Adaptive Filtering for Non-Gaussian Stable Processes

Orhan Arikan, A. Enis Çetin, and Engin Erzin

Abstract—A large class of physical phenomenon observed in practice exhibit non-Gaussian behavior. In this letter, 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 a noise is a requirement of many practical problems. Since direct application of commonly used adaptation techniques fail in these applications, new algorithms for adaptive filtering for α-stable random processes are introduced.

I. 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]–[4]. There exists an important class of distributions known as α-stable distributions [5] that 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:

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

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 that 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 letter, we introduce new algorithms for adaptive filtering under additive α-stable noise with finite mean corresponding to $1 \le \alpha < 2$. The performance of these approaches will be compared with those of the already existing approaches in the literature. The algorithms are presented in Section II, and simulation results are given in Section III.

II. 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

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

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

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

where μ is the step size that 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]:

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

which is also known as the signed-LMS algorithm.

In this letter, we introduce two normalized adaptation algorithms with the motivation of the normalized-LMS algorithm. The first one (the normalized least mean p-norm (NLMP) algorithm) uses the following update:

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

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 input. For $p = 2$, NLMP (8) reduces to the normalized-LMS algorithm [8].

The second algorithm (the normalized least mean absolute deviation (NLMAD)) corresponds to the case of $p = 1$ in (8) with the following time update:

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

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]. These algorithms are different from ours and they are developed in different context.

III. SIMULATION STUDIES

In [5], the performance of the LMAD algorithm is compared with that of the LMS algorithm for a first order α-stable AR process and it was observed that LMAD outperforms LMS especially for low α values. In this section, we compare the performances of the new algorithms with the LMAD, LMP, and LMS algorithms.

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 ($\gamma = 1$)
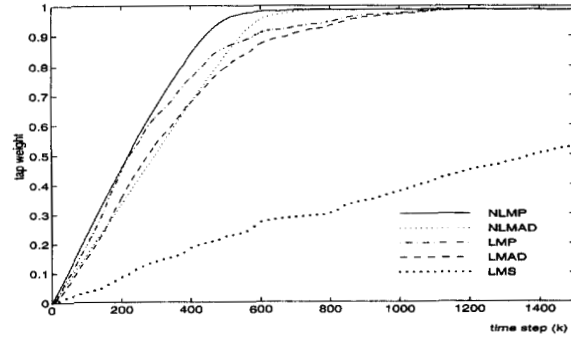

Fig. 1. Transient behavior of tap weight adaptations in the NLMP, NLMAD, LMAD, LMP, and LMS algorithms with $\alpha = 1.2$ for $AR(1)$ process. AR parameter is chosen as $a_1 = 0.99$.

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 are compared for the cases of first- and second-order α-stable AR processes with a fixed characteristic exponent $\alpha = 1.2$. In the second set, the performances of LMAD, NLMAD, and NLMP algorithms are compared when a fourth-order α-stable AR process with different values of the characteristic exponent is used. For both sets, the tap weights are obtained by averaging 100 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 size of the LMS is adjusted as large as possible while ensuring the convergence. Then the step sizes of other algorithms are chosen in such a way that they all had a comparable steady-state error.

In the first simulation set, tap weight adaptation is performed for $AR(1)$ and $AR(2)$ processes with first- and second-order LMP, LMAD, NLMAD, and LMS algorithms, respectively. The coefficient of $AR(1)$ process is chosen as $a_1 = 0.99$. In Fig. 1, the transient behaviors of the tap weight adaptations for $AR(1)$ process are plotted. The NLMP and NLMAD algorithms introduced in this study has a better convergence behavior than the LMAD [5], LMP [5], and the LMS algorithms.

For $AR(2)$ process, the coefficients are chosen as $a_1 = 0.99$ and $a_2 = -0.1$. In Fig. 2, the transient behaviors of the tap weight adaptations for $AR(2)$ process are plotted. In this case, the NLMP and NLMAD algorithms again show faster convergence.

Second simulation set tests the performances of the LMAD, NLMAD, and NLMP algorithms for α values $1.2, 1.5$, and 1.9 with an $AR(4)$ process. The $AR(4)$ process is chosen as a fourth-order LPC synthesis filter of a voiced speech frame [10] $A(z) = 1.323z^{-1} - 0.152z^{-2} - 0.097z^{-3} - 0.115z^{-4}$. The adaptation performances are plotted in Fig. 3. The NLMAD and NLMP algorithms have comparable performances for small α values both of which converge faster than the LMAD
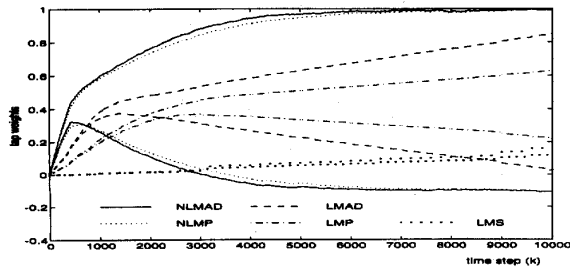
Fig. 2. Transient behavior of tap weight adaptations in the NLMP, NLMAD, LMAD, LMP, and LMS algorithms with $\alpha = 1.2$ for $AR(2)$ process. AR parameters are chosen as $a_1 = 0.99$ and $a_2 = -0.1$.
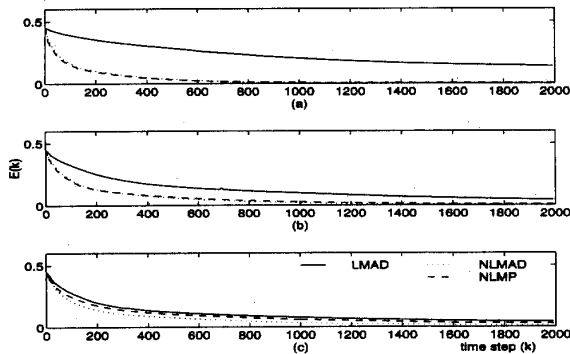

Fig. 3. Tap weight error powers for (a) $\alpha = 1.2$, (b) $\alpha = 1.5$, and (c) $\alpha = 1.9$ in LMAD, NLMAD, and NLMP algorithms. $E(k) = \|\underline{w}(k) - \underline{w}^*\|^2$, where $\underline{w}(k)$ and $\underline{w}^*$ are the current tap weight and optimal solution vectors, respectively.

algorithm. When α gets larger, e.g., $\alpha = 1.9$, the performances of all the three algorithms are comparable to each other.

IV. CONCLUSION

In this paper, new adaptive filtering algorithms in the presence of α-stable random processes are introduced. These

algorithms are developed with the motivation of p-norm normalization in ℓ_p spaces $1 \leq p \leq 2$. The performances of these normalized algorithms are found to be superior to that of LMS and LMAD algorithms in simulation studies. Based on the experience gained in the simulation studies, it is observed that a safe choice of p value is 1 in the case of imprecise knowledge of α. This corresponds to the use of NLMAD algorithm in such cases.

Transformation of both the input and the desired signal by using a reversible nonlinearity to another domain in which Hilbert space framework exists is under investigation. The effects of the adaptation in the transform domain will be considered. In addition, a generalization of recursive least squares (RLS) algorithm to α-stable processes will be investigated as a future work.

REFERENCES

[1] B. Mandelbrot and J. W. Van Ness, "Fractional brownian motions, fractional noises, and applications," *SIAM Rev.*, vol. 10, pp. 422-437, 1968.
[2] S. S. Pillai and M. Harisankar, "Simulated performance of a DS spread spectrum system in impulsive atmospheric noise," *IEEE Trans. Electromagn. 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 Signal Processing*, vol. 37, pp. 621-626, 1989.
[4] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. 19, pp. 106-127, 1977.
[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. Acoust, Speech, Signal Processing*, vol. 36, no. 3, pp. 365-371, Mar. 1988.
[8] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[9] S. C. Douglas, "A family of normalized LMS algorithms," *Signal Processing Lett.*, vol. 1, no. 3, pp. 49-51, Mar. 1994.
[10] J. R. Deller *el al.*, *Discrete-Time Processing of Speech Signals*. New York, Macmillan, 1993.