ADAPTIVE FILTERING APPROACHES FOR NON-GAUSSIAN STABLE PROCESSES

Orhan Arıkan[1], Murat Belge[1], A. Enis Çetin[2] and Engin Erzin[1]

[1]Bilkent University, Ankara, TURKEY.
[2]Koç University, İstanbul, TURKEY.

ABSTRACT

A large class of physical phenomenon observed in practice exhibit non-Gaussian behavior. In this paper, α-stable distributions, which have heavier tails than Gaussian distribution, are considered to model non-Gaussian signals. Adaptive signal processing in the presence of such kind of noise is a requirement of many practical problems. Since, direct application of commonly used adaptation techniques fail in these applications, new approaches for adaptive filtering for α-stable random processes are introduced.

1. INTRODUCTION

In many signal processing applications the noise is modeled as a Gaussian process. This assumption has been broadly accepted because of the Central Limit Theorem. However, a large class of physical observations exhibit non-Gaussian behavior, such as low frequency atmospheric noise, many types of man-made noise and underwater acoustic noise [1]-[3]. There exists an important class of distributions known as α-stable distributions [5] which can be used to model this type of noises. These distributions have heavier tails than those of Gaussian distribution, and they exhibit sharp spikes or occasional bursts in their realizations. A random variable is called α-stable if its characteristic function has the following form:

$$\phi(t) = exp\{i\alpha t - \gamma|t|^{\alpha}[1 + i\beta sign(t)\omega(t, \alpha)]\} \quad (1)$$

where $-\infty < a < \infty$, $\gamma > 0$, $0 < \alpha \le 2$, $-1 \le \beta \le 1$ and

$$\omega(t, \alpha) = \begin{cases} \tan(\alpha\pi/2) & \text{for } \alpha \ne 1 \\ \frac{2}{\pi}\log|t| & \text{for } \alpha = 1. \end{cases} \quad (2)$$

There is no compact expression for the probability density function of these random variables except $\alpha = 1$ and 2 cases which correspond to the Cauchy and Gaussian distributions, respectively.

Members of stable distributions also satisfy a generalized central limit theorem which states that if the sum of i.i.d. random variables converges then the limit distribution is a stable one. If individual distributions are of finite variance then the limit distribution is Gaussian. Tails of this type of distributions are characterized with the α parameter $(0 < \alpha \le 2)$ which is called as the characteristic exponent (α values close to 0 indicates impulsive nature and α values close to 2 indicates a more Gaussian type of behavior).

With the Gaussian assumption, signals could be treated in a Hilbert space framework which would allow the use of L_2 (or ℓ_2) norm in various optimization criteria. Whereas, the linear vector space generated by α-stable distributions is a Banach space when $(1 \le \alpha < 2)$. In the linear space of stable processes only p-norms exists for $p \le \alpha$, hence, ℓ_2 norm cannot be used with an α-stable processes. Modeling α-stable processes under a Gaussian assumption leads to unacceptable results as is reported in [5].

In this paper, various approaches to adaptive filtering is investigated under additive α-stable noise with finite mean corresponding to case of $1 \le \alpha < 2$. These approaches are also compared to recently introduced p-norm algorithms [4, 5]. The p-norm algorithms are presented in Section 2 and the use of pre-nonlinearity in adaptive filtering is investigated in Section 3. The simulation results are given in Section 4.

2. ADAPTIVE FILTERING FOR α-STABLE PROCESSES

The objective for a general filtering application is to find an FIR filter of length N, $\underline{w}$, that relates the input, $x(n)$ to the desired signal $d(n)$:

$$\hat{d}(k) = \underline{x}(k)' \underline{w} \quad (3)$$

where $\hat{d}(k)$ is the estimate of the desired signal at time instant k, and

$$\underline{x}(k) = [x(k) \ x(k-1) \ \cdots \ x(k-N+1)]' . \quad (4)$$

Commonly used adaptive filtering algorithms utilize the Hilbert space framework. This allows the use of least squares cost function whose solution can be found either exactly as in Recursive Least Squares (RLS) algorithms or approximated by Least-Mean-Squares (LMS) type methods [7, 8]. However, in the existence of α-stable processes least squares cost function cannot be defined because the variance of the error is not finite. Hence a new cost function other than least squares should be used.

In this work, we consider an adaptation algorithm for an FIR filter of length N. The problem is to adaptively update the tap weights of the FIR filter, $\underline{w}$, such that given an input sequence $x(n)$, the output of the filter is close to the desired response $d(n)$, both of which is assumed to be α-stable. In this case, it is appropriate to minimize the dispersion of the error function [5].

This adaptation problem can be solved asymptotically by using the stochastic gradient method with the motivation of the LMS algorithm [8]. Such an algorithm, least mean p-norm (LMP) algorithm, is proposed in [5]. This algorithm is a generalization of instantaneous gradient descent algorithm to α-stable processes, where the gradient of the p-norm of the error,

$$
\begin{aligned}
J &= E[|e(k)|^p] \\
&= E[|d(k) - \underline{w}(k)' \underline{x}(k)|^p], \quad 0 < p < \alpha \quad (5)
\end{aligned}
$$

is used, and the tap weights, $\underline{w}$, are adapted at time step $k+1$ as follows:

$$
\underline{w}(k+1) = \underline{w}(k) + \mu \, |e(k)|^{p-1} \, sgn(e(k)) \, \underline{x}(k) \quad (6)
$$

where μ is the step size which should be appropriately determined. Note that, for $p = \alpha = 2$ the LMP algorithm reduces to the well-known LMS algorithm [8]. When p is chosen as 1, the LMP algorithm is called the Least Mean Absolute Deviation (LMAD) algorithm [5]:

$$
\underline{w}(k+1) = \underline{w}(k) + \mu \, sgn(e(k)) \, \underline{x}(k) \quad (7)
$$

which is also known as the signed-LMS algorithm.

In this paper we introduce two normalized adaptation algorithms with the motivation of the Normalized-LMS algorithm. The first one, Normalized Least Mean p-Norm (NLMP) algorithm, uses the following update:

$$
\underline{w}(k+1) = \underline{w}(k) + \beta \, \frac{|e(k)|^{p-1} \, sgn(e(k))}{\|\underline{x}(k)\|_p^p + \lambda} \, \underline{x}(k) \quad (8)
$$

where $\beta, \lambda > 0$ are appropriately chosen update parameters. In (8) normalization is obtained by dividing the update term by the p-norm of the input vector, $\underline{x}(k)$. The regularization parameter, λ, is used to avoid excessively large updates in case of an occasionally small inputs. For $p = 2$, NLMP (8) reduces to the Normalized-LMS algorithm [8].

The second algorithm, Normalized Least Mean Absolute Deviation (NLMAD), corresponds to the case of $p = 1$ in (8) with the following time update:

$$
\underline{w}(k+1) = \underline{w}(k) + \beta \, \frac{sgn(e(k))}{\|\underline{x}(k)\|_1 + \lambda} \, \underline{x}(k). \quad (9)
$$

This adaptation scheme is especially useful when the characteristic exponent, α, either is unknown or varying in time. Among the stable distributions the heaviest tail occur for the Cauchy distribution, $\alpha = 1$. By selecting $p = 1$ the update term is guaranteed to have a finite magnitude for all $1 < \alpha \leq 2$. Due to the above reasons NLMAD is a safe choice for the adaptation.

Recently, another class of normalized LMS type algorithms are also reported in [9, 10]. These algorithms are different from ours and they are developed in different context for white Gaussian input and Laplacian noise.

3. USE OF PRE-NONLINEARITY IN ADAPTIVE FILTERING

In this section the performance of LMS and RLS algorithms running on nonlinearly transformed data will be investigated. In this paper, we consider the use of a softlimiter as

shown in Figure 1. The motivation behind this approach is able to reduce the effect of spiky characteristic of the α-stable data. This type of regularization have been used in robust signal processing applications [11]. It can be easily shown that any random process which is passed through a softlimiter has finite variance. Thus, the LMS and RLS algorithms can be used in adaptation process after the input and reference signals have been soft-limited. The optimal filter coefficients which LMS and RLS converge are biased. However, the bias so introduced can be kept at a reasonably small level by a proper selection of threshold value. The use of softlimiter reduces the spiky characteristics of input data hence a much smoother convergence can be expected. Because of the use of a nonlinear mapping we call the well-known LMS and RLS algorithms as NMLMS and NMRLS. One noteworthy feature of this technique is that it has the same computational complexity as well-known LMS and RLS algorithms. Because of the nonlinear mapping involved we call the proposed algorithms as NMLMS and NMRLS. A sample sequence of AR process disturbed by α-stable ($\alpha = 1.8$) noise and the output sequence after the soft limiter are shown in Figure 2.
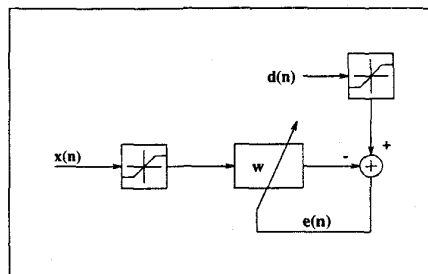

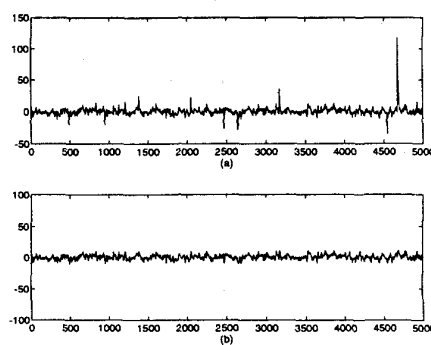Figure 1: Transform domain adaptive filtering block diagram.

Figure 2: A sample AR process disturbed by α-stable ($\alpha = 1.8$) noise (a), and the output process after the soft limiter (b).

1401

4. SIMULATION STUDIES

In simulation studies we consider $AR(N)$ α-stable processes, which are defined as follows,

$$x(n) = \sum_{i=1}^{N} a_i x(n-i) + u(n) \qquad (10)$$

where $u(n)$ is a α-stable sequence of i.i.d random variables. The common distribution of $u(n)$ is chosen to be an even function ($\beta = 0$), and the gain factors are all set to one ($\gamma = 1$) without loss of generality. It can be shown that $x(n)$ will also be a α-stable random variable with the same characteristic exponent when $\{a_i\}$ is an absolutely summable sequence [5, 6].

Two sets of simulation studies are performed. In the first set, the adaptation algorithms NLMAD, NLMP, LMAD, LMP and LMS are compared for a second order α-stable AR process with a fixed characteristic exponent, $\alpha = 1.2$. In the second set the performances of NLMAD, NLMP, NMLMS and NMRLS algorithms are compared for a second order α-stable AR process with different values of the characteristic exponent. For both sets, the tap weights are obtained by averaging 40 independent trials of the experiment and for each trial, a different computer realization of the process $\{u(n)\}$ is used. To get a fair comparison between algorithms the step sizes of adaptive algorithms are chosen in such a way that they all had a comparable steady-state variance. For both simulation set the coefficients of AR(2) is chosen as $a_1 = 0.99$ and $a_2 = -0.1$.
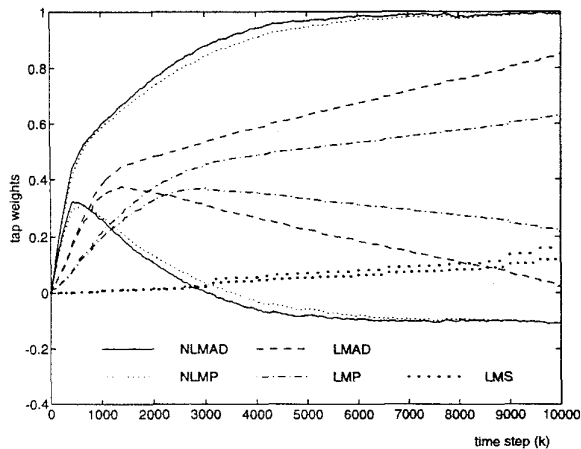

Figure 3: *Transient behavior of tap weights in the NLMAD, NLMP, LMAD, LMP and LMS algorithms with $\alpha = 1.2$.*

In the first part of the simulations, AR parameters $\underline{a}$ are estimated by a 2nd order LMP, LMAD, NLMP, NLMAD and LMS algorithms. The plot of the tap weights is given in Figure 3. In the first part we observed that the normalized algorithms NLMAD and NLMP outperformed other algorithms. Therefore, in the second part the performances of NMLMS and NMRLS are only compared to NLMAD and NLMP algorithms.

In the second part of the simulations, AR parameters $\underline{a}$ are estimated by a 2nd order NLMP, NLMAD, NMLMS and NMRLS algorithms for two different α-stable AR processes with $\alpha = 1.2$ and $\alpha = 1.8$. The plots of the tap weights for NLMAD, NLMP and NMLMS algorithms are given in Figure 4 and Figure 5 for $\alpha = 1.2$ and $\alpha = 1.8$, respectively. The tap weights convergence performance of the NMRLS is given in Figure 6 for $\alpha = 1.2$ and $\alpha = 1.8$.
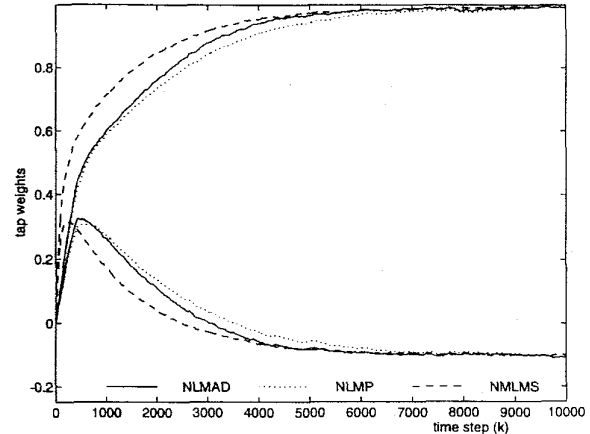

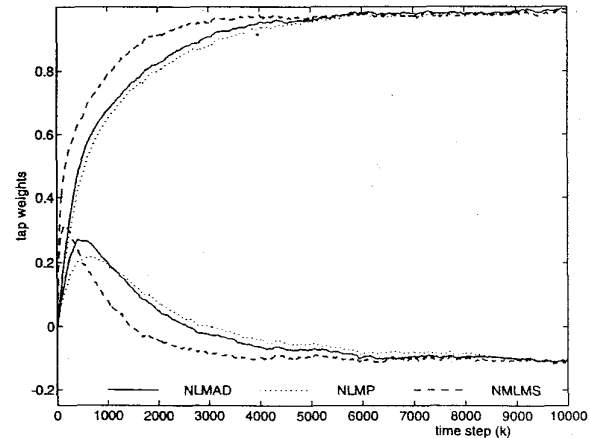Figure 4: *Transient behavior of tap weights in the NMLMS, NLMAD, NLMP algorithms with $\alpha = 1.2$.*

Figure 5: *Transient behavior of tap weights in the NMLMS, NLMAD, NLMP algorithms with $\alpha = 1.8$.*

5. CONCLUSION

In this paper, new adaptive filtering approaches in the presence of α-stable random processes are introduced. These approaches are developed with the motivation of p-norm normalization in ℓ_p spaces $1 \leq p \leq 2$, and the use of prenonlinearity in adaptive filtering. In our simulation studies the normalized algorithms NLMAD and NLMP outperform

the LMAD, LMP and LMS type algorithms. The use of pre-nonlinearity in LMS type algorithm exhibits a faster convergence than NLMAD and NLMP algorithms in the tap weight adaptations. However, pre-nonlinearity introduced an off-set to the steady state values of the tap weights. In our simulation examples, these off-set values are negligible, but it should be observed for higher order α-stable processes. The convergence of the RLS algorithm with the pre-nonlinearity outperforms other algorithms with a higher steady state variance. Also, NMRLS algorithm introduces an off-set, especially for low α values. The use of other non-linearities and the effect of off-set for higher order systems will be investigated as a future work.
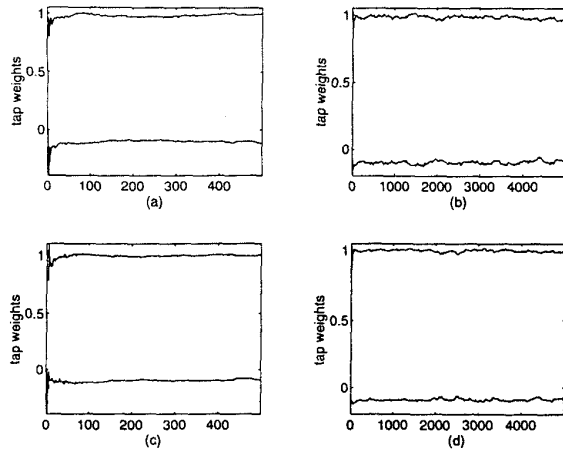

Figure 6: *Transient behavior of tap weights in the NMRLS algorithm with $\alpha = 1.8$ (a),(b), and $\alpha = 1.2$ (c),(d).*

6. REFERENCES

[1] B. Mandelbrot and J. W. Van Ness, "Fractional brownian motions, fractional noises, and applications," *SIAM Review*, vol. 10, pp. 422-437, 1968.

[2] S.S. Pillai, M. Harisankar, "Simulated performance of a DS spread spectrum system in impulsive atmospheric noise," *IEEE Trans. Electromagnetic Compat.*, vol. 29, pp. 80-82, 1987.

[3] M. Bouvet and S. C. Schwartz," Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acoust. Speech and Signal Proc.*, vol. 37, pp. 621-626, 1989.

[4] O. Arıkan, A.E. Çetin and E. Erzin, "Adaptive filtering for non-Gaussian stable processes," presented in *Twenty-eight Annual Conference on Information Sciences and Systems*, Princeton, N.J., March 1994.

[5] M. Shao and C. L. Nikias, "Signal Processing with fractional lower order moments: Stable Processes and their applications," *Proc. IEEE*, vol. 81, pp. 986-1009, 1993.

[6] Y. Hosoya, "Discrete-time stable processes and their certain properties," *Ann. Prob.*, vol. 6, no. 1, pp. 94-105, 1978.

[7] J. M. Cioffi, "An Unwindowed RLS Adaptive Lattice Algorithm," *IEEE Trans. on ASSP*, vol. 36, no. 3, pp. 365-371, March 1988.

[8] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Prentice Hall, NJ, 1985.

[9] N.L. Freire and S.C. Douglas, "Adaptive cancellation of geomagnetic background noise using a sign-error normalized LMS algorithm," *Proc. IEEE International Conf. on Acoustics, Speech, and Signal Processing*, vol 3, pp. 523-526, April 1993. ¿

[10] S.C. Douglas, "A Family of Normalized LMS Algorithms," *Signal Processing Letters*, vol. 1, no. 3, pp. 49-51, March 1994.

[11] S.A. Kassam and H.V. Poor, "Robust techniques for signal processing: a survey," *Proc. of IEEE*, vol. 73, pp. 433-481, March 1985.