Novel Adaptive Algorithm Based on Least Mean p-Power Error Criterion for Fourier Analysis in Additive Noise

Yegui Xiao and *Katsunori Shida*
Dept. of Electrical Engineering, Saga University,
Honjo-machi, Saga 840-8502, Japan
Tel: +81 952 288640; Fax: +81 952 288651
e-mail: xiao@ee.saga-u.ac.jp

ABSTRACT

This paper presents a novel adaptive algorithm for the estimation of discrete Fourier coefficients (DFC) of sinusoidal and/or quasi-periodic signals in additive noise. The algorithm is derived using a least mean p-power error criterion. It reduces to the conventional LMS algorithm when p takes on 2. It is revealed by both analytical results and extensive simulations that the new algorithm for $p = 3, 4$ generates much improved DFC estimates in moderate and high SNR environments compared to the LMS algorithm, while both have similar degrees of complexity. Assuming the Gaussian property of the estimation error, the proposed algorithm including the LMS algorithm is analyzed in detail. Elegant dynamic equations and closed form noise misadjustment expressions are derived and clarified.

1 Introduction

Adaptive estimation of nonstationary sinusoidal signals or quasi-periodic signals with arbitrary known frequencies or periods is of essential importance in many diverse engineering fields, such as digital communications, power systems, biomedical engineering, pitch detection in transcription and so forth [1-7]. So far, Kalman filtering based techniques [2,4], recursive least square (RLS)[8], simplified RLS (SRLS)[9], LMS-type algorithms [5-7], and sliding algorithms based on FIR or IIR notch filters [10,11], for examples, have been developed for this purpose.

All the above-mentioned adaptive algorithms have both advantages and drawbacks of their own. It seems that the LMS-type (gradient-based) algorithms are perhaps the most frequently used ones in real-life applications due to their low computational requirements and good performances. Algorithm that posesses similar complexity to, but enjoy better performance than the LMS algorithm will have great appeal in real-life applications. This work is devoted to presenting such a new algorithm.

When the frequencies of the sinusoidal signal are not known in advance, FIR type adaptive line enhancer (ALE) and adaptive IIR notch filters can be used to produce very good estimates of frequencies [12, and references therein]. Of course, there are a lot of cases where the frequencies are given *a prior* [6,11].

This paper delivers two major contributions. First, the new algorithm using the least mean p-power error criterion is developed, which reduces to a new sign algorithm and the conventional LMS algorithm when p takes on 1 and 2, respectively. It should be noted that this work is the first to mention the sign algorithm in the context of adaptive estimation of sinusoidal signals. It is found by both analytical results and extensive simulations that the new algorithm with $p = 3, 4$ works much better than the sign and the LMS algorithms for a very wide range of estimation scenarios. Second, we present an elegant performance analysis for both the new algorithm and the LMS algorithm. The analysis for $p = 1, 2, 3$, and 4 is carried out along the same procedures under the assumption that the estimation error signal is Gaussian distributed. The dynamic properties (convergence in the mean and convergence in the mean square) and closed form noise misadjustment expressions are derived and clarified.

Let the noisy measurement be

$$\begin{aligned} d(n) &= s(n) + v(n) \\ &= \sum_{i=1}^{q} (a_i \cos \omega_i n + b_i \sin \omega_i n) + v(n) \quad (1) \end{aligned}$$

where ω_i is the arbitrary frequency of the i-th component, known in advance or estimated by some adaptive frequency estimator. $v(n)$ is the additive white noise with zero mean and variance σ_v^2. It is required to design an adaptive algorithm that is capable of estimating the DFCs of $s(n)$ contaminated by $v(n)$. Here, the p-power error criterion is used to design the adaptive algorithm.

$$J(n) = E[|e(n)|^p], \quad p \geq 1 \quad (2)$$

$$e(n) = d(n) - \hat{s}(n), \quad (3)$$

$$\hat{s}(n) = \sum_{i=1}^{q} \left(\hat{a}_i \cos \omega_i n + \hat{b}_i \sin \omega_i n \right). \quad (4)$$

Fig.1 demonstrates the adaptive scheme. The new algorithm is given by

$$\hat{\mathbf{A}}_i(n+1) = \hat{\mathbf{A}}_i(n) + \mu_i e^{p-1}(n)\mathbf{X}_i(n) \qquad (5)$$

for even p, and

$$\hat{\mathbf{A}}_i(n+1) = \hat{\mathbf{A}}_i(n) + \mu_i sgn(e(n)) e^{p-1}(n)\mathbf{X}_i(n) \qquad (6)$$

for odd p, where

$$\begin{aligned} \mathbf{X}_i(n) &= [\cos\omega_i n \ \sin\omega_i n]^T, & (7)\\ \hat{\mathbf{A}}_i(n) &= [\hat{a}_i(n)\ \hat{b}_i(n)]^T. & (8) \end{aligned}$$

This paper is organized as follows. In Section 2, we provide an extensive and elegant performance analysis for the new class of algorithms including the conventional LMS algorithm. Some representative simulation results will be given in Section 3 to prove the better performance of the new algorithm and to support the analytical findings. Section 4 gives the conclusions.

2 Performance analysis

In this section, we give the performance analysis of the new algorithm. The analysis is based on the assumption that the error signal is Gaussian distributed along the convergence process. Similar assumption has been successful for the performance analysis of an echo canceler [13]. We first confirm the validity of this assumption.

2.1 Gaussianity of the error signal

The Gaussianity of the estimation error $e(n)$ has been confirmed at many points of the transition and steady-state stages in an ensemble-averaging sense. We omit the simulation results due to the space limitation. This assumption will make the analysis simpler and elegant, especially for odd p.

2.2 Difference equations

Let

$$\begin{aligned} \varepsilon_{a_i}(n) &= a_i - \hat{a}_i(n), & (9)\\ \varepsilon_{b_i}(n) &= b_i - \hat{b}_i(n), & (10)\\ J_{a_i}(n) &= E[(\hat{a}_i(n) - a_i)^2] = E[\varepsilon_{a_i}^2(n)], & (11)\\ J_{b_i}(n) &= E[(\hat{b}_i(n) - b_i)^2] = E[\varepsilon_{b_i}^2(n)]. & (12) \end{aligned}$$

The mean and variance of $e(n)$ are calculated by

$$\mu_{e(n)} = \sum_{i=1}^{q} (E[\varepsilon_{a_i}(n)]\cos\omega_i n + E[\varepsilon_{b_i}(n)]\sin\omega_i n), \quad (13)$$

$$\sigma_{e(n)}^2 = \sum_{i=1}^{q} (\sigma^2_{\varepsilon_{a_i}(n)} \cos^2\omega_i n + \sigma^2_{\varepsilon_{b_i}(n)}\sin^2\omega_i n) + \sigma_v^2 \quad (14)$$

where

$$\sigma^2_{\varepsilon_{a_i}(n)} = E[\varepsilon^2_{a_i}(n)] - E[\varepsilon_{a_i}(n)]^2, \qquad (15)$$

$$\sigma^2_{\varepsilon_{b_i}(n)} = E[\varepsilon^2_{b_i}(n)] - E[\varepsilon_{b_i}(n)]^2. \qquad (16)$$

In the analysis followed, we assume $p = 3$. Through the same derivation process, the analysis for $p = 1, 2, 4, \cdots$ can be performed.

The difference equations for the convergence in the mean are derived as, after some complicated calculations,

$$\begin{aligned} E[\varepsilon_{a_i}(n+1)] &= E[\varepsilon_{a_i}(n)] - 2\mu_i\ I_3(n)\ \cos\omega_i n, & (17)\\ E[\varepsilon_{b_i}(n+1)] &= E[\varepsilon_{b_i}(n)] - 2\mu_i\ I_3(n)\ \sin\omega_i n & (18) \end{aligned}$$

where

$$\begin{aligned} I_3(n) &= \int_{-\infty}^{\infty} sgn(e(n)) e^2(n) p(e(n))\, de(n) \\ &= \sqrt{\frac{2}{\pi}} \mu_{e(n)}\sigma_{e(n)} e^{-\frac{1}{2}\left(\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right)^2} + 2\left(\mu^2_{e(n)} + \sigma^2_{e(n)}\right) \\ &\quad \times sgn\left(\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right) erf\left(\left|\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right|\right), \qquad (19) \end{aligned}$$

$$erf(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \qquad (20)$$

and $p(e(n))$ is the Gaussian PDF of the error signal $e(n)$.

For the convergence in the mean square, we have

$$\begin{aligned} J_{a_i}(n+1) &= J_{a_i}(n) - 2\mu_i E[sgn(e(n))e^2(n)\varepsilon_{a_i}(t)] \\ &\quad \times \cos\omega_i n + \mu_i^2 E[e^4(n)]\cos^2\omega_i n, & (21)\\ J_{b_i}(n+1) &= J_{b_i}(n) - 2\mu_i E[sgn(e(n))e^2(n)\varepsilon_{b_i}(t)] \\ &\quad \times \sin\omega_i n + \mu_i^2 E[e^4(n)]\sin^2\omega_i n. & (22) \end{aligned}$$

Here, we also assume that $e(n)$ and $\varepsilon_{a_i}(n)$, $e(n)$ and $\varepsilon_{b_i}(n)$ are joint-Gaussian distributed. After some complicated algebra, we obtain

$$\begin{aligned} &E[sgn(e(n))e^2(n)\varepsilon_{a_i}(n)] \\ &= A e^{-\frac{1}{2}\left(\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right)^2} + B sgn\left(\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right) erf\left(\left|\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right|\right), \quad (23) \end{aligned}$$

$$A = \sqrt{\frac{2}{\pi}} \sigma_{e(n)} \left(\mu_{e(n)} E[\varepsilon_{a_i}(n)] + \sigma^2_{\varepsilon_{a_i}(n)} \cos\omega_i n\right), \quad (24)$$

$$B = 2\left(\sigma^2_{e(n)} + \mu^2_{e(n)}\right) E[\varepsilon_{a_i}(n)] + 4\mu_{e(n)}\sigma^2_{\varepsilon_{a_i}(n)}\cos\omega_i n, \quad (25)$$

$$\begin{aligned} &E[sgn(e(n))e^2(n)\varepsilon_{b_i}(n)] \\ &= A' e^{-\frac{1}{2}\left(\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right)^2} + B' sgn\left(\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right) erf\left(\left|\frac{\mu_{e(n)}}{\sigma_{e(n)}}\right|\right), \quad (26) \end{aligned}$$

$$A' = \sqrt{\frac{2}{\pi}} \sigma_{e(n)} \left(\mu_{e(n)} E[\varepsilon_{b_i}(n)] + \sigma^2_{\varepsilon_{b_i}(n)} \sin\omega_i n\right), \quad (27)$$

$$B' = 2\left(\sigma^2_{e(n)} + \mu^2_{e(n)}\right) E[\varepsilon_{b_i}(n)] + 4\mu_{e(n)}\sigma^2_{\varepsilon_{b_i}(n)}\sin\omega_i n, \quad (28)$$

$$E[e^4(n)] = 3\sigma^4_{e(n)} + 6\mu^2_{e(n)}\sigma^2_{e(n)} + \mu^4_{e(n)}. \qquad (29)$$

Substitution of (23)-(29) to (21) and (22) yields the difference equations. Note that the difference equations for the convergences in the mean and mean square are connected with each other, and all the dynamics of the algorithm can be numerically obtained from them.

2.3 Noise misadjustment

Here, the steady-state property of the algorithm is considered. At steady-state, it can be found, by close inspections of (17), (18), that the algorithm is unbiased:

$$E[\varepsilon_{a_i}(n)]|_{n\to\infty} = E[\varepsilon_{b_i}(n)]|_{n\to\infty} = 0. \qquad (30)$$

It is also proved by using (21)-(29) that

$$J_{a_i}(n)|_{n\to\infty} = J_{b_i}(n)|_{n\to\infty}, \qquad (31)$$

regardless of the signal frequencies and DFC magnitudes. From the difference equations for the mean square error, one obtains

$$4\sqrt{\frac{2}{\pi}}J_{a_i}(\infty) = 3\mu_i(\sum_{j=1}^{q} J_{a_j}(\infty) + \sigma_v^2)^{\frac{3}{2}}. \qquad (32)$$

Expanding the above equation in Taylor series expansion and keeping the first-order term lead to

$$4\sqrt{\frac{2}{\pi}}J_{a_i}(\infty) - \frac{9}{2}\mu_i\sigma_v\sum_{j=1}^{q} J_{a_j}(\infty) \approx 3\mu_i\sigma_v^3. \qquad (33)$$

Then, for $i = 1, 2, \cdots, q$, the following simultaneous equations are generated.

$$\begin{bmatrix} 4\sqrt{\frac{2}{\pi}} - \frac{9}{2}\mu_1\sigma_v & -\frac{9}{2}\mu_1\sigma_v & \cdots & -\frac{9}{2}\mu_1\sigma_v \\ -\frac{9}{2}\mu_2\sigma_v & 4\sqrt{\frac{2}{\pi}} - \frac{9}{2}\mu_2\sigma_v & \cdots & -\frac{9}{2}\mu_2\sigma_v \\ \vdots & \vdots & \vdots & \vdots \\ -\frac{9}{2}\mu_q\sigma_v & -\frac{9}{2}\mu_q\sigma_v & \cdots & 4\sqrt{\frac{2}{\pi}} - \frac{9}{2}\mu_q\sigma_v \end{bmatrix}$$

$$\times \begin{bmatrix} J_{a_1}(\infty) \\ J_{a_2}(\infty) \\ \vdots \\ J_{a_q}(\infty) \end{bmatrix} = 3\sigma_v^3 \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_q \end{bmatrix}. \qquad (34)$$

Letting $J_{a_i}(\infty) = 3\mu_i\sigma_v^3 x$, we have from the above equations

$$x = \frac{1}{4\sqrt{\frac{2}{\pi}} - \frac{9}{2}\sigma_v\sum_{j=1}^{q}\mu_j}. \qquad (35)$$

Therefore, the noise misadjustment is finally obtained

$$J_{a_i}(\infty) = J_{b_i}(\infty) = \frac{3\mu_i\sigma_v^3}{4\sqrt{\frac{2}{\pi}} - \frac{9}{2}\sigma_v\sum_{j=1}^{q}\mu_j}. \qquad (36)$$

For other algorithms, we have

Sign algorithm ($p = 1$) :

$$J_{a_i}(\infty) = J_{b_i}(\infty) = \frac{\mu_i\sigma_v}{2\sqrt{\frac{2}{\pi}} - \frac{1}{2}\sigma_v^{-1}\sum_{j=1}^{q}\mu_j}; \qquad (37)$$

LMS algorithm ($p = 2$) :

$$J_{a_i}(\infty) = J_{b_i}(\infty) = \frac{\mu_i\sigma_v^2}{2 - \sum_{j=1}^{q}\mu_j}; \qquad (38)$$

Proposed algorithm ($p = 4$) :

$$J_{a_i}(\infty) = J_{b_i}(\infty) = \frac{5\mu_i\sigma_v^4}{2 - 10\sigma_v^2\sum_{j=1}^{q}\mu_j}. \qquad (39)$$

From these elegant explicit expressions, we have the following remarks in order:

- The numerators of the noise misadjustments derived above are proportional to the p-power of the noise standard deviation σ_v. Their denominators are negatively proportional to the (p-2)-power of the noise standard deviation σ_v.
- For small noise environment, the algorithms with larger p will produce much smaller noise misadjustments. However, for larger additive noise ($\sigma_v >> 1$), they will have large noise misadjustments, which is undesirable.
- Loose stability bounds can be derived from the above noise misadjustment expressions by letting them positive. For example, for $p = 4$, we have

$$\mu_i > 0, \quad 0 < \sum_{j=1}^{q}\mu_j < \frac{1}{5\sigma_v^2}. \qquad (40)$$

3 Simulation results

To compare the four algorithms ($p = 1, 2, 3, 4$), we let the steady-state MSEs of the LMS algorithm be the base. The MSEs of the other three algorithms are adjusted to have the same values as the base by using different step size parameters of their own. The analytical dynamics of all the algorithms can then be easily obtained by solving the established difference equations numerically. Simulations are also carried out for these step size parameters. For concise notation, here we let the MSEs of the LMS algorithm be $J_{a_i}^{lms}(\infty)$, and $\mu_{i,sgn}, \mu_{i,lms}, \mu_{i,p3}, \mu_{i,p4}$ indicate the step size parameters of the sign algorithm, the LMS, and the algorithms with $p = 3$ and 4, respectively.

Letting the i-th MSE of the algorithm with $p = 3$ be equal to that of the LMS yields

$$3\mu_{i,p3}\sigma_v^3 = J_{a_i}^{lms}(\infty)\left(4\sqrt{\frac{2}{\pi}} - \frac{9}{2}\sigma_v\sum_{j=1}^{q}\mu_{j,p3}\right) \qquad (41)$$

Summing the above equation leads to

$$\sum_{j=1}^{q} \mu_{j,p3} = \frac{4\sqrt{\frac{2}{\pi}} \sum_{i=1}^{q} J_{a_i}^{lms}(\infty)}{3\sigma_v^3 + \frac{9}{2}\sigma_v \sum_{i=1}^{q} J_{a_i}^{lms}(\infty)}. \tag{42}$$

Then, each step size parameter $\mu_{i,p3}$ can be readily obtained by substituting (42) to (41). Step size parameters for the sign algorithm and algorithm with $p = 4$ can be readily obtained in the same way.

Here, we show a typical result selected from a vast number of simulations conducted to confirm the superiority of the new algorithm over the LMS. Fig.2 demonstrates the analytical dynamics and their simulated results of the four algorithms for a moderate noise situation. It is obvious that the new algorithm shows faster convergence rate than the LMS, and the theory fits the simulation excellently.

A representative comparison between algorithms with p=2, 3, 4, 5, and 6 is given in Fig.3. It is obvious that the new algorithm with $p = 3, 4$ presents much improved performance than the LMS, and the algorithms with $p = 5$ and 6 work poorer than the new algorithm. That is why we have not paid attention to the algorithms with $p \geq 5$.

4 Conclusions

To conclude, the proposed p-power (p=3,4) adaptive algorithm works much better than the sign and LMS algorithms. Its dynamic behavior and steady-state properties have been extensively investigated. The analytical results show excellent fits to their simulated values.

References

[1] International Conference on Harmonics in Power Systems, University of Manchester Inst. of Science and Technology (Power Systems Eng. Series), 1981, Manchester.

[2] R. R. Bitmead, A. C. Tsoi and P. J. Parker, "A Kalman filtering approach to short-term Fourier analysis," IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-34, No.6, pp.1493-1501(Dec., 1986).

[3] S. Osowski, "Neural network for estimation of harmonic components in a power system," IEE PROCEEDINGS-C, Vol.139, No.2, pp.129-135(Mar., 1992).

[4] P. Gruber and J. Todtli, "Estimation of quasiperiodic signal parameters by means of dynamic signal models," IEEE Trans. Signal Processing, Vol.42, No.3, pp.552-562(Mar., 1994).

[5] C. A. Vaz and N. V. Thakor, "Adaptive Fourier estimation of time-varying evoked potentials," IEEE Trans. Biomed. Eng., Vol.36, pp.448-455(1989).

[6] Y. Xiao and Y. Tadokoro, "LMS-based notch filter for the estimation of sinusoidal signals in noise," Signal Processing, Vol.46, No.2, pp.223-231(1995).

[7] Y. Xiao, Y. Tadokoro, and K. Iwamoto, "Real-valued LMS Fourier analyser for sinusoidal signals in noise," Signal Processing, (revised version submitted, 1997).

[8] N. K. M'sirdi et al., "An RML algorithm for retrieval of sinusoids with cascaded notch filters," ICASSP-1988, pp.2484-2487(1988).

[9] S. Pei and C. Tseng, "Real time cascade adaptive notch filter scheme for sinusoidal parameter estimation," Signal Processing, Vol.39, pp.117-130(1994).

[10] Y. Tadokoro and K. Abe, "Notch Fourier transform," IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-35, No.9, pp.1282-1288(Setp., 1987).

[11] M. T. Kilani and J. F. Chicharo, "A constrained notch Fourier transform," IEEE Trans. Signal Processing, Vol.43, No.9, pp.2058-2067(Setp., 1995).

[12] Y. Xiao, Y. Tadokoro, and Y. Kobayashi, "A new memoryless nonlinear gradient algorithm for a second-order adaptive notch filter and its performance analysis," IEEE Trans. Circuits, Syst., Vol.45, No.4 (regular paper in press, 1998).

[13] K. Koike, "Convergence analysis of a data echo canceller with a stochastic gradient adaptive FIR filter using the sign algorithm," IEEE Trans. Signal Processing, Vol.43, No.12, pp.2852-2861(Dec. 1995).

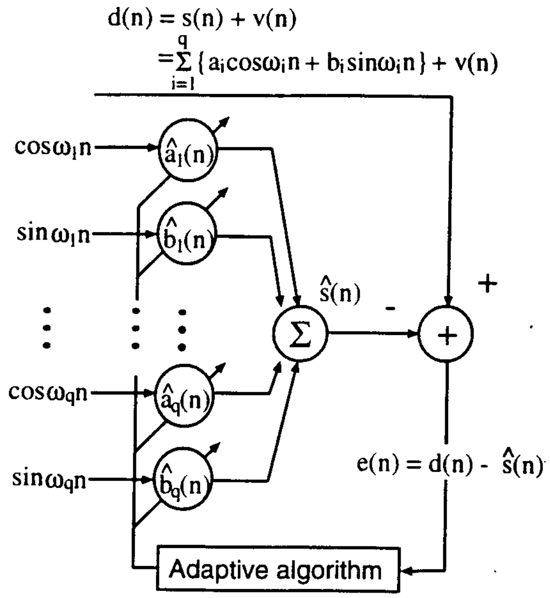

Figure 1: Block diagram of the scheme.

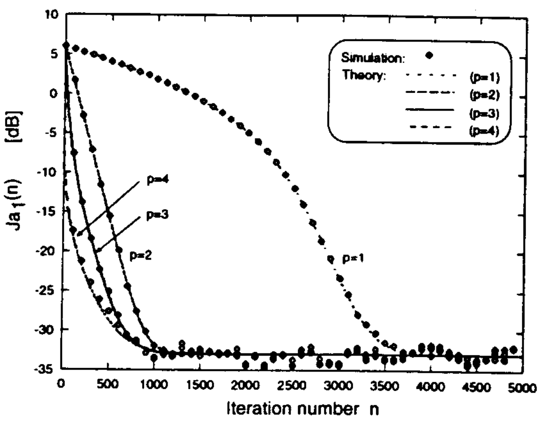

Figure 2: Comparisons between the theoretical difference equations and their simulated values of the four algorithms ($\sigma_v^2 = 0.1, q = 3, 100$ runs).

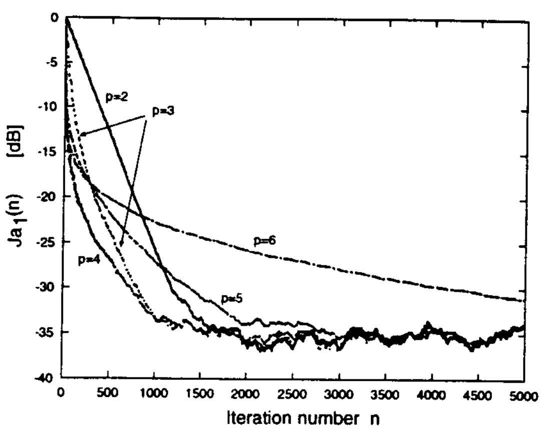

Figure 3: Comparisons among five algorithms with $p = 2, 3, 4, 5, 6$ ($\sigma_v^2 = 0.1, q = 3, 100$ runs).