Least Mean Square Nonlinear Regressor Algorithm

Shin'ichi Koike
Consultant
Tokyo, Japan
shin-koike@mub.biglobe.ne.jp

***Abstract*—This paper proposes a new adaptation algorithm named Least Mean Square Nonlinear Regressor Algorithm (LMS-NRA) that makes adaptive filters highly robust against impulse noise at the filter input for which a stochastic model is presented. The proposed algorithm uses a simple nonlinear function of the regressor. A Statistical analysis of the LMS-NRA is developed to calculate theoretical filter convergence. Through numerical experiments, we demonstrate that the proposed algorithm is effective in realizing a robust adaptive filter which is convergent as fast as with the LMS algorithm. Good agreement between simulated and theoretical filter convergence curves shows the validity and accuracy of the analysis.**

***Index Terms*—adaptive filter, LMS algorithm, impulse noise, nonlinear regressor, robust filtering**

I. INTRODUCTION

Impulse noise is often observed in adaptive filters used in practical communication systems. It is known that impulse noise present at the adaptive filter input, as is found in "active noise canceller," may significantly degrade the performance of the adaptive filter using the Least Mean Square Algorithm (LMSA) [1], resulting in a large steady-state error or possible instability, or even filter divergence [2].

To make adaptive filters robust against the impulse noise at the filter input, we introduce a normalizing factor as in the NLMS algorithm [3]. Use of a "signed regressor" in the real-number domain [4] or a "regressor phase" in the complex-number domain [5] is also effective, but it makes adaptive filter convergence much slower.

In order to preserve the convergence speed of the LMSA while protecting the adaptive filter from the impulse noise, we propose, with reference to [6], a simple nonlinear function of the regressor. This proposed algorithm is named Least Mean Square Nonlinear Regressor Algorithm (LMS-NRA) which will be proven highly effective.

For a new adaptation algorithm, it is required to develop a theoretical analysis of the algorithm to verify simulation results. Although the LMS-NRA proposed above is very simple, we find that the theoretical analysis of the algorithm is quite complicated, because impulse noise appears stochastically at each tap of the adaptive filter and the nonlinearity is introduced in the regressor. Thus, the main goal of this paper is to develop a rigorous analysis of the LMS-NRA based on reasonable assumptions. Through numerical experiments, we demonstrate the effectiveness of the proposed algorithm in enhancing the performance of adaptive filters in the presence of impulse noise at the filter input.

The remaining part of the paper is organized as follows. Section II presents a stochastic model for the impulse noise at the filter input. In Section III the LMS-NRA is proposed, and in Section IV we develop its statistical analysis. Section V provides the results of numerical experiments. Finally, Section VI gives some concluding remarks.

II. IMPULSE NOISE AT FILTER INPUT

In this section, we present a stochastic model for impulse noise which is found at the filter input [7].

A "noisy" filter input $b(n)$ with impulse noise $\nu_a(n)$ added to the regressor $a(n)$ is expressed as

$$b(n) = a(n) + \tau(n)\, \nu_a(n), \tag{1}$$

where n is the time instant and $\tau(n)$ is an independent Bernoulli random variable defined by

$$\tau(n) = \begin{cases} 1 & \text{with probability of occurrence } p_{\nu a} \\ 0 & \text{with probability of occurrence } 1 - p_{\nu a}. \end{cases} \tag{2}$$

The impulse noise itself is assumed to be a *White & Gaussian* process independent of the regressor, and its variance is given by $\sigma^2_{\nu a} = E[|\nu_a(n)|^2]/2$ which is much greater than the variance of the regressor $\sigma^2_a = E[|a(n)|^2]/2$.

III. LEAST MEAN SQUARE NONLINEAR REGRESSOR ALGORITHM

A. Nonlinear Regressor

Define a "nonlinear" function of a complex variable z by

$$\phi_{NL}(|\,z\,|; A) = \begin{cases} 1 & \text{for } |\,z\,| \le A \\ 0 & \text{for } |\,z\,| > A, \end{cases} \tag{3}$$

where $|\cdot|$ denotes *modulus* of a complex number and A is a threshold.

Then, the *Nonlinear Regressor* at the kth tap is given by

$$\phi_{NL}[|\,a(n-k)\,|; A_r]\, a(n-k) \quad (k = 0, \cdots, N-1)$$

with N being the number of taps and A_r a fixed threshold.

B. Least Mean Square Nonlinear Regressor Algorithm

Using the nonlinear function (3) for the filter input $b(n-k)$ which contains impulse noise, we derive an update equation for the kth tap weight $c_k(n)$ as given by

$$c_k(n+1) = c_k(n) + \alpha_c\, e^*(n)\, \phi_{NL}[|\,b(n-k)\,|; A_r]\, b(n-k), \tag{4}$$

where $e(n)$ is the error, α_c is the step size and $(\cdot)^*$ denotes complex conjugate.

The adaptation algorithm defined by (4) is named *Least Mean Square Nonlinear Regressor Algorithm* (LMS-NRA). The factor $\phi_{NL}[|\,b(n-k)\,|; A_r]$ effectively suppresses the peaks of the impulse noise at the filter input by adequately choosing the threshold A_r, making the filter robust against the impulse noise. Usually, we select

$$A_r = M_r\, \sigma_a \tag{5}$$

with $M_r = 3$ (three-sigma of $a(n)$). For $M_r \to \infty$, the algorithm becomes the LMS algorithm. See Fig. 1 for a schematic diagram of the LMS-NRA.

IV. Analysis of LMS-NRA

In this section, a statistical analysis of the LMS-NRA is developed, assuming the presence of the impulse noise at the filter input. For ease of reading, we avoid lengthy derivation processes and only summarize the main results derived.

A. Assumptions

A1: The regressor $a(n)$ is a zero-mean complex-valued colored Gaussian process with a covariance matrix $\mathbf{R_a} = E[\mathbf{a}(n)\, \mathbf{a}^H(n)]/2$ for a regressor vector $\mathbf{a}(n)$ and a variance σ^2_a. The real and imaginary parts of $a(n)$ are mutually uncorrelated.

A2: The filter input with impulse noise is modeled in Section II.

A3: The additive observation noise $\nu(n)$ is independent zero-mean Gaussian noise with a variance $\sigma^2_\nu = E[|\nu(n)|^2]/2$.

A4: The regressor $a(n)$ and the tap weights $\mathbf{c}(n)$ are mutually independent (*Independence Assumption*).

B. Difference Equations for Tap Weight Misalignment

First, the error is given by $e(n) = \in(n) + \nu(n)$, where $\in(n) = \boldsymbol{\theta}^H(n)\, \mathbf{b}(n)$ is the excess error, $\mathbf{b}(n) = [b(n) \cdots b(n-k) \cdots b(n-N+1)]^T$ is the filter input vector, $\boldsymbol{\theta}(n) = \mathbf{h} - \mathbf{c}(n)$ is the tap weight misalignment vector, $\mathbf{h}$ is the impulse response vector of the unknown system, and $\mathbf{c}(n)$ is the tap weight vector.

For the kth tap weight misalignment, we find

$$\theta_k(n+1) = \theta_k(n) - \alpha_c\, e^*(n)\, \phi_{NL}[|\, b(n-k)\,|; A_r]\, b(n-k),$$

from which we can derive a set of difference equations for the mean vector $\mathbf{m}(n) = E[\boldsymbol{\theta}(n)]$ and the second-order moment matrix $\mathbf{K}(n) = E[\boldsymbol{\theta}(n)\, \boldsymbol{\theta}^H(n)]$ as follows.

$$\mathbf{m}(n+1) = \mathbf{m}(n) - \alpha_c\, \mathbf{p}(n) \tag{6}$$

and

$$\mathbf{K}(n+1) = \mathbf{K}(n) - \alpha_c\, [\mathbf{V}(n) + \mathbf{V}^H(n)] + \alpha^2_c\, \mathbf{S}(n), \tag{7}$$

where for the kth element of $\mathbf{p}(n)$, $p_k(n) = E\{e^*(n)\, \phi_{NL}[|\, b(n-k)\,|; A_r]\, b(n-k)\}$ and for the (k, κ)th element of $\mathbf{V}(n)$ and $\mathbf{S}(n)$, $V_{k\kappa}(n) = E\{e^*(n)\, \phi_{NL}[|\, b(n-k)\,|; A_r]\, b(n-k)\, \theta_\kappa^*(n)\}$ and $S_{k\kappa}(n) = E\{|\, e(n)\,|^2\, \phi_{NL}[|\, b(n-k)\,|; A_r]\, b(n-k)\, \phi_{NL}[|\, b(n-\kappa)\,|; A_r]\, b^*(n-\kappa)\}$.

C. Calculation of Matrix $\mathbf{W}(n)$ for $\mathbf{p}(n)$ and $\mathbf{V}(n)$

If the variance of the tap weights is small, we find that $e(n)$ given $b(n-k)$ is a Gaussian random variable whose mean is $\mu_{b\{\tau\}k}^*(n) / \sigma^2_{b\{\tau\}k} \cdot b(n-k)$, where $\boldsymbol{\mu}_{\mathbf{b}\{\tau\}}(n) = \mathbf{R}_{\mathbf{b}\{\tau\}}\, \mathbf{m}(n)$, $\mathbf{R}_{\mathbf{b}\{\tau\}} = \mathbf{R_a} + \mathrm{diag}\{\tau(n)\, \sigma^2_{\nu a}, \cdots, \tau(n-k)\, \sigma^2_{\nu a}, \cdots, \tau(n-N+1)\, \sigma^2_{\nu a}\}$, and $\sigma^2_{b\{\tau\}k} = \sigma^2_a + \tau(n-k)\, \sigma^2_{\nu a}$. Then we have $p_k(n) = \mu_{b\{\tau\}k}(n) / \sigma^2_{b\{\tau\}k} \cdot E[|\, b(n-k)\,|^2;\ |\, b(n-k)\,| \le A_r]$. Knowing that $|\, b(n-k)\,|$ is subject to Rayleigh Distribution, we derive $p_k(n) = \Sigma_{l=0}^{N-1}\, W_{\{\tau\}kl}\, m_l(n)$ and $W_{\{\tau\}k\kappa} = 2\, \mathrm{H_c}(A_r/\sigma_{b\{\tau\}k})\, R_{b\{\tau\}k\kappa}$, where

$$\mathrm{H_c}(r) = \int_0^r t^3/2 \cdot \exp(-t^2/2)dt = 1 - (1 + r^2/2)\exp(-r^2/2).$$

Thus, we obtain

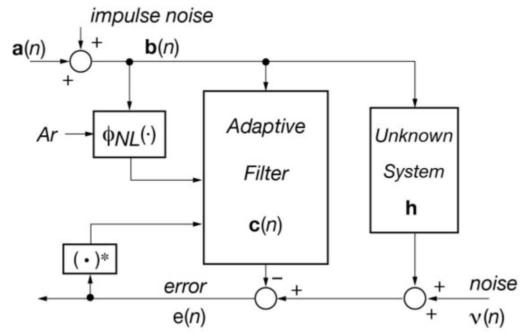

Fig. 1. Schematic diagram for LMS-NRA.

$$\mathbf{p}(n) = \mathbf{W}\, \mathbf{m}(n) \tag{8}$$

and

$$\mathbf{V}(n) = \mathbf{W}\, \mathbf{K}(n), \tag{9}$$

where

$$W_{kk} = 2[(1-p_{\nu a})\, \mathrm{H_c}(A_r/\sigma_a)\, \sigma^2_a + p_{\nu a}\, \mathrm{H_c}(A_r/\sigma_s)\, \sigma^2_s] \tag{10}$$

and

$$W_{k\kappa} = 2[(1-p_{\nu a})\, \mathrm{H_c}(A_r/\sigma_a) + p_{\nu a}\, \mathrm{H_c}(A_r/\sigma_s)]\, R_{ak\kappa} \quad (k \ne \kappa) \tag{11}$$

with $\sigma^2_s = \sigma^2_a + \sigma^2_{\nu a}$.

D. Calculation of Matrix $\mathbf{S}(n)$

Next, we have $S_{k\kappa}(n) = 2\, E\{\sigma^2_{eb}(n)\, \phi_{NL}[|\, b(n-k)\,|; A_r]\, b(n-k)\, \phi_{NL}[|\, b(n-\kappa)\,|; A_r]\, b^*(n-\kappa)\}$, where $\sigma^2_{eb}(n) = \sigma^2_e(n) + p_{\nu a}\sigma^2_{\nu a}\, \Sigma_{l=0}^{N-1}\, K_{ll}(n)$, and $\sigma^2_e(n) = \varepsilon(n) + \sigma^2_\nu$ is the error variance. As a performance measure, we define *Excess Mean Square Error* (EMSE) by

$$\varepsilon(n) = E[|\in(n)\,|^2]/2 = \mathrm{tr}[\mathbf{R_a}\, \mathbf{K}(n)]. \tag{12}$$

Then, for $k = \kappa$, given $\tau(n-k)$, we derive $S_{\{\tau\}kk}(n) = 2E\{\sigma^2_{eb\{\tau\}k}(n)\, \phi_{NL}[|\, b(n-k)\,|; A_r]\, |\, b(n-k)\,|^2\}$, where $\sigma^2_{eb\{\tau\}k}(n) = \sigma^2_{eb}(n) - p_{\nu a}\, \sigma^2_{\nu a}\, K_{kk}(n) + \tau(n-k)\, \sigma^2_{\nu a}\, K_{kk}(n)$ with which we obtain

$$\begin{aligned} S_{kk}(n) &= 4\, [\sigma^2_{eb}(n) - p_{\nu a}\sigma^2_{\nu a}\, K_{kk}(n)] \\ &\quad \cdot [(1-p_{\nu a})\, \mathrm{H_c}(A_r/\sigma_a)\, \sigma^2_a + p_{\nu a}\, \mathrm{H_c}(A_r/\sigma_s)\, \sigma^2_s] \\ &\quad + 4\, p_{\nu a}\sigma^2_{\nu a}\, K_{kk}(n)\, \mathrm{H_c}(A_r/\sigma_s)\, \sigma^2_s \\ &= 4\, [\sigma^2_{eb}(n) - p_{\nu a}\sigma^2_{\nu a}\, K_{kk}(n)]\, W_{kk} \\ &\quad + 4\, p_{\nu a}\sigma^2_{\nu a}\, K_{kk}(n)\, \mathrm{H_c}(A_r/\sigma_s)\, \sigma^2_s. \end{aligned} \tag{13}$$

For $k \ne \kappa$, given $\tau(n-k)$ and $\tau(n-\kappa)$, referring to APPENDIX for the function $\mathrm{H_{c2}}(r_z, r_w; \sin\alpha)$ (see (15)), we obtain

$$\begin{aligned} S_{k\kappa}(n) &= 4\, \{\sigma^2_{eb}(n) - p_{\nu a}\sigma^2_{\nu a}\, [K_{kk}(n) + K_{\kappa\kappa}(n)]\} \\ &\quad \cdot \{\, (1-p_{\nu a})^2\, \mathrm{H_{c2}}(A_r/\sigma_a, A_r/\sigma_a; R_{ak\kappa}/\sigma^2_a)\, \sigma^2_a \\ &\quad + 2\, (1-p_{\nu a})\, p_{\nu a}\, \mathrm{H_{c2}}[A_r/\sigma_a, A_r/\sigma_s; R_{ak\kappa}/(\sigma_a\, \sigma_s)]\, (\sigma_a\, \sigma_s) \\ &\quad + p^2_{\nu a}\, \mathrm{H_{c2}}(A_r/\sigma_s, A_r/\sigma_s; R_{ak\kappa}/\sigma^2_s)\, \sigma^2_s\, \} \\ &\quad + 4\, p_{\nu a}\sigma^2_{\nu a}\, [K_{kk}(n) + K_{\kappa\kappa}(n)] \\ &\quad \cdot \{\, \mathrm{H_{c2}}[A_r/\sigma_a, A_r/\sigma_s; R_{ak\kappa}/(\sigma_a\, \sigma_s)]\, (\sigma_a\, \sigma_s) \\ &\quad + \mathrm{H_{c2}}(A_r/\sigma_s, A_r/\sigma_s; R_{ak\kappa}/\sigma^2_s)\, \sigma^2_s\, \}. \end{aligned} \tag{14}$$

V. Experiments - Numerical Results

In this section, numerical experiments with two examples below are carried out to examine the performance of the proposed LMS-NRA. In the experiments, the simulation result is calculated as an ensemble average of the squared excess error $<|\in(n)\,|^2> / 2$ over 1000 filter runs, and the theoretical convergence is drawn in terms of EMSE $\varepsilon(n)$.

Example #1

number of taps: $N = 4$

regressor: AR1 Gaussian process with variance $\sigma^2_a = 1$ (0 dB) and regression coefficient 0.5 (eigenvalue spread $= \lambda_{max} / \lambda_{min} \cong 5.56$)

unknown system response:

$\mathbf{h} = [0.05–j0.05\ 0.994+j0.01\ 0.01+j0.1\ –0.1+j0.02]^T$

Gaussian observation noise: $\sigma^2_v = 0.01$ (–20 dB)

algorithm: LMSA

LMS-NRA with threshold $M_r = 3$

impulse noise at filter input:

$\sigma^2_{va} = 0$, 100 (+20 dB), 200 (+23 dB), 500 (+27 dB)

$p_{va} = 0.1$

step size: $\alpha_c = 2^{-9}$

Example #2

number of taps: $N = 32$

regressor: AR1 Gaussian process with variance $\sigma^2_a = 1$ (0 dB) and regression coefficient 0.2 (eigenvalue spread $= \lambda_{max} / \lambda_{min} \cong 2.24$)

unknown system response:

$\mathbf{h} = [0.01–j0.05\ \ 0.758–j0.02\ \ 0.05+j0.05\ \ –0.5+j0.1\ \ –0.25+j0.05\ h_5 \cdots h_{31}]^T\ h_k = 0.8\ h_{k–1}, k = 5, \cdots, 31$

Gaussian observation noise: $\sigma^2_v = 0.1$ (–10 dB)

algorithm: LMS-NRA with threshold $M_r = 3$

impulse noise at filter input:

$\sigma^2_{va} = 0$, 100 (+20 dB), 200 (+23 dB), 500 (+27 dB)

$p_{va} = 0.1$

step size: $\alpha_c = 2^{-12}$

In *Example #1*, the number of taps N is small, the regressor is moderately correlated, the SNR ($= \sigma^2_a/\sigma^2_v$) is medium-valued, and the step size is small. The variance of the impulse noise at the filter input σ^2_{va} is chosen 0 (no impulse noise), 100, 200 and 500. First, for the conventional LMSA, the results of experiments are shown in Fig. 2, where for $\sigma^2_{va} = 0$ the filter converges normally to the steady-state value (−41.0 dB). For $\sigma^2_{va} = 100$, the simulated filter convergence is stable but the steady-state error is a few dB larger. With $\sigma^2_{va} = 200$, the filter seems non-divergent but very unstable. When $\sigma^2_{va} = 500$, the filter is completely divergent. Thus, with the LMSA, the adaptive filter is quite vulnerable to the impulse noise at the filter input.

Also for *Example #1*, Fig. 3 shows the results for the proposed LMS-NRA. With no impulse noise at the filter input, the filter convergence is slightly slower than that for the LMSA due to the finite threshold A_r. As σ^2_{va} increases, the filter convergence becomes slower and the steady-state error larger, but even for $\sigma^2_{va} = 500$ the filter converges stably, that shows the effectiveness of the "Nonlinear Regressor" in enhancing the robustness against the impulse noise at the filter input.

In *Example #2*, N is large, the regressor is weakly correlated, the SNR is low, and the step size is small. The results are shown in Fig. 4, where we again observe successful robust filtering.

Generally, we see good agreement between simulated and theoretically calculated filter convergence curves.

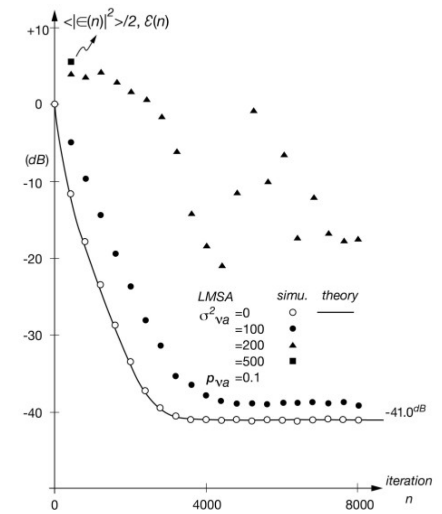

Fig. 2. Adaptive filter convergence.
(*Example #1*, $N = 4$, LMSA, $\sigma^2_{va} = 0$, 100, 200 & 500)

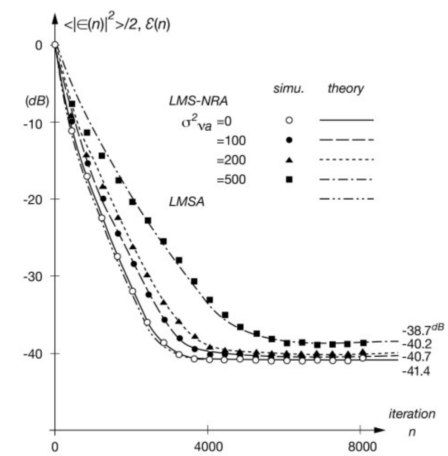

Fig. 3. Adaptive filter convergence.
(*Example #1*, $N = 4$, LMS-NRA, $\sigma^2_{va} = 0$, 100, 200 & 500)

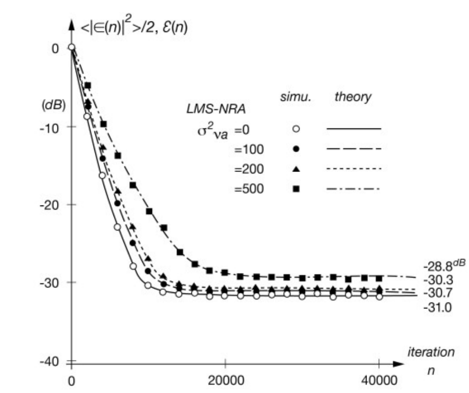

Fig. 4. Adaptive filter convergence.
(*Example #2*, $N = 32$, LMS-NRA, $\sigma^2_{va} = 0$, 100, 200 & 500)

VI. CONCLUSION

In this paper, we have proposed a new adaptation algorithm LMS-NRA for adaptive filters with impulse noise at the filter input, employing a simple nonlinear function of the regressor in the tap weight update equation.

A statistical analysis of the LMS-NRA has been developed which is considerably complicated, though the nonlinear function $\phi_{NL}(\ .\ )$ is very simple. This is because we have dealt with stochastic appearance of the impulse noise at each tap.

Through numerical experiments, we have demonstrated that the proposed algorithm makes adaptive filters convergent as fast as the LMSA, yet sufficiently robust in the presence of impulse noise added to the regressor at the input of the adaptive filter.

In the experiments we also observe that the theoretically calculated filter convergence curves are in good agreement with the simulation results that shows the validity and accuracy of the analysis.

Comparison with other impulse noise robust algorithms is left for further study.

APPENDIX

Derivation of Function $H_{c2}(r_z, r_w; \sin\alpha)$

Let $z = x + j\,y$ and $w = u + j\,v$ be zero-mean complex-valued Gaussian random variables, for which $E(x^2) = E(y^2) = \sigma^2_z$, $E(u^2) = E(v^2) = \sigma^2_w$, $E(x\,u) = E(y\,v) = R_{zw}$, and $E(x\,y) = E(u\,v) = E(x\,v) = E(y\,u) = 0$ hold. Here, the real and imaginary parts are uncorrelated, and the real (or imaginary) parts of z and w are correlated with covariance R_{zw}.

The joint PDF of x and u is given by [8]

$$p(x, u) = (2\pi)^{-1}(\sigma^2_z\,\sigma^2_w - R^2_{zw})^{-1/2} \cdot \exp[-(\sigma^2_w\,x^2 - 2\,R_{zw}\,x\,u + \sigma^2_z\,u^2) / (\sigma^2_z\,\sigma^2_w - R^2_{zw}) / 2].$$

Likewise, the joint PDF of y and v is given by

$$p(y, v) = (2\pi)^{-1}(\sigma^2_z\,\sigma^2_w - R^2_{zw})^{-1/2} \cdot \exp[-(\sigma^2_w\,y^2 - 2\,R_{zw}\,y\,v + \sigma^2_z\,v^2) / (\sigma^2_z\,\sigma^2_w - R^2_{zw}) / 2].$$

Then the joint PDF of z and w is:

$$p(z, w) = p(x, u)\,p(y, v) = (2\pi)^{-2}(\sigma^2_z\,\sigma^2_w - R^2_{zw})^{-1} \cdot \exp\{-[\sigma^2_w\,|z|^2 - 2\,R_{zw}\,(x\,u + y\,v) + \sigma^2_z\,|w|^2] / (\sigma^2_z\,\sigma^2_w - R^2_{zw}) / 2\}.$$

Converting the variables (z, w) to polar coordinates $(\zeta, \omega, \varphi, \psi)$ via $z = \sigma_z\,\zeta\,\exp(j\varphi)$ and $w = \sigma_w\,\omega\,\exp(j\psi)$, and defining $\sin\alpha = R_{zw} / (\sigma_z\,\sigma_w)$ where we assume $|\sin\alpha| < 1$, we obtain the following joint PDF.

$$p(\zeta, \omega, \varphi, \psi) = \sec^2\alpha\,\zeta\,\omega\,\exp[-(\zeta^2 + \omega^2)\sec^2\alpha/2] \cdot (2\pi)^{-2}\exp[\zeta\,\omega\,\sec^2\alpha\,\sin\alpha\,\cos(\varphi - \psi)],$$

where $0 \le \zeta, \omega < \infty$ and $0 \le \varphi, \psi < 2\pi$.

Using the above PDF, let us calculate expectation

$$E[\phi_{NL}(|z|; A)\,z\,\phi_{NL}(|w|; A)\,w^*] = E(z\,w^*; |z| \le A,\ |w| \le A).$$

We find, with $r_z = A / \sigma_z$ and $r_w = A / \sigma_w$,

$$\begin{aligned} &E(z\,w^*; |z| \le A,\ |w| \le A) \\ &= \sigma_z\,\sigma_w\,E\{\zeta\,\omega\,\exp[j(\varphi-\psi)];\ \zeta \le r_z,\ \omega \le r_w\} \\ &= \sigma_z\,\sigma_w \int_0^{r_z} d\zeta \int_0^{r_w} d\omega \int_0^{2\pi} d\varphi \int_0^{2\pi} d\psi \cdot \zeta\,\omega\,\exp[j(\varphi-\psi)]\,p(\zeta, \omega, \varphi, \psi) \\ &= \sigma_z\,\sigma_w\,\sec^2\alpha \int_0^{r_z} d\zeta \int_0^{r_w} d\omega\,\zeta^2\,\omega^2\,\exp[-(\zeta^2+\omega^2)\sec^2\alpha/2] \\ &\quad \cdot (2\pi)^{-1}\int_0^{2\pi} d\varphi\,(2\pi)^{-1}\int_0^{2\pi} d\psi \cdot \exp[j(\varphi-\psi)]\,\exp[\zeta\,\omega\,\sec^2\alpha\,\sin\alpha\,\cos(\varphi-\psi)] \\ &= \sigma_z\,\sigma_w\,\sec^2\alpha \int_0^{r_z} d\zeta \int_0^{r_w} d\omega\,\zeta^2\,\omega^2 \cdot \exp[-(\zeta^2+\omega^2)\sec^2\alpha/2]\,I_1(\zeta\,\omega\,\sec^2\alpha\,\sin\alpha), \end{aligned}$$

where $I_1(x)$ is the First-order Modified Bessel Function of the First Kind [9].

Thus we finally obtain

$$E[\phi_{NL}(|z|; A)\,z\,\phi_{NL}(|w|; A)\,w^*] = 2\,\sigma_z\,\sigma_w\,H_{c2}(r_z, r_w; \sin\alpha),$$

for which we define a function

$$H_{c2}(r_z, r_w; \sin\alpha) = \sec^2\alpha / 2 \cdot \int_0^{r_z} \zeta^2 \exp(-\zeta^2\sec^2\alpha/2)\,d\zeta \cdot \int_0^{r_w} \omega^2 \exp(-\omega^2\sec^2\alpha/2)\,d\omega\,I_1(\zeta\,\omega\,\sec^2\alpha\,\sin\alpha). \quad (15)$$

Note that $H_{c2}(0, r_w; \sin\alpha) = H_{c2}(r_z, 0; \sin\alpha) = 0$, $H_{c2}(r_z, r_w; 0) = 0$ and $H_{c2}(\infty, \infty; \sin\alpha) = \sin\alpha$ hold.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice-Hall, 1996.

[2] S. Koike, "Effects of impulse noise at filter input on performance of adaptive filters using the LMS and signed LMS algorithms," in *Proc. IEEE ISPACS 2006*, Tottori, Japan, Dec. 2006, pp. 829-832.

[3] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. ASSP*, vol. 34, no. 8, pp. 793-806, Aug. 1986.

[4] S. Koike, "Analysis of adaptive filters using normalized signed regressor LMS algorithm," *IEEE Trans. Signal Processing*, vol. 47, no. 10, pp. 2710-2723, Oct. 1999.

[5] S. Koike, "Adaptive threshold nonlinear errors - regressor phase algorithm," in *Proc. IEEE ISCIT 2019*, Ho Chi Minh City, Vietnam, Sept. 2019, pp. 1-4.

[6] S. Koike, "Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulse noise," *IEEE Trans. Signal Processing*, *Corresp*., vol. 45, no. 9, pp. 2391-2395, Sept. 1997.

[7] S. Koike, "Analysis of normalized least mean modulus algorithm for adaptive filters in impulsive noise environments," in *Proc. IEEE ISPACS 2010*, Chengdu, China, Dec. 2010, TA1.3.3.

[8] A. Papoulis, *Probability*, *Random Variables and Stochastic Processes*. Singapore: McGraw-Hill, 1984.

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York; Dover, 1964.