A theoretical analysis of the median LMF adaptive algorithm

Bysted, Tommy Kristensen; Rusu, C.

Link back to DTU Orbit

A THEORETICAL ANALYSIS
OF THE MEDIAN LMF ADAPTIVE ALGORITHM

Tommy K. Bysted[1], Corneliu Rusu[2]

[1]Department of Information Technology, Technical University of Denmark
DK-2800 Lyngby, DENMARK
[2]Signal Processing Laboratory, Digital Media Institute, Tampere University of Technology
P.O.BOX 553, SF-33101, Tampere, FINLAND

ABSTRACT

Higher order adaptive algorithms are sensitive to impulse interference. In the case of the LMF (Least Mean Fourth), an easy and effective way to reduce this is to median filter the instantaneous gradient of the LMF algorithm. Although previous published simulations have indicated that this reduces the speed of convergence, no analytical studies have yet been made to prove this. In order to enhance the usability this paper presents a convergence and steady-state analysis of the Median LMF adaptive algorithm. As expected this proves that the Median LMF has a slower convergence and a lower steady-state error compared to the LMF.

1. INTRODUCTION

In 1984 Walach and Widrow introduced the LMF algorithms as a supplement to the widely used LMS algorithm [1]. Since then it has been extensively used especially for data echo cancellation and channel equalization. In these applications the involving signals have short tailed distributions and consequently the LMF outperforms the conventional LMS algorithm. Although widely used, an overlooked problem of the LMF algorithm is its high distribution sensitivity which was theoretically shown in [2]. Due to this sensitivity the performance of the LMF algorithm is inferior for long tailed distributed data. This was illustrated by an example in [3] where binary distributed data was corrupted by impulses, which completely destroyed the convergence of the conventional LMF algorithm. In order to reduce this problem a new adaptive algorithm combining the LMF algorithm and the nonlinear median filter was presented. The Median LMF algorithm is defined as the conventional LMF algorithm where the instantaneous gradient is replaced by its median:

$$\nabla_{MED} = MED \{ \nabla (k) \}_N \qquad (1)$$

where N is the length of the median filter and is usually chosen odd [4], and

$$\nabla (k) = -e (k)^3 X (k),$$
$$MED \{ \nabla (k) \}_N = MED \{ \nabla (k), ..., \nabla (k-N+1) \}. \qquad (2)$$

This gives the following equation for the coefficients update of the adaptive filter:

$$W (k + 1) = W (k) - 4\mu \nabla_{MED} (k). \qquad (3)$$

The behaviour of this new algorithm (called Median LMF adaptive algorithm) is of course different from the LMF algorithm, and in [2] the following three characteristics were illustrated by simulations:

- Impulses are removed from the instantaneous gradient, giving a more regular convergence;
- The speed of convergence is reduced;
- A larger convergence factor can be chosen.

The first characteristic is of course obvious due to the median filtering, which removes large gradient values in the update of the coefficients. In order to verify the last two characteristics a convergence and steady-state analyses will be performed in the following Section. Because of the lowpass nature of the median filter the gradient variance will be reduced. This is more pronounced for long tailed distributions, where the median filter has a high attenuation [4]. This suggests that Median LMF is less distribution sensitive than the LMF algorithm. This is illustrated in Section 3.

2. CONVERGENCE AND STEADY-STATE STUDY

2.1 Convergence Analysis

The convergence properties of adaptive algorithms are important in order to understand the transient behaviour. For most algorithms the convergence analysis is relatively simple, but due to the nonlinearity of the median operator, this is quite difficult for the Median LMF algorithm. A way to make the analysis tractable is to use the following two assumptions:

- **Assumption 1.** The coefficients $W(k)$ are slowly varying and can be considered constant for N consecutive samples.
- **Assumption 2.** If u and v are two (non-correlated)

i.i.d. sequences, each having symmetrical pdf and mean value 0, and k, respectively, then:

$$E\{\text{MED}\{u+v\}_N\} = E\{\text{MED}\{v\}_N\} = k \qquad (4)$$

If only u is symmetrical distributed, then the equation (4) is only approximately true. Although a complete convergence analysis has been made for the LMF (see [5]), it is customary to restrict it to the vicinity of the true solution, i.e. the values of the weight deviation vector $V(k)=W(k)-W_*$ are small [1]. Consequently the higher order terms of $V(k)$ can be neglected and the mean value of the weight deviation vector can be approximated as

$$E\{V(k+1)\} \approx E\{V(k)\}$$
$$-12\mu E\{\text{MED}\{n^2(k)X(k)X^T(k)V(k)\}_N\}, \qquad (5)$$

where the approximation in assumption 2 has been used[1]. The argument to the i'th median operator of the gradient vector, called $\nabla_i(k)$ can be written as:

$$n^2(k)\left(x^2(k-i)v_i(k) + \sum_{m=0,m\neq 1}^{L-1} x(k-i)x(k-m)v_m(k)\right)$$

If the input sequence $x(k)$ is assumed to be symmetrical i.i.d. having zero mean, then the use of assumption 1, assumption 2 and the independence of $n(k)$, $x(k)$ and $V(k)$ will lead to the equation:

$$E\{V(k+1)\} \approx E\{V(k)\}$$
$$-12\mu E\{\text{MED}\{n^2(k)x^2(k)\}_N\}E\{V(k)\} \qquad (6)$$

We obtained a very similar formula with the one from conventional LMF algorithm, and consequently the Median LMF will converge in the mean if the convergence factor is chosen in the interval:

$$0 < \mu < \frac{1}{6E\{\text{MED}\{n^2(k)x^2(k)\}_N\}} \qquad (7)$$

It is difficult to compute the upper limit of this inequality, since it requires the pdf of the product $n^2(k)x^2(k)$. For ordinary symmetrical distributions like double exponential, Gaussian and uniform, we have

$$E\{\text{MED}\{n^2(k)x^2(k)\}_N\} < E\{n^2(k)x^2(k)\}. \qquad (8)$$

Thus the convergence factor of the Median LMF can be chosen larger than for the LMF algorithm, especially for long tailed distributions. In practice it is though easier to choose the same step-size as for the LMF algorithm. The time-constant for the adaptation can easily be found

---

1. Here $x(k)$ is the near-end signal, $n(k)$ is the far-end signal, and L is the length of the adaptive echo canceller [3].

from equation (6) as:

$$\tau = \frac{1}{12\mu E\{\text{MED}\{n^2(k)x^2(k)\}_N\}} \qquad (9)$$

The Median LMF will therefore have a slower convergence than the conventional LMF for the same μ, as illustrated by simulations in [3]. This is of course most pronounced for long tailed distributions and large values of N where the median has the best attenuation.

2.2 Steady-State Analysis

We expect a better steady-state performance than for the LMF due to the reduction of the sensitivity of the gradient estimate. We follow once again [1], by finding an approximation of the covariance matrix of the weight deviation vector:

$$\text{cov}V(k) \approx \frac{2\mu E\{\left(\text{MED}\{n^3(k)x(k)\}_N\right)^2\}}{3E\{\text{MED}\{n^2(k)x^2(k)\}_N\}}I. \qquad (10)$$

The covariance matrix has a diagonal structure, thus the components of the weight deviation vector are non-correlated. This can be exploited in the definition of the misadjustment to ignore the cross-terms [1]:

$$M \approx \frac{E\{\left(V^T(k)X(k)\right)^2\}}{E\{n^2(k)\}} \qquad (11)$$

It is possible to approximate this misadjustment by:

$$M \approx \frac{2\mu L}{3} \cdot \frac{E\{\left(\text{MED}\{n^3(k)x(k)\}_N\right)^2\}E\{x^2(k)\}}{E\{\left(\text{MED}\{n^2(k)x^2(k)\}_N\right)^2\}E\{n^2(k)\}} \qquad (12)$$
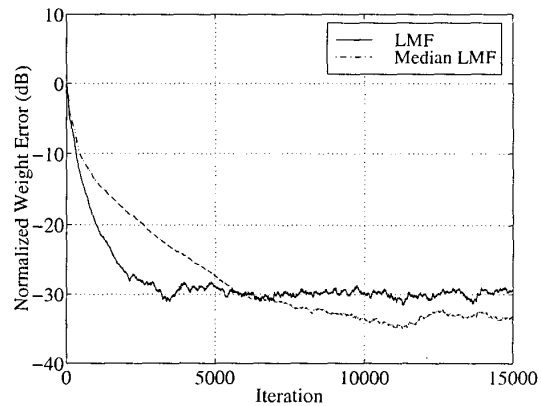
We can compare this equation with LMF case where we
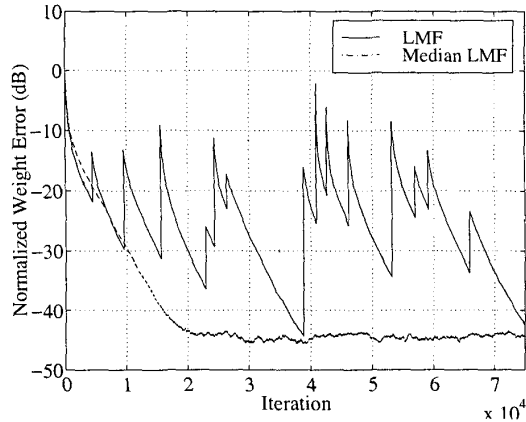

Fig.1. Gaussian distributed signals

Fig.2: Signals corrupted by impulsive noise

Fig.3: Double exponential distributed signals

have

$$M \approx \frac{2\mu L}{3} \cdot \frac{E\{\left(n^3(k)x(k)\right)^2\}E\{x^2(k)\}}{E\{\left(n^2(k)x^2(k)\right)^2\}E\{n^2(k)\}} \quad (13)$$

A comparison of these two equations demonstrates that the theoretical misadjustment of the Median LMF algorithm is lower than the misadjustment of the LMF algorithm (except for binary distributed data where the two algorithms have the same misadjustment). This lower misadjustment is expected due to the variance reduction of the instantaneous gradient made by the median filter, and is of course most pronounced for long tailed pdf's.

3. SIMULATIONS

Two different examples are presented in this section. First the sensitivity of the LMF algorithm to impulse noise and the robustness of the Median LMF is illustrated. Secondly the performance characteristics of the Median LMF adaptive algorithm for long tailed distributed signals are shown (slower convergence and lower misadjustment). The performance measure used in all of the figures in this section is the normalized norm *p(k)* of the weight deviation vector [3].

3.1 Data Echo Cancellation

The setup used in this example is the data echo cancellation, which is an application where the LMF algorithm has been widely used [6]. The main purpose of an echo canceller for data transmission is to cancel the echo generated by the imperfect impedance matching at a hybrid transformer. This echo will, if not removed, disturb or destroy the desired signal sent from the other end of the line (called far-end signal). Here the echo path is modelled as a simple low-pass filter [6], where
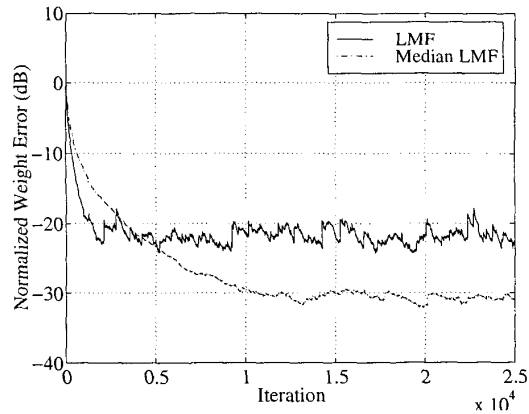
$a=0.800250$. Both the input signal to the adaptive filter and to the echo path are modelled as random binary sequences having the values $\{+1,-1\}$ and $\{+f,-f\}$ respectively. Impulses are added to the received sequence to simulate transmission and decoding errors. This is done simply by changing the polarity of the received sample if an error occur (the probability of error is *p*). For this example L=32, p=0.0005, f^2=0.1 and step-size have been chosen 0.002. The performance curve is depicted in Figure 2, and as expected this demonstrates the noisy convergence of the LMF when transmission or decoding errors occur. This is avoided by the Median LMF (plotted in the same figure for N=3), and this demonstrates the robustness of the Median LMF compared to the conventional LMF algorithm.

3.2 Distribution Sensitivity

In this example the LMF and the Median LMF algorithms are used to identify the system [1]:

$$H(z) = 0.1 + 0.2z^{-1} + 0.3z^{-2} + 0.4z^{-3} + 0.5z^{-4} \\ + 0.4z^{-5} + 0.3z^{-6} + 0.2z^{-7} + 0.1z^{-8}. \quad (14)$$

To illustrate the dependency of the convergence and the steady-state performance upon the signal distribution, two different distributions of the input and noise signals are used. In Figure 1 the signals are Gaussian distributed whereas double exponential signals are used in Figure 3. In both examples SNR=10 dB, N=3, L=9 and step-sizes are 0.001. These two figures clearly demonstrates the lower misadjustment and the slower convergence of the Median LMF algorithm. This is more obvious for the double exponential distributed signals in Figure 3. Having the attenuation capabilities of the median filter in mind (best attenuation for long tailed noise), this behaviour is expected.

In Section 2.1 it was mentioned that the Median LMF algorithm is less distribution sensitive than the
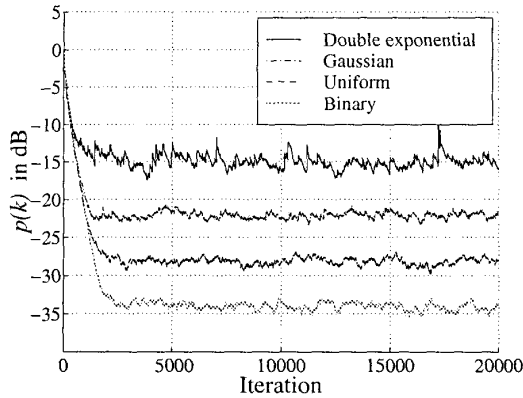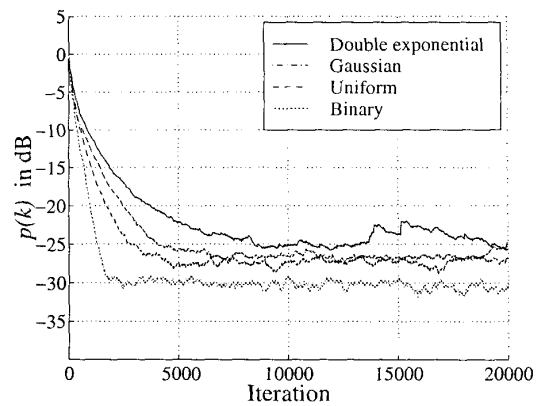
Fig.4 Different distributions (LMF)

Fig.5. Different distributions (Median LMF)

conventional LMF algorithm. The high sensitivity of the LMF algorithm is due to the variance of the instantaneous gradient estimate, which is large for long tailed distributions. Because the median filter has a good attenuation of these distributions the inclusion of the median filtering of the gradient reduces this problem. This is also illustrated in this section. To demonstrate we use the same setup, i.e. a system identification where the system to be identified is described by same equation. The convergence factor is identical 0.005 for both algorithms: Median LMF (N=3) and LMF. Thus we eliminate its effect on the steady-state error. In Figure 4 and 5 the results of the simulation are plotted, and as described the Median LMF has a more uniform steady-state performance. While the steady-state error for the LMF has a 20 dB spread, this is now reduced to 5 dB without changing the convergence factor. The cost for this is a slower convergence and an increase of the steady-state error for the binary distributed data, which can be explained by introduction of correlation in the gradient [1]. If the length of the median filter is increased, the spread in performance for small median filter windows (N=5 to N=9) is even smaller. Unfortunately will the spread increase again for large median filter windows, due to the increased steady-state error for binary sequences. Therefore, the Median LMF should be restricted to small median filters.

4. CONCLUSIONS

The Median LMF adaptive algorithm combines the conventional LMF algorithm and the median filter. by combining these two the sensitivity of the LMF algorithm to long tailed distributed signals is reduced. In this paper a convergence and a steady-state analysis of the Median LMF algorithm is presented. Due to the

nonlinear nature of the median filter these analyses are difficult and in order to make them feasible some assumptions were necessary. Both analyses were verified in examples, which showed that the Median LMF has a slower convergence than the conventional LMF. This can though be reduced by choosing a larger convergence factor.

References

[1] E. Walach, B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and Its Family", *IEEE Transactions Inf. Theory*, vol. IT-30, pp.275-283, 1984.

[2] T.K. Bysted, "Error analysis and error reduction in cepstrum transformation and the Median LMF adaptive algorithm", Technical University of Denmark, 1999.

[3] T.K. Bysted, C. Rusu, "Improving the stability of the LMF adaptive algorithm using the median filter", *IEEE Norsig98*, Vigso, Denmark, June 1998, pp.157-160.

[4] I. Pitas, A.N. Venetsanopoulos, "Order Statistics in Digital Image Processing" *Proc.IEEE*, vol.80, pp.1893, 1992.

[5] O. Tanrikulu, A.G. Constantinides, "A Mean Square Convergence Analysis of the Least Mean Fourth (LMF) Adaptive Algorithm", *ICT-96*, pp. 114-117, 1996.

[6] S.A.Shah, C.F.N. Cowan, "Modified Stochastic Gradient Algorithm Using Nonquadratic Cost Functions for Data Echo Cancellation", *IEE Proc.Vis.Image Signal Process.*, vol.142, pp.187-191, 1995.