entry, Descent and Landing," in *Proc. SSP*, Madison, WI, 2007.  
 [26] J. Arenas-Garcia and A. R. Figueiras-Vidal, "Adaptive combination of IPNLMS filters for robust sparse echo cancellation," in *Proc. MLSP*, Cancún, Mexico, 2008.