ADAPTIVE STEP SIZE SIGN ALGORITHM FOR APPLICATION TO ADAPTIVE FILTERING IN DIGITAL QAM COMMUNICATIONS

Shin'ichi Koike

NEC Corporation
7-1, Shiba 5-chome, Minato-ku, Tokyo 108-8001 Japan

ABSTRACT

A new Adaptive Step Size control algorithm is proposed to be combined with the Sign Algorithm for use in complex-valued adaptive filters for application to QAM communications. The algorithm, ASSSA, is fully analyzed to yield a set of difference equations for calculating the transient behavior, hence the steady-state performance, of the filter convergence in terms of excess mean squared error (EMSE). An approximation method for multilevel QAM is further proposed to reduce the amount of computation. The results of experiment with some examples verify the effectiveness of the proposed ASSSA in significantly improving the convergence rate, and also show that the theoretical convergence is in good agreement with that of simulations, which validates the analysis.

1. INTRODUCTION

Quadrature Amplitude Modulation (QAM) is a basic modulation scheme most widely adopted in wired and wireless digital communications.

In the wired communications, the QAM is applied to voiceband modems, xDSL (a family of high speed digital subscriber line transmission systems) modems, cable modems for CATV networks, *etc*. In the wireless world, terrestrial and satellite digital radio communications, terrestrial digital TV broadcasting, mobile communications including the next generation CDMA, *etc*., are all based on the QAM technology.

Adaptive filtering constitutes the core technology in Digital Signal Processing, and plays a key role in the QAM communications as well. In fact, adaptive filters are used for echo cancellation, channel distortion equalization, antenna array control, *etc*. In QAM applications, the adaptive filters must be 2D, or *complex-valued*, because the baseband signal of the QAM is expressed as a set of in-phase and quadrature components.

In practical adaptive filtering systems, the Least Mean Square Algorithm (LMSA) is widely used as the tap weight adaptation algorithm. However, the Sign Algorithm (SA), that is derived from the LMSA, is also attractive for its simplicity in implementation, robustness against disturbances and assured convergence [1]-[3]. While the SA outperforms the LMSA in impulse noise environment, the drawback of using the SA in certain applications appears to be its fairly slow convergence rate as compared with the LMSA [4]. Some methods for improving the convergence rate of the SA using an adaptively controlled step size for tap weight adaptation have been proposed [5] [6].

Based upon the above observations, this paper proposes and analyzes a new Adaptive Step Size Sign Algorithm (ASSSA) for use in complex-valued adaptive filters, particularly focusing on application to digital QAM communications.

2. ASSSA FOR COMPLEX-VALUED ADAPTIVE FILTERS

2.1 Tap Weight Update Equation for the SA

Fig.1 depicts the basic structure of an FIR adaptive filter for identification of an unknown system, where

$\boldsymbol{a}(n) = \boldsymbol{a}^R(n) + j\boldsymbol{a}^I(n)$ complex-valued reference input vector (length N),

$\boldsymbol{c}(n) = \boldsymbol{c}^R(n) + j\boldsymbol{c}^I(n)$ complex-valued tap weight vector (N taps),

$$d(n) = d^R(n) + jd^I(n) = \boldsymbol{h}(n)^H \boldsymbol{a}(n) + \nu(n) \quad (1)$$

desired signal (complex),

$\boldsymbol{h}(n) = \boldsymbol{h}^R(n) + j\boldsymbol{h}^I(n)$ unknown system response vector (complex),

$$e(n) = e^R(n) + je^I(n) = d(n) - \boldsymbol{c}(n)^H \boldsymbol{a}(n) \quad (2)$$

error signal (complex),

$\nu(n) = \nu^R(n) + j\nu^I(n)$ additive noise (complex),

n time instant, N number of taps, and $j = \sqrt{-1}$,

where $(\bullet)^R$ and $(\bullet)^I$ indicate real and imaginary part of a complex number, respectively, and $(\bullet)^H$ denotes *transpose with complex conjugation* or *Hermitian*.

The tap weight update equation for the complex-valued Sign Algorithm is given by

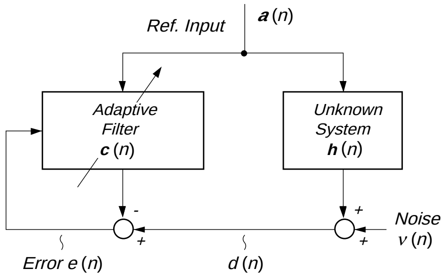

Fig.1 Structure of adaptive filter for system identification.

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) + \alpha_c\, sgn\{\, e(n)\}^* \boldsymbol{a}(n), \tag{3}$$

in which α_c is *real-valued* step size for tap weight adaptation, $sgn(\bullet)$ is *signum* function, $sgn(x + jy) = sgn(x) + j\, sgn(y)$, and $(\bullet)^*$ denotes *complex conjugate*.

2.2 Assumptions

Prior to the analysis to be developed in the subsequent sections, the following assumptions are made.

<u>*Assumptions*</u>

(A) The reference inputs $\boldsymbol{a}^R(n)$ and $\boldsymbol{a}^I(n)$ are stationary *independent and identical digital data* processes, each having zero mean, covariance $\mathbf{R_a} = E[\boldsymbol{a}^R(n)\, \boldsymbol{a}^R(n)^T] = E[\boldsymbol{a}^I(n)\, \boldsymbol{a}^I(n)^T]$ and variance σ_a^2.

(B) The additive noise components $v^R(n)$ and $v^I(n)$ are stationary zero mean *i. & i. Gaussian* distributed processes, each having variance σ_v^2.

(C) The reference input $\boldsymbol{a}(n)$ and the tap weight $\boldsymbol{c}(n)$ are mutually independent (*Independence Assumption*).

In *Assumption* (A), $a^R(n)$ or $a^I(n)$ takes on finite number of *discrete* values, and the value $tr(\mathbf{R_a}^2)$ is known. *Assumption* (C) is adopted to simplify the analysis as is done in many papers.

2.3 Transient Analysis of the SA

Assuming the unknown system to be time-invariant, let "tap weight error" vector be defined by $\boldsymbol{\theta}(n) = \boldsymbol{h} - \boldsymbol{c}(n) = \boldsymbol{\theta}^R(n) + j\boldsymbol{\theta}^I(n)$, and its mean and covariance by

$$\boldsymbol{m}(n) = E[\boldsymbol{\theta}(n)] = \boldsymbol{m}^R(n) + j\, \boldsymbol{m}^I(n) \tag{4}$$

and

$$\mathbf{R}(n) = E[\,(\boldsymbol{\theta}(n) - \boldsymbol{m}(n))\,(\boldsymbol{\theta}(n) - \boldsymbol{m}(n))^H] = \mathbf{R}^{RR}(n) + \mathbf{R}^{H}(n) + j\{-\mathbf{R}^{RI}(n) + \mathbf{R}^{IR}(n)\}, \tag{5}$$

where $\boldsymbol{m}^R(n) = E[\boldsymbol{\theta}^R(n)]$, $\mathbf{R}^{RI}(n) = E[(\boldsymbol{\theta}^R(n) - \boldsymbol{m}^R(n))(\boldsymbol{\theta}^I(n) - \boldsymbol{m}^I(n))^T]$, *etc.*

The excess mean squared error (EMSE) per channel is calculated as

$$\varepsilon(n) = E[\,|\,\boldsymbol{\theta}(n)^H \boldsymbol{a}(n)\,|^2\,]/2 = tr\{\mathbf{R}a\,(\boldsymbol{m}^R(n)\, \boldsymbol{m}^R(n)^T + \boldsymbol{m}^I(n)\, \boldsymbol{m}^I(n)^T + \mathbf{R}^{RR}(n) + \mathbf{R}^{II}(n))\}, \tag{6}$$

where $|\bullet|^2$ is *squared norm* of a complex number and $tr(\bullet)$ denotes *trace* of a matrix.

The author proposed an efficient method of calculating the convergence of an adaptive filter with digital data input $\boldsymbol{a}(n)$ taking on discrete values, based on the assumption that the conditional probability distribution of $e(n)$ given the reference input at the kth tap, $a(n-k)$, is approximately Gaussian [7]. Applying this method yields the following difference equations for $\boldsymbol{m}^R(n)$, $\boldsymbol{m}^I(n)$, $\mathbf{R}^{RR}(n)$ and $\mathbf{R}^{II}(n)$.

$$\boldsymbol{m}^R(n+1) = \boldsymbol{m}^R(n) - 2\,\alpha_c\, \boldsymbol{p}^R(n), \tag{7}$$

$$\boldsymbol{m}^I(n+1) = \boldsymbol{m}^I(n) - 2\,\alpha_c\, \boldsymbol{p}^I(n), \tag{8}$$

$$\mathbf{R}^{RR}(n+1) = \mathbf{R}^{RR}(n) - 2\,\alpha_c\,(\mathbf{W}^R(n)\,\mathbf{R}^{RR}(n) + \mathbf{R}^{RR}(n)\mathbf{W}^R(n)) + 2\,\alpha_c^2\,\mathbf{R_a} - 4\,\alpha_c^2\, \boldsymbol{p}^R(n)\, \boldsymbol{p}^R(n)^T \tag{9}$$

and

$$\mathbf{R}^{II}(n+1) = \mathbf{R}^{II}(n) - 2\,\alpha_c\,(\mathbf{W}^I(n)\,\mathbf{R}^{II} + \mathbf{R}^{II}(n)\,\mathbf{W}^I(n)) + 2\,\alpha_c^2\,\mathbf{R_a} - 4\,\alpha_c^2\, \boldsymbol{p}^I(n)\, \boldsymbol{p}^I(n)^T, \tag{10}$$

where we calculate, for the kth element of $\boldsymbol{p}^R(n)$ and $\boldsymbol{p}^I(n)$ and for the (k,κ) th element of $\mathbf{W}^R(n)$ and $\mathbf{W}^I(n)$,

$$p^R{}_k(n) = 2E_a{}^R{}_{(n-k)}\,[a^R(n-k)\, erf(a^R(n-k)\,\rho^R{}_k(n))], \tag{11}$$

$$p^I{}_k(n) = 2E_a{}^I{}_{(n-k)}\,[a^I(n-k)\, erf(a^I(n-k)\,\rho^I{}_k(n))], \tag{12}$$

$$\mathrm{W}^R{}_{k\kappa}(n) = 2E_a{}^R{}_{(n-k)}\,[a^R(n-k)^2\, p_N(a^R(n-k)\,\rho^R{}_k(n))] \times \mathrm{R}_{a\,k\kappa}/(\sigma_a^2\,\sigma^R{}_k(n)) \tag{13}$$

and

$$\mathrm{W}^I{}_{k\kappa}(n) = 2E_a{}^I{}_{(n-k)}\,[a^I(n-k)^2\, p_N(a^I(n-k)\,\rho^I{}_k(n))] \times \mathrm{R}_{a\,k\kappa}/(\sigma_a^2\,\sigma^I{}_k(n)), \tag{14}$$

with

$$\rho^R{}_k(n) = \mu^R{}_k(n)/(\sigma_a^2\,\sigma^R{}_k(n)),$$

$$\rho^I{}_k(n) = \mu^I{}_k(n)/(\sigma_a^2\,\sigma^I{}_k(n)),$$

$$\boldsymbol{\mu}^R(n) = \mathbf{R_a}\, \boldsymbol{m}^R(n) = [\,\ldots,\, \mu^R{}_k(n),\, \ldots\,]^T,$$

$$\boldsymbol{\mu}^I(n) = \mathbf{R_a}\, \boldsymbol{m}^I(n) = [\,\ldots,\, \mu^I{}_k(n),\, \ldots\,]^T,$$

$$\sigma^R{}_k(n)^2 = \varepsilon(n) - \mu^R{}_k(n)^2/\sigma_a^2 + \sigma_v^2, \text{ and}$$

$$\sigma^I{}_k(n)^2 = \varepsilon(n) - \mu^I{}_k(n)^2/\sigma_a^2 + \sigma_v^2.$$

Here, $E_a{}^X{}_{(n-k)}[\bullet]$ means averaging over all possible values of $a^X(n-k)$, $\mathrm{R}_{\mathbf{a}\,k\kappa}$ is the (k,κ) th element of $\mathbf{R_a}$, $erf(x) = \int_0^x p_N(t)\, dt$ (*Error Function*) and $p_N(x) = exp(-x^2/2)/\sqrt{2\pi}$ (*Normal Density*).

If (6) is combined with (7) through (10), theoretically expected transient behavior of the EMSE can be calculated recurrently.

2.4 Approximation Method for Multilevel QAM

For multilevel QAM, such as 64 QAM, 256 QAM, *etc*., the amplitude distribution of the reference input signal $a^R(n-k)$ or $a^I(n-k)$ can be approximated by *Uniform* distribution $U(0, \sigma_a^2)$ whose *pdf* is given by

$$p(a) = \begin{cases} 1/(2\sqrt{3}\,\sigma_a) & \text{for } |a| < \sqrt{3}\,\sigma_a \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

If we use this Uniform *pdf* in the expectation calculation $E_a{}^X{}_{(n-k)}[\bullet]$ in (11) through (14), we obtain

$$E_a{}^R{}_{(n-k)}\,[a^R(n-k)\, erf(a^R(n-k)\,\rho^R{}_k(n))] = \sqrt{3}\,\sigma_a\, Ferf(\sqrt{3}\,\sigma_a\,\rho^R{}_k(n)), \tag{16}$$

$$E_a{}^I{}_{(n-k)}\,[a^I(n-k)^2\, p_N(a^I(n-k)\,\rho^I{}_k(n))] = 3\,\sigma_a^2\, Gp_N(\sqrt{3}\,\sigma_a\,\rho^I{}_k(n)), \tag{17}$$

and *so on*, where we define functions

$$Ferf(x) = \{(x^2-1)\, erf(x) + x\, p_N(x)\}/(2x^2) \tag{18}$$

and

$$Gp_N(x) = \{erf(x) - x\, p_N(x)\}/x^3. \tag{19}$$

Use of (16), (17), *etc*. significantly reduces the amount of computation required for the averaging with respect to a large number of QAM levels of $a^X(n-k)$.

2.5 Adaptive Step Size Sign Algorithm (ASSSA)

The author proposed a new adaptive step size control algorithm for realizing fast convergent adaptive filters in which the theoretically optimum step size is approximated using leaky accumulators [6]. If the adaptive step size control algorithm above is applied to the complex-valued Sign Algorithm, the tap weights and the step size are to be adapted through the following set of equations.

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) + \alpha_c(n)\, sgn\{e(n)\}^* \boldsymbol{a}(n), \qquad (20)$$

$$\alpha_c(n) = Re\{\boldsymbol{q}_0(n)^H \boldsymbol{q}(n)\} / (8\, tr(\mathbf{R_a}^2)) \qquad (21)$$

$$\boldsymbol{q}_0(n+1) = (1-\rho)\, \boldsymbol{q}_0(n) + \rho\, e(n)^* \boldsymbol{a}(n), \qquad (22)$$

$$\boldsymbol{q}(n+1) = (1-\rho)\, \boldsymbol{q}(n) + \rho\, sgn\{e(n)\}^* \boldsymbol{a}(n), \qquad (23)$$

where $\boldsymbol{q}_0(n)$ and $\boldsymbol{q}(n)$ are vectors of length N, $Re\{\bullet\}$ means taking the real part of a complex number, ρ is leakage factor.

The total number of *Real-valued Multiplications* required for calculating (20) through (23) is found to be $7N+1$.

For the adaptive step size in (20), the expectation $E[\alpha_c(n)]$ and $E[\alpha_c(n)^2]$ can be recurrently calculated with a set of difference equations for $E[\boldsymbol{q}_0(n)]$, $E[\boldsymbol{q}(n)]$, *etc*., which are derived from (22) and (23). Due to space limitation, only a few of them are given below. Here, it is assumed that the step size is independent of the tap weights and the input data in (20) [6].

$$\text{(a)}\ E[\boldsymbol{q}_0(n+1)] = (1-\rho)\, E[\boldsymbol{q}_0(n)] + 2\rho\, \boldsymbol{\mu}(n), \qquad (24)$$

$$\text{(b)}\ E[\boldsymbol{q}(n+1)] = (1-\rho)\, E[\boldsymbol{q}(n)] + 2\rho\, \boldsymbol{p}(n), \qquad (25)$$

$$\begin{aligned}\text{(c)}\ E[\boldsymbol{q}_0(n+1)^H \boldsymbol{q}(n+1)] &= (1-\rho)^2 E[\boldsymbol{q}_0(n)^H \boldsymbol{q}(n)] \\ &+ 2(1-\rho)\rho\,(\boldsymbol{\mu}(n)^H E[\boldsymbol{q}(n)] + E[\boldsymbol{q}_0(n)]^H \boldsymbol{p}(n)) \\ &+ \rho^2\, tr(\mathbf{S}(n)), \qquad (26)\end{aligned}$$

where

$$\boldsymbol{\mu}(n) = \mathbf{R_a}\, \boldsymbol{m}(n) = \boldsymbol{\mu}^R(n) + j\, \boldsymbol{\mu}^I(n),$$

$$\boldsymbol{p}(n) = \boldsymbol{p}^R(n) + j\, \boldsymbol{p}^I(n) \quad \text{(see (11) \& (12))},$$

and

$$\begin{aligned}\mathbf{S}(n) &= E[sgn\{e(n)\}^* e(n)\, \boldsymbol{a}(n)\, \boldsymbol{a}(n)^H] \\ &\cong 4\sqrt{2/\pi}\sqrt{\varepsilon(n) + \sigma_v^2}\, \mathbf{R_a}.\end{aligned}$$

Assuming the filter convergence as $n \to \infty$, the steady-state EMSE for the ASSSA is solved to be

$$\varepsilon(\infty) = \delta / (1-\delta) \times \sigma_v^2 \qquad (27)$$

with

$$\begin{aligned}\delta &= (\rho/4)\, N \sigma_a^4\, N / tr(\mathbf{R_a}^2) \\ &\times \{1 + (1 + \pi/2) / (2N) \times tr(\mathbf{R_a}^2) / \sigma_a^4 N\}. \qquad (28)\end{aligned}$$

3. EXPERIMENT

Simulations and theoretical calculations are performed for the following three examples with different filter parameters, where filter convergence with the proposed ASSSA is compared to that with a fixed step size (FSSSA). In the experiment, the simulation result is given as an ensemble average of the squared error over *1000* independent runs of the filter convergence.

Example #1 *16QAM*

$N = 4$, $\sigma_a^2 = 1$ (0 dB), $\sigma_v^2 = .01$ (– 20 dB)

FSSSA : $\alpha_c = 2^{-12}$ $\quad \to \varepsilon(\infty) = -39$ dB

ASSSA : $\rho = 2^{-7}$ $\quad \to \varepsilon(\infty) = -41$ dB

Example #2 *64QAM*

$N = 32$, $\sigma_a^2 = 1$ (0 dB), $\sigma_v^2 = 1$ (0 dB)

FSSSA : $\alpha_c = 2^{-12}$ $\quad \to \varepsilon(\infty) = -20$ dB

ASSSA : $\rho = 2^{-10}$ $\quad \to \varepsilon(\infty) = -21$ dB

Example #3 *AMI-QAM*

$N = 8$, $\sigma_a^2 = 1$ (0 dB), $\sigma_v^2 = .1$ (– 10 dB)

FSSSA : $\alpha_c = 2^{-12}$ $\quad \to \varepsilon(\infty) = -31$ dB

ASSSA : $\rho = 2^{-8}$ $\quad \to \varepsilon(\infty) = -32$ dB

Fig.2 shows the results of the experiment with *Example #1*, where the digital data signal is 16QAM (4 × 4), *SNR* (= σ_a^2 / σ_v^2) is high (*20 dB*) and the EMSE is about – *40 dB*. It is found that the proposed ASSSA significantly improves the convergence rate of the FSSSA in the transient phase. The theoretically calculated convergence curves for both FSSSA and ASSSA agree with those of the simulations with sufficient accuracy, validating the analysis.

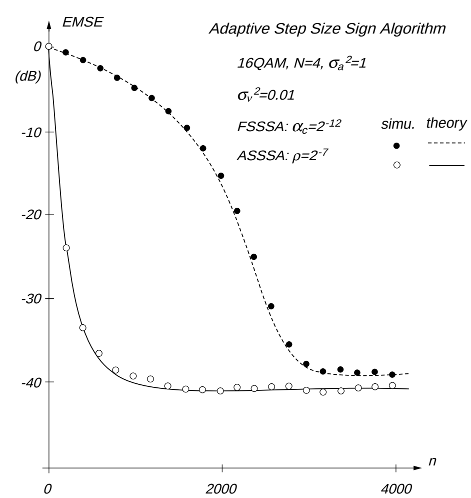

Fig.2 Convergence of adaptive filter – ASSSA versus FSSSA / simulation versus theory
(*Example #1* ; *16QAM, N = 4, SNR = – 20 dB*).

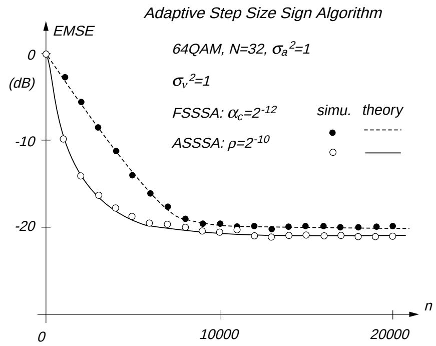

Fig.3 Convergence of adaptive filter – ASSSA versus FSSSA / simulation versus theory (*Example #2* ; *64QAM, N = 32, SNR = 0 dB*).

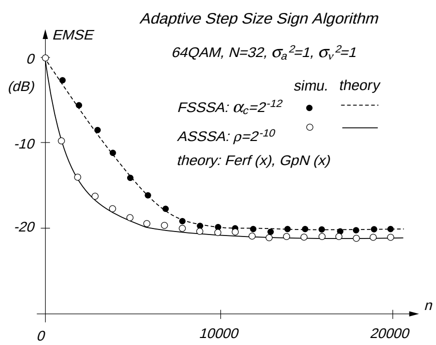

Fig.4 Convergence of adaptive filter – ASSSA versus FSSSA / simulation versus theory (functions *Ferf* (x) & Gp_N (x) used) (*Example #2* ; *64QAM, N = 32, SNR = 0 dB*).

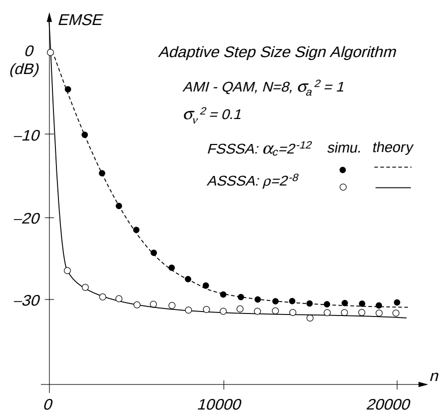

Fig.5 Convergence of adaptive filter – ASSSA versus FSSSA / simulation versus theory (*Example #3* ; *AMI-QAM, N = 8, SNR = – 10 dB*).

In Fig.3 the results for *Example #2* are shown. Again we observe that the ASSSA makes the filter convergence considerably faster.

Fig.4 depicts the theoretical convergence compared with that of the simulation for *Example #2*, in which the approximation using the functions *Ferf* (x) and Gp_N (x) as described in 2.4 is applied to 64QAM. Again we find a good match between the theoretical and empirical convergence curves, showing the validity of the approximation method proposed.

Finally, Fig.5 shows the results for *Example #3*, where the digital data signal is a Quadrature Amplitude Modulated Alternate Mark Inversion code (AMI-QAM). Even for a correlated input, the filter convergence is highly accelerated with the ASSSA.

4. CONCLUSION

A new Adaptive Step Size Sign Algorithm (ASSSA) has been proposed for use in complex-valued adaptive filters to be applied to digital QAM communications.

The results of the experiment with some examples show that the proposed ASSSA is highly effective in improving the convergence rate and that the theoretically calculated convergence and the simulated one exhibit good agreement with sufficient accuracy, validating the analysis.

Further study is required for implementation of the proposed ASSSA in specific applications to QAM systems such as Decision Feedback Equalizer.

5. REFERENCES

[1] D.L.Duttweiler, "Adaptive Filter Performance with Nonlinearities in the Correlation Multipliers," *IEEE Trans. on ASSP*, vol.30, no.4, pp.578-586, August 1982.

[2] W.A.Sethares, "Adaptive Algorithms with Nonlinear Data and Error Functions," *IEEE Trans. on SP*, vol.40, no.9, pp.2199-2206, September 1986.

[3] V.J.Mathews and S.H.Cho, "Improved Convergence Analysis of Stochastic Gradient Adaptive Filters Using the Sign Algorithm," *IEEE Trans. on ASSP*, vol.35, no.4, pp.450-454, April 1987.

[4] S.Koike, "Adaptive Threshold Nonlinear Algorithm for Adaptive Filters with Robustness against Impulse Noise," *IEEE Trans. on SP*, *Correspondence*, vol.45, no.9, pp.2391-2395, September 1997.

[5] A.Kanemasa and K.Niwa, "An Adaptive-Step Sign Algorithm for Fast Convergence of a Data Echo Canceller," *IEEE Trans. on Commun.*, vol.35, no.10, pp.1102-1108, October 1987.

[6] S.Koike, "A Novel Adaptive Step Size Control Algorithm for Adaptive Filters," *Proceedings IEEE ICASSP99*, vol.4, pp.1845-1848, Phoenix, AZ, March 1999.

[7] S.Koike, "A New Efficient Method of Convergence Calculation for Adaptive Filters Using the Sign Algorithm with Digital Data Inputs," *Proceedings IEEE ICASSP97*, vol.III, pp.2333-2336, Munich, Germany, April 1997.