14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

ANALYSIS OF STEADY-STATE EXCESS MEAN-SQUARE-ERROR OF THE LEAST MEAN KURTOSIS ADAPTIVE ALGORITHM

Junibakti Sanubari

Department of Electronics Engineering, Satya Wacana University
Jalan Diponegoro 52 − 60, Salatiga, Indonesia
Phone: + (62)298-311884. Fax: + (62)298-321433. Email: jsawitri@hotmail.com
Web: www.uksw.edu

ABSTRACT

In this paper, the average of the steady state excess mean square error (ASEMSE) of the least mean kurtosis (LMK) adaptive algorithm is theoretically derived. It is done by applying the energy conservation behavior of adaptive filters and it is based on the n-th order correlations and cumulants theory. By doing so, the behavior of the recently proposed LMK can be predicted, so that it can be widely used. The behavior is compared with the various adaptive algorithms. Our study shows that it is possible to adjust the performance of the LMK. When the step size μ is carefully selected, the performance of the LMK can outperform the LMS.

1. INTRODUCTION

Traditionally, the coefficient of the adaptive filter $\mathbf{w}$ is adjusted using the least mean square (LMS) algorithm. The LMS algorithm is largely used in real-time adaptive filtering applications for communications, control engineering, processing of biological signals and more recently active noise and vibration controls [1]. It is very popular due to its simplicity. In the LMS, the coefficient of the adaptive filter is adjusted so that the square of the error $e(n)$ between the output of the adaptive filter $y(n)$ and the desired signal $d(n)$ is minimal[1]; minimizing $\sum_n e^2(n)$.

In applications, the input signal of the adaptive filter $x(n)$ is contaminated with additive noise $v(n)$. By minimizing $\sum_n e^2(n)$, we minimize the power of the noise; σ_v^2. No special consideration is taken regarding the detail statistical behavior of the noise. All parts of input signal are given an equal weighting. The obtained solution is more determined by large amplitude noise portion than that by small amplitude noise portion. The obtained solution is very much affected by large amplitude noise portion. Further improvements is hoped to be obtained when the statistical behavior of the noise is considered.

The above mentioned problems are encountered by assuming that remaining error $e(n)$ is Gaussian. Remembering the fact that the signal $e(n)$ is a summation of many stochastic processes and the central limit theory [2], assuming that $e(n)$ is Gaussian is reasonable. Since many stochastic processes with different probability density function (PDF) can have equal mean or variance and no stochastic processes can have normalized kurtosis

$$\gamma_e^4 = \frac{E\{e^4(n)\}}{E^2\{e^2(n)\}} = 3, \tag{1}$$

except Gaussian process [3], in this paper we evaluate $\gamma_e^4(n)$ to test the Gaussianity of $e(n)$. We define the fourth-order cumulants of the error as $E\{e^4(n)\}$ and its variance as σ_e^2. In this paper we study an adaptive algorithm where $\mathbf{w}$ is adjusted by forcing $\gamma_e^4 = 3$; which is named least mean kurtosis (LMK) [8]. Thus, in LMK algorithm, the coefficient of the adaptive filter is solved by forcing the error signal to be Gaussian.

Many adaptive algorithms that are based on higher order moments of the error signal due exist [4]−[8] and LMK is one of such algorithm [8]. Those algorithms have been shown to outperform LMS in some important applications. The practical use of such algorithms, however, has been largely restricted due to lack of accurate analytical models to predict their behavior. Since some approximations have also been introduced to reduce the calculation complexity of the LMK, it is interesting to study its analytic behavior to further open the possible applications of the algorithm.

The model of the behavior of the LMK adaptive algorithm with Gaussian input has been studied [9]. The behavior is studied for zero-mean contaminated input signals. Since the probability density function (PDF) of the noise is unknown, in this paper the average of the steady-state excess mean square error (ASEMSE) of LMK algorithm is studied for both the symmetrical and the non-symmetrical PDF noises.

2. THE BASIC ADAPTIVE SYSTEM

Consider a noisy measurements $d(n)$ that arise from the linear model

$$d(n) = \mathbf{x}^T(n)\mathbf{w}^0 + v(n) \tag{2}$$

where the input data vector $\mathbf{x}(n)$ and desired unknown parameter vector $\mathbf{w}^0$ are

$$\mathbf{x}(n) = [\ x(n) \quad x(n-1) \quad \cdots \quad x(n-M)\ ]^T \tag{3}$$

and

$$\mathbf{w}^0 = [\ w_1^0(n) \quad w_2^0(n) \quad \cdots \quad w_M^0(n)\ ]^T, \tag{4}$$

respectively. The superscript T denotes the vector or matrix transpose. $v(n)$ accounts for the noise signal and the order of the system is M. Both $v(n)$ and $x(n)$ are stochastic in nature. The so-called estimation error is defined as

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n), \tag{5}$$

where $\mathbf{w}(n)$ is the estimate of $\mathbf{w}^0$ at iteration n. Since (5) shows that indeed the error signal $e(n)$ is resulted from summing many stochastic process; $d(n) and x(n-l)$; for $0 \leq l \leq M$, it is reasonable to assume that $e(n)$ is Gaussian. Therefore, it is expected to gain improvements when the optimal solution is calculated by forcing $\gamma_e^4 = 3$.

By defining

$$\mathbf{q}(n) = \mathbf{w}^0 - \mathbf{w}(n), \qquad (6)$$

we can obtain

$$e(n) = e_a(n) + v(n), \qquad e_a(n) = \mathbf{x}^T(n)\mathbf{q}(n) \qquad (7)$$

The estimation of $\mathbf{w}^0$ at time index $(n+1)$ is recursively determined using the gradient method [1]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_a J'_{\text{LMK}}(e_n)\mathbf{x}(n). \qquad (8)$$

The symbol $'$ indicates the first derivatives; i.e. $f'(x) = \frac{\partial f(x)}{\partial x}$. From the definition of $q(n)$, $e_a(n)$ is also called the error signal due to parameter mismatch and the convergence constant is denoted as μ_a.

3. THE LEAST MEAN KURTOSIS CRITERION

Instead of using the definition of the normalized kurtosis of a stochastic in (1), this paper uses

$$J_{\text{LMK}}(e(n)) = 3E^2\{e^2(n)\} - E\{e^4(n)\}. \qquad (9)$$

as the objective function to be minimized to obtain the optimal solution. It is selected, so that no unknown parameters are in the denominator.

Since all $e(n)$ has to be available to evaluate $E\{e^2(n)\}$ and $E\{e^4(n)\}$ and it is impossible in adaptive algorithm, in this paper the short average is used as an approximation

$$E\{e^3(n)\} \approx e^3(n), \qquad (10)$$

and

$$E\{e^2(n)\} \approx e^2(n) + \beta e^2(n-1) + \beta^2 e^2(n-2). \qquad (11)$$

Substituting the approximation in (10) into (8), the updating is done using

$$\mathbf{w}(n+1) \;=\; \mathbf{w}(n) + \mu_{\text{LMK}}\kappa(e(n))\,\mathbf{x}(n). \qquad (12)$$

where $\mu_{\text{LMK}} = 4\mu_a$ and

$$\kappa(e(n)) = 2e^2(n) + 3\beta e^2(n-1) + 3\beta^2 e^2(n-2), \qquad (13)$$

or by using (7), we get

$$\kappa(e(n)) = (e_a(n) + v(n))\,(\alpha_1 + \alpha_2 + \alpha_3), \qquad (14)$$

where

$$\begin{aligned}
\alpha_1 &= 2e_a^2(n) + 3\beta e_a^2(n-1) + 3\beta^2 e_a^2(n-2) \\
\alpha_2 &= 2v^2(n) + 3\beta v^2(n-1) + 3\beta^2 v^2(n-2) \\
\alpha_3 &= 4e_a(n)v(n) + 6\beta v(n-1)e_a(n-1) + \\
&\quad 6\beta^2 e_a(n-2)v(n-2).
\end{aligned} \qquad (15)$$

The selection of the positive constant β is not critical and will be discussed later. The theoretical performance of the LMK is evaluated in the following section.

4. THE STEADY STATE PERFORMANCE OF THE LEAST MEAN KURTOSIS ALGORITHM

One of the important performance of the adaptive algorithm is the steady-state of the remaining error; $\lim_{n\to\infty} q(n)$. Since the input is stochastic processes, we evaluate $\lim_{n\to\infty} E\{q(n)\}$ for all possible inputs. The evaluation is easily done by evaluating the average of the steady-state excess mean-square-error (*ASEMSE*) which is defined as

$$ASEMSE_{\text{LMK}} = \lim_{n\to\infty} E\left(|e(n)|^2\right). \qquad (16)$$

Under the often realistic assumption that [1]:

1. The noise sequence $\{v(n)\}$ is identical and independent distributed (IID).
2. The noise $\{v(n)\}$ and the input sequences $\{x(n)\}$ are independent each other. Utilizing those two assumptions and (7) and assuming that the noise is stationary the *ASEMSE* is given by:

$$ASEMSE_{\text{LMK}} = \sigma_v^2 + \lim_{n\to\infty} E\left(e_a(n)\right) \qquad (17)$$

where σ_v^2 is the variance of the noise sequence and $e_a(n)$ is defined in (7).
3. Generally, the PDF of the noise is unknown. In this paper, we assumed that the PDF of the noise $v(n)$ can be non-symmetrical; $f(-v(n)) \neq f(v(n))$; or symmetrical distributed; $f(-v(n)) = f(v(n))$.
4. The PDF of the input signals $x(n)$ is IID.

Utilizing the definition of $\mathbf{q}(n)$ in (6), (12) can be rewritten to be

$$\mathbf{q}(n+1) = \mathbf{q}(n) - \mu_{\text{LMK}}\kappa(e(n))\,\mathbf{x}(n) \qquad (18)$$

that can be taken its L_2 norm. After some algebraic manipulations, the relation in (19) can be obtained.

$$\begin{aligned}
||\mathbf{q}(n+1)||^2 \;&=\; ||\mathbf{q}(n)||^2 - 2\mu_{\text{LMK}}\kappa(e(n))\,e_a(n) \\
&\quad + \mu_{\text{LMK}}^2\kappa^2(e(n))\,\mathbf{x}^T(n)\mathbf{x}(n)
\end{aligned} \qquad (19)$$

The second term of (17) is solved by calculating

$$\begin{aligned}
E\left\{||\mathbf{q}(n+1)||^2\right\} \;&=\; E\left\{||\mathbf{q}(n)||^2\right\} \\
&\quad -2\mu_{\text{LMK}} E\left\{\kappa(e(n))\,e_a(n)\right\} \\
&\quad +\mu_{\text{LMK}}^2 E\left\{\kappa^2(e(n))\,\mathbf{x}^T(n)\mathbf{x}(n)\right\}
\end{aligned} \qquad (20)$$

When $n \to \infty$, a steady state has been reached, so that

$$E\{||\mathbf{q}(n+1)||^2\} = E\{||\mathbf{q}(n)||^2\}; \qquad (21)$$

showing that no further improvement can be gained. Therefore, from (20) we get

$$E\left(\frac{|e_a(n)|^2}{X_n}\right) = E\left(\frac{|e_a(n) - \mu_{\text{LMK}}\kappa(e(n))\,X_n|^2}{X_n}\right), \qquad (22)$$

where $X_n = \mathbf{x}^T(n)\mathbf{x}(n) = \text{tr}\{R_{xx}\}$. The symbol tr indicates the trace of the autocorrelation matrix of the input signal $x(n)$. Expanding the right hand side of (22), we obtain

$$2\mu_{\text{LMK}} E\left(e_a(n)\kappa(e(n))\right) = \mu_{\text{LMK}}^2 E\left(\kappa^2(e(n))\,X_n\right) \qquad (23)$$

Using (7), (14) and the definition of α_1, α_2 and α_3 in (15), we can obtain

$$e_a(n)\kappa(e(n)) = e_a(n)\left\{(e_a(n)+v(n))(\alpha_1+\alpha_2+\alpha_3)\right\}. \tag{24}$$

On the other hand from the definition of $\kappa(e(n))$ in (14) we can obtain

$$\kappa^2(e(n)) = \left\{(e_a(n)+v(n))(\alpha_1++\alpha_2+\alpha_3)\right\}^2, \tag{25}$$

Using the facts that:

1. M is large. Thus, based on the central limit theory [2] and (7), $e_a(n)$ can be safely assumed to be zero mean Gaussian and unity standard deviation (SD). Thus all its odd moment and its cumulant higher than two vanish.
2. $v(n)$ and $e_a(n)$ are independent
3. The number of samples is large; so that we can have $E\left\{v^2(n)\right\} \approx E\left\{v^2(n-1)\right\} \approx E\left\{v^2(n-2)\right\} = \sigma_v^2$ and $E\left\{e_a^2(n)\right\} \approx E\left\{e_a^2(n-1)\right\} \approx E\left\{e_a^2(n-2)\right\} = E\left\{e_a^2(n)\right\}$

we will obtain

$$\begin{aligned} E\left\{e_a(n)\kappa(e(n))\right\} &= 2E\left\{e_a^4(n)\right\} \\ &+ 12E\left\{v^2(n)\right\}E\left\{e_a^2(n)\right\} \end{aligned} \tag{26}$$

and

$$\begin{gathered} E\left(\kappa^2(e(n))\right) = 4E\left\{e_a^6(n)\right\} + \\ \gamma_1 E\left\{v^4(n)\right\}E\left\{e_a^2(n)\right\} + \\ \gamma_1 E\left\{e_a^4(n)\right\}E\left\{v^2(n)\right\} + 4E\left\{v^6(n)\right\} \\ +\gamma_2 E\left\{v^2(n)v^2(n-1)\right\}E\left\{e_a^2(n-1)\right\} + \\ \gamma_3 E\left\{e_a^2(n-2)\right\}E\left\{v^2(n)v^2(n-2)\right\} + \\ 9\beta^2 E\left\{v^2(n)v^4(n-1)\right\} + \\ 9\beta^4 E\left\{v^2(n)v^4(n-2)\right\}, \end{gathered} \tag{27}$$

where

$$\begin{aligned} \gamma_1 &= 60+21\beta^2+21\beta^4, \\ \gamma_2 &= 60\beta+54\beta^2+18\beta^3 \\ \gamma_3 &= 60\beta^2+18\beta^3+54\beta^4 \end{aligned} \tag{28}$$

Using the moment theorem [2] that stated $E(x^4) = 3E(x^2)$ and $E(x^6) = 15E(x^2)$ and the cumulants theory [3], we get

$$E\left\{e_a(n)\kappa(e(n))\right\} = 6E\left\{e_a^2(n)\right\} + 12\sigma_v^2 E\left\{e_a^2(n)\right\}, \tag{29}$$

and

$$\begin{gathered} E\left(\kappa^2(e(n))\right) = 60E\left\{e_a^2(n)\right\} + 6\gamma_1\sigma_v^2 E\left\{e_a^2(n)\right\} + \\ 60\sigma_v^2 + \gamma_2 m_4^v(0,1,1)E\left\{e_a^2(n)\right\} + \\ \gamma_3 E\left\{e_a^2(n)\right\} m_4^v(0,2,2) + 9\beta^2 m_6^v(0,1,1,1,1) + \\ 9\beta^4 m_6^v(0,2,2,2,2) \end{gathered} \tag{30}$$

where the N-th order moment of the noise signal $v(n)$ is defined as

$$m_N^v\left(\overbrace{s,q,\cdots}^{(N-1)\ terms}\right) = E\left\{\overbrace{v(n)v(n-s)v(n-q)\cdots}^{(N-1)\ terms}\right\}. \tag{31}$$

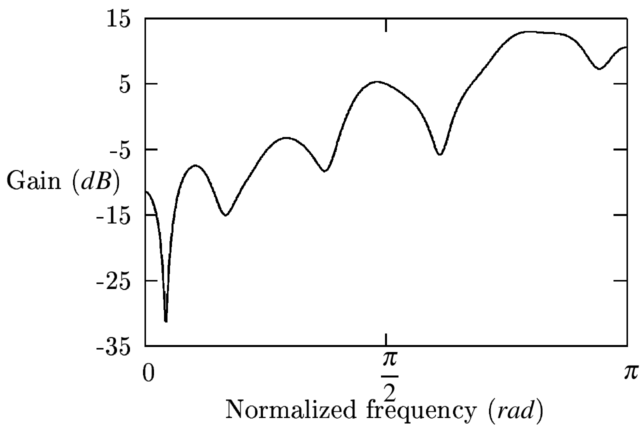

Figure 1: The frequency response of the 20th order system that was used in the simulation to obtain $ASEMSE$.

Substituting (29) and (30) into (17), we can obtaine $ASEMSE$ as given in (32). Since the numerator of (32) is always positive, we have to limit μ_{LMK} as in (34) to be sure that the denominator of (32) is positive. By doing so we can be sure that $ASEMSE$ in (32) is positive and valid. This is consistent with the limitation of the step size μ that will be explained in the next section.

Equation (32) shows that the selection of μ_{LMK} determine the resulted $ASEMSE$. $ASEMSE$ was experimentally solved from (17) by generating contaminated signals to be applied as $x(n)$. Those contaminated signals were generated by adding two signals. The first signal was generated by exciting an unknown model with a zero-mean stochastic process $g(n)$ and the second signal was the noise $v(n)$. The power of $g(n)$ was fixed to be 1.0 and the power of $v(n)$ were adjusted to obtain $S/N = -5$ dB in all experiments. To get unbiased modeling, the order of the unknown and the adaptive systems were set to be 20. The frequency response of the 20th order system that was used in experiment is shown in Fig. 1. We evaluated the figure of merit ρ using

$$\rho = \frac{ASEMSE_{\mathrm{LMS}}}{ASEMSE_{\mathrm{LMK}}} \tag{32}$$

The $ASEMSE_{\mathrm{LMS}}$ is the ensemble average of the steady state excess mean square error of the standard LMS algorithm [10] and $ASEMSE_{\mathrm{LMK}}$ is evaluated from (32). The value of $ASEMSE_{\mathrm{LMS}}$ and $ASEMSE_{\mathrm{LMK}}$ were obtained from averaging 200 times independent experiments. We set various

$$k_\mu = \frac{\mu_{\mathrm{LMS}}}{\mu_{\mathrm{LMK}}}. \tag{33}$$

$\rho > 1$ means that $ASEMSE_{\mathrm{LMK}} < ASEMSE_{\mathrm{LMS}}$. On the other hand $\rho < 1$ means that $ASEMSE_{\mathrm{LMK}} > ASEMSE_{\mathrm{LMS}}$. Table 1 indicates that when we hope to achieve $ASEMSE_{\mathrm{LMK}} < ASEMSE_{\mathrm{LMS}}$, μ_{LMK} has to be selected lower than that of LMS. Our experiments using various noises indicates that this is true for all kind of noises. LMK has to utilize lower step size than that of LMS, but since the time constant of LMK is smaller than that of LMS, LMK can still be faster than LMS. Furthermore our experiments show that the selection of β is not critical.

$$ASEMSE = \sigma_v^2 + \frac{\mu_{\text{LMK}}\left\{\beta^2\left(9m_6^v(0,1,1,1,1) + 9\beta^2 m_6^v(0,2,2,2,2)\right) + 60\sigma_v^2\right\} tr\{\mathbf{R}_{xx}\}}{12 + 24\sigma_v^2 - \mu_{\text{LMK}}\left(60 + 6\gamma_1\sigma_v^2 + \gamma_2 m_4^v(0,1,1) + \gamma_3 m_4^v(0,2,2)\right) tr\{\mathbf{R}_{xx}\}}, \quad (32)$$

$$E\{\|\mathbf{q}(n+1)\|\} - \mu_{\text{LMK}}^2 tr\{\mathbf{R}_{xx}\}\left(9\beta^2 m_6^v(0,0,1,1,1,1) + 9\beta^4 m_6^v(0,0,2,2,2,2)\right) =$$
$$E\{\|\mathbf{q}(n)\|\} + \mu_{\text{LMK}} E\{e_a^2(n)\}\left\{12 + 24\sigma_v^2 - \mu_{\text{LMK}}\left(60 + 6\sigma_v^2\lambda_1 + \lambda_2 m_4^v(0,0,1,1) + \lambda_3 m_4^4(0,0,2,2)\right) tr\{\mathbf{R_{xx}}\}\right\} \quad (33)$$

$$\mu_{\text{LMK}} \leq \frac{12 + 24\sigma_v^2}{\left(60 + 6\gamma_1\sigma_v^2 + \gamma_2 m_4^v(0,1,1) + \gamma_3 m_4^v(0,2,2)\right) tr\{\mathbf{R_{xx}}\}} \quad (34)$$

5. THE STEP SIZE LIMIT

By substituting (14) and (25) into (20), we will obtain equation (33). Based on (33), we have to limit μ_{LMK} as in (34). This limit is consistent with the pretending (32) to be positive.

6. THE SPEED OF CONVERGENCE

Another important parameter to measure the performance of the adaptive filter is the convergence speed; showing how fast the algorithm adapts when disturbances occur. Assuming that $\mu_{\text{LMK}} << 1$, all terms of (33) containing μ_{LMK} can be neglected. Thus from (20) we can write

$$E\left\{||\mathbf{q}(n+1)||^2\right\} = E\left\{||\mathbf{q}(n)||^2\right\} - E\left\{2\mu_{\text{LMK}}\kappa(e(n))\, e_a(n)\right\} \quad (35)$$

Using (29) we can obtain

$$E\left\{||\mathbf{q}(n+1)||^2\right\} = E\left\{||\mathbf{q}(n)||^2\right\} \left[1 - 2\mu_{\text{LMK}}\left\{6 + 12\sigma_v^2\right\}\mathbf{R_{xx}}\right] \quad (36)$$

Using the method in [1], we can obtain

$$\tau_{\text{LMK}} = \frac{1}{\mu_{\text{LMK}}\left(12 + 24\sigma_v^2\right)\mathbf{R_{xx}}}. \quad (37)$$

The convergence speed figure of merit is evaluate using

$$k_\tau = \frac{\tau_{\text{LMS}}}{\tau_{\text{LMK}}} = 12 + 24\sigma_v^2, \quad (38)$$

where τ_{LMS} is the time constant of the LMS algorithm [1]. The resulted k_τ for various conditions are also included in tabel 1. $k_\tau > 1$ means that the time constant of LMK algorithm $\tau_{\text{LMK}} < \tau_{\text{LMS}}$, meaning that LMK will converges faster than that of LMS.

7. CALCULATION COMPLEXITY

Table 2 shows the required multiplications (mpx) and additions ($addt$) to implement various adaptive algorithms. The table shows that LMK has the heaviest burden among LMS, NLMS and LMK. This fact indicates that more powerful processors have to be utilized for implementing the proposed LMK algorithm than in case of LMS, NLMS. This complexity is compensated by the fact that the performance of the LMK is the best among those three algorithms.

Noise Type	k_μ	ρ_μ	k_τ
Gaussian distributed	1	0.85	18
	2	1.75	18
	8	2.39	18
	12	3.60	18
	20	6.04	18
Flat distributed	1	0.17	14
	2	0.32	14
	8	1.38	14
	12	2.03	14
	20	3.34	14
Non-symmetrical distributed	1	0.17	18
	2	0.34	18
	8	1.38	18
	12	2.06	18
	20	3.45	18

Table 1: The table of k_μ, ρ_μ and k_τ for various adaptive algorithms and various noises. All $S/N = -5dB$.

Method	Formula	Complexity
LMS	$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{x}(n)e(n)$	$2M+2$ mpx $2M$ addt
NLMS	$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu\mathbf{x}(n)e(n)}{X_n}$	$3M+2$ mpx 3M addt
LMK	$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{x}(n)\kappa(e(n))$	$2M+7$ mpx $2M+2$ addt

Table 2: The total multiplications (mpx) and additions ($addt$) for various adaptive algorithms.

8. THE SIMULATION RESULTS

In the following experiment, the unknown model is excited by a zero-mean noise. The order of the unknown system was set to be 20, while the order of the adaptive system was also selected to be 20. We measured $e^2(n)$ as a function of time index n. We compare the result of the LMK and the conventional LMS algorithm [1] methods. The level of the noise $v(n)$ was adjusted so that $S/N = -5$ dB. In the first experiment, we used Gaussian noise. The average progression of $e^2(n)$ that was obtained from 200 independent runs is plotted in Fig. 2. We used the LMK and the LMS algorithms to adapt the coefficient of the filter. We select $k_\mu = 10$. The progression of $e^2(n)$ for both algorithms are plotted in Fig. 2.

Those plots indicate that the LMK algorithm converges faster than that of LMS and also to a lower error as it is expected in section **II**.

In the second experiment, a non Gaussian noise was applied.

14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

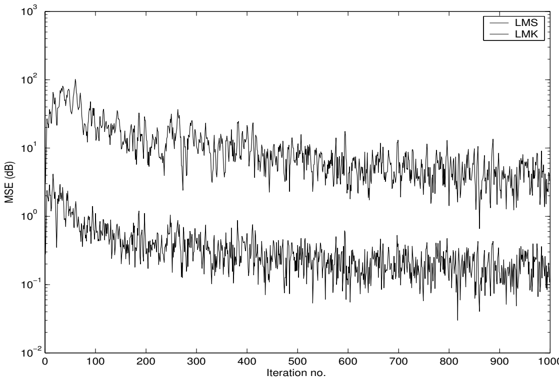

Figure 2: The progression of $e^2(n)$ when the noise is Gaussian using LMS and LMK algorithms.

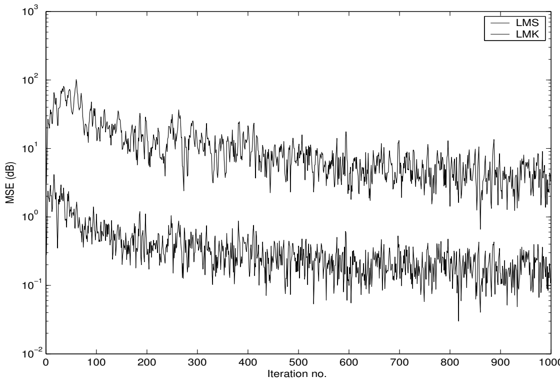

Figure 3: The progression of $e(n)$ for contaminated non-Gaussian symmetrical PDF additive noise.

The non Gaussian noise was generated using the random noise generator with flat probability density function $-1 \leq f(x) \leq 1$. In the experiments we applied $S/N = -5$ dB.

The obtained $e^2(n)$ as in Fig. 3 shows the same behavior as before. We obtain smaller final error by using LMK than that when LMS is used.

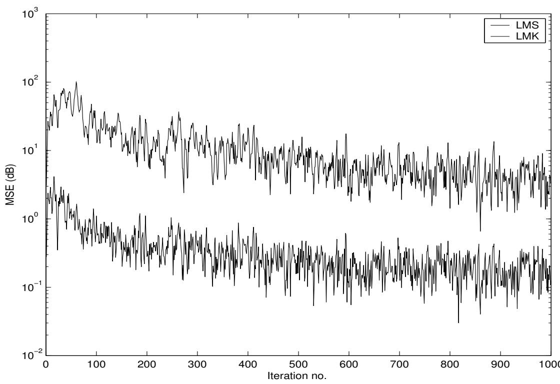

Figure 4: The progression of $e(n)$ for contaminated non-Gaussian non-symmetrical PDF additive noise.

In the third experiment, a non Gaussian noise with non-symmetrical PDF was applied. The non Gaussian noise was generated using the random noise generator with flat probability density function $-1 \leq f(x) \leq 1$. To get non-symmetrical PDF will only take signal that are positive and use them as noise. In the experiments the power of the noise was adjusted to get $S/N = -5$ dB.

The obtained $e^2(n)$ as in Fig. 4 shows the same behavior as before. We obtain smaller final error by using LMK than that when LMS is used.

All our other simulation results show the same behavior for both Gaussian and impulsive additive noise cases. We can always achieve smaller error by using LMK. The LMK algorithm convergences is faster than that when LMS is used. Because of space limitation, those results are not presented here.

In many applications, the statistical behavior of noise is unknown and it could be impulsive. Thus, finally we recommend that LMK is widely used as an alternative adaptive algorithm than the conventional LMS.

9. CONCLUSIONS

A theoretical analysis of the LMK algorithm performance has been presented in this paper. The theoretical and experimental study results indicate that LMK can outperforms the conventional LMS method; i.e for Gaussian and non-Gaussian noises by adjusting μ_{LMK}. Selected the appropriate μ_{LMK} is still studied and will be reported later. Since calculation burden of the LMK and LMS are not much different, LMK can be considered as an alternative algorithm to LMS.

Real applications and also efficient hardware implementation of LMK are still studied and will be reported somewhere else.

REFERENCES

[1] S. Haykin, Adaptive Filter Theory, 2nd Edition, Prentice Hall Inc, 1990.

[2] A. Papoulis, Random Variables and Stochastic Processes, Mc Graw Hill Inc, 1965.

[3] C.L. Nikias and M.R. Raghuveer, "Bispectrum Estimation: A Digital Signal Processing Framework", *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869 − 891, July 1987.

[4] J. A. F. Fonolossa, "Adaptive system Identification Based on High Order Statistics", *Proc. ICASSP 1991*, pp. 3437 − 3440, Toronto, Canada, 1991.

[5] D. I. Pazaitis and A. G. Constantinides, "A novel kurtosis driven variable step-size adaptive algorithm", *IEEE Trans. Signal Processing*, vol. 47, pp. 864 − 872, no. 3, March 1999.

[6] D. I. Pazaitis and A. G. Constantinides, "A novel kurtosis driven variable step-size adaptive algorithm", *IEEE Trans. Signal Processing*, vol. 47, pp. 864 − 872, no. 3, March 1999.

[7] D.I. Pazaitis and A.G. Constantinides, "An Intelligent LMS+F Algorithm", *Proc* 8th *IEEE Signal Processing Workshop, Statistical Signal and Array Processing* , pp. 486 − 489, June 1996.

[8] O. Tanrikulu, A.G. Constantinendes, "Least Mean Kurtosis: A novel higher-order statistics based adaptive filtering algorithm", *Electronics Letters*, vol. 30, no. 3, pp. 189 − 190, February 1984.

[9] P.I. Hubsher and J.C. Bermudez, " A Model for the Behavior of the Least Mean Kurtosis (LMK) Adaptive Algorithm with Gaussian Input, *Proc. International Telecommunication Symposium - ITS 2000,* Natal, Brasil.

[10] Nabil R. Yousef and Ali H Sayed, "A Unified Approach to the Steady State and Tracking Analyses of Adaptive Filters", *IEEE Trans. on Signal Processing*, vol. 49, no. 2, pp. 314 − 324, Feb. 2001.