Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/

ANALYSIS OF SPARSE CONTROLLED PROPORTIONATE RECURSIVE LEAST SQUARES ALGORITHM

Shiv Ram Meena* **and C.S. Rai**
University School of Information, Communication & Technology,
Guru Gobind Singh Indraprastha University, Dwarka, New Delhi (INDIA).
*E-mail: srm.iitr@gmail.com

1. Abstract

The sparse controlled proportionate recursive least square algorithm performs well for the systems with variable sparsity and for systems with fixed sparsity. The algorithm updates each filter weight iteratively with different gain factors to increase the rate of convergence. It uses the degree of sparseness to calculate the gain factor during iterations. The estimated degree of sparseness is calculated for the estimated weight vector. The proportionate matrix assigns independent gain to each filter weight while updating the filter taps. For fixed sparse systems, the degree of sparseness converges to the original degree of sparseness faster than the filter weight vector, which increases the algorithm's convergence rate. The convergence controlling parameter μ is crucial to the SC-PRLS's effectiveness. Therefore, Analysis of both transient and steady-state performance is necessary. This study investigates the performance of the mean square error of the SC-PRLS algorithm. Energy conservation principle is applied to calculate the mean square performance of transient and steady-state stages. Explicit conditions are obtained to ensure better performance of the algorithm. The selection of convergence controlling parameter μ also depends on the number of filter weights. A larger tap length requires higher μ for faster convergence, while lower μ is needed for lower steady-state error. The optimum value is 0.65, and the range of convergence controlling parameter μ is 30-65 for better performance.

Key points: Analysis of SC-PRLS, convergence controlling parameter, MSE (mean square error), MMSE.

2. Introduction

Sparse system identification (SSI) has various applications where unknown systems can be modeled as FIR systems having finite memory, including telecommunications, control engineering, sensing, and acoustics [1-3]. Few applications have variable sparsity, like channel estimation for underwater acoustic (UWA) communication and wireless channel estimation, etc., where a large number of parameters and time-varying dynamics makes the channel impulse response (CIR) sparse [4-6]. Most sparse adaptive filtering techniques are based either on the zero-attractor (ZA) or proportionate update (PU) concept [3]. Zero attractor-based algorithms perform better for strict sparse systems, while PU-type algorithms are useful for relatively sparse systems [6]. The sparse aspect of the system is used by proportionate-type algorithms to increase convergence [7]. However, if system behavior changes from sparse to non-sparse or there is significant change in sparsity, the performance of PU-type algorithms degrades [6]. So, their improved versions were developed. Examples of PU-type algorithms are Proportionate normalized least mean square (PNLMS), Improved PNLMS (IPNLMS), Proportionate recursive least square (PRLS), L0-PRLS, etc. [5-8]. But these algorithms do not work well for systems having variable sparsity. There are many practical applications where sparsity changes. For example, In an underwater acoustic channel movement in the position of transmitter or receiver, variable environmental noise, fluctuating water surface, and depth-varying sound speed profile result in the time-varying sparseness of the system [6]. So, sparse controlled algorithms are much needed for applications where the degree of sparseness changes [4-6]. The sparse controlled-IPNLMS (SC-IPNLMS) and sparse-controlled proportionate recursive least square (SC-PRLS) algorithms were developed for such systems, and their performance was also analyzed [9-10].

Performance analysis is available in the literature for many sparse algorithms [5-17]. White input signals were considered to overcome the nonlinearity problem of sparsity regularizes in ZA-type algorithms for transient behavior analysis [17]. Wagner presented an analytical model in [7] that characterizes the transient and steady-state performance for PU-type sparse adaptive filtering algorithms and performance study of the PNLMS method. The performance analysis of the compressed distributed least square algorithm is explained in [11]. Ayoub Tedjani et. al. in [14] compare the performance in terms of mean square error (MSE) for IPNLMS, SC-IPNLMS, and compressive sensing based VSS-RZA-NLMS algorithm for echo cancellation. Zhen Qin investigated how the PRLS algorithm performed for time-invariant systems in both the transient and steady-state stages [6]. Z.Qin also analyzed the performance of PMCC [15],l1-norm with PRLS [5], l1-RLS[8] and PRLS algorithms [6].

This work explains the performance analysis of transient and steady state errors (SSE) for SC-PRLS. Excess MSE (EMSE) eMSExpression for the algorithm is also derived. Effect of convergence controlling parameter μ in EMSE is also discussed. Numerical simulations and experimental sparse system identification results verify theoretical results. Explicit conditions are developed to increase the efficiency of the SC-PRLS algorithm.

3. SC-PRLS Algorithm

To analyse SC-PRLS, standard system identification settings are considered. The desired output of the system to be identified is given as

$$d(k) = \omega_M^H x(k) + \vartheta(k) \tag{1}$$

Where vector $\omega_M = [\omega_1, \omega_2, \omega_3, \cdots, \omega_M]^T$ is the impulse response of system of unknown system. This unknown system is to be identified in the system identification problem. And input at time instant k is given by $x(k) = [x(k), x(k-1), x(k-3), \cdots, x(n-M+1)]^T$. The scalar $\vartheta(k)$ is additive noise. The update equation given to track the system parameters in system identification problems. The update equation for estimated filter weights $\varpi(k)$ in the SC-PRLS algorithm is given by

$$\varpi(k) = \varpi_M(k-1) + G(k-1)\mathcal{K}(k)e^*(k|k-1) \tag{2}$$

Where $e(k|k-1)$ is the a priori error and calculated as

$$e(k|k-1) = d(k) - \varpi_M^H(k-1)x(k), \tag{3}$$

Kalman gain vector is $\mathcal{K}(k) = \frac{\mathcal{P}(k-1)x(k)}{\lambda + x^H(k)\mathcal{P}(k-1)x(k)}$ (4)

And forgetting factor λ is defined in the range $(0,1)$. Inverse of input covariance matrix is calculated iteratively as

$$\mathcal{P}(k) = \lambda^{-1}[\mathcal{P}(k-1) - \mathcal{K}(k)x^H(k)\mathcal{P}(k-1)] \quad (5)$$

The proportionate matrix G(k-1) is diagonal matrix written as

$$G(k-1) = diag\{g_1(k-1), g_2(k-1), \cdots, g_M(k-1)\} \quad (6)$$

The nth element of proportionate matrix can be calculated as

$$g_n = \mu \frac{\left(1-\frac{\xi(k)}{2}\right)}{M}\left(\frac{1-\alpha_{sc}}{2M}\right) + \mu \frac{\left(1+\frac{\xi(k)}{2}\right)}{M} \frac{(1+\alpha_{sc})|\varpi_k(n)|}{\|\varpi_M(k)\|_1 + \varepsilon_{sc}} \quad (7)$$

Where regularization parameter ε_{sc} is a small positive constant $\varepsilon_{sc} \epsilon [-1,1]$ and $\xi(k)$ is the degree of sparseness that can be calculated as

$$\xi(k) = \frac{M}{M-\sqrt{M}}\left(1 - \frac{\|\varpi_k(n)\|_1}{2\|\varpi_k(n)\|_2}\right) \quad (8)$$

Where $\|.\|_1$ is l_1- norm and $\|.\|_2$ is l_2- norm. T

For an impulse response of FIR filter the degree of sparseness for depends on the number of total taps and sum of absolute value of these taps as well as magnitude of filter coefficients. Therefore, it can be understood that degree of sparseness depends only on magnitude of filter taps (non-zero filter coefficients) and length of FIR filter.

4. Performance Analysis of the Algorithm

Statistically steady state behaviour and transient response of the SC-PRLS can be evaluated by principle of energy conservation.

Kalman gain vector can be rearranged using equation (4) as follows

$$\lambda\mathcal{K}(k) + \mathcal{K}(k)x^H(k)\mathcal{P}(k-1)x(k) = \mathcal{P}(k-1)x(k)$$

$$\lambda\mathcal{K}(k) = \mathcal{P}(k-1)x(k) - \mathcal{K}(k)x^H(k)\mathcal{P}(k-1)x(k)$$

$$\mathcal{K}(k) = \lambda^{-1}[\mathcal{P}(k-1) - \mathcal{K}(k)x^H(k)\mathcal{P}(k-1)]x(k)$$

Using equation (5) here, Kalman gain vector can be written as

$$\mathcal{K}(k) = \mathcal{P}(k)x(k) \quad (9)$$

Now weight update equation can be rewritten using equation (2) and (9) as

$$\varpi_M(k) = \varpi_M(k-1) + G(k-1)\mathcal{P}(k)x(k)e^*(k|k-1) \quad (10)$$

Now, Define weight error vector $\hbar_M(k)$ for the SC-PRLS algorithms as

$$\hbar_M(k) = \omega_M - \varpi_M(k) \quad (11)$$

Where ω_M is impulse response vector need to be identified and $\varpi_M(k)$ is estimated impulse response vector by the algorithm.

subtracting equation (10) from ω_M both side

$$\omega_M - \varpi_M(k) = \omega_M - \varpi_M(k-1) - G(k-1)\mathcal{P}(k)x(k)e^*(k|k-1) \quad (12)$$

Using equation (11) and (12) weight error vector can be written as

$$\hbar_M(k) = \hbar_M(k-1) - G(k-1)\mathcal{P}(k)x(k)e^*(k|k-1) \quad (13)$$

Let, Σ be a Hermitian nonnegative-definite matrix, then Pre-multiplying $x^H(k)\mathcal{P}(k)G(k-1)\Sigma$ to the equation (13) both side

$$x^H(k)\mathcal{P}(k)G(k-1)\Sigma\hbar_M(k) = x^H(k)\mathcal{P}(k)G(k-1)\Sigma\hbar_M(k-1) - x^H(k)\mathcal{P}(k)G(k-1)\Sigma G(k-1)\mathcal{P}(k)x(k)e^*(k|k-1)$$

This can be rewritten using Hermitian matrix properties as

$$[\hbar_M^H(k)\Sigma G(k-1)\mathcal{P}(k)x(k)]^* = [\hbar_M^H(k-1)\Sigma G(k-1)\mathcal{P}(k)x(k)]^* - x^H(k)(\mathcal{P}(k)G(k-1)\Sigma G(k-1)\mathcal{P}(k))x(k)e^*(k|k-1) \quad (14)$$

As $G(n-1)$ is diagonal matrix, $\mathcal{P}(k)$ is inverse of input correlation matrix so taking complex conjugate will not affect them. Because $A^H = [A^*]^T = [A^T]^*$
Therefore $G^H(k-1) = G(k-1)$ and $\mathcal{P}^H(k) = \mathcal{P}(k)$
Using properties of Hermitian matrix, let define new Hermitian matrixes

$$\Sigma_1 = \Sigma G(k-1)\mathcal{P}(k)\mathcal{P}(k), \tag{15}$$

$$\Sigma_2 = \mathcal{P}(k)G(k-1)\Sigma G(k-1)\mathcal{P}(k) \tag{16}$$

Equation (14) can be written as

$$[\hbar_M^H(k)\Sigma_1 x(k)]^* = [\hbar_M^H(k-1)\Sigma_1 x(k)]^* - x^H(k)\Sigma_2 x(k)e^*(k|k-1) \tag{17}$$

Now, weighted error can be refined as a priori and p posteriori as
Weighted a priori error $e_a(k) = \hbar_M^H(k-1)\Sigma_1 x(k),$ (18)
Weighted posteriori error $e_p(k) = \hbar_M^H(k)\Sigma_1 x(k)$ (19)
Now, equation (17) can be rewritten as

$$e_p(k)^* = e_a(k)^* - x^H(k)\Sigma_2 x(k)e^*(k|k-1)$$

$$e_p(k)^* = e_a(k)^* - \|x(k)\|_{\Sigma_2}^2 e^*(k|k-1) \tag{20}$$

From equation (20), substitute $e^*(k|k-1)$ into (13)

$$\hbar_M(k) = \hbar_M(k-1) - G(k-1)\mathcal{P}(k)x(k)\left(\frac{e_a(k)^* - e_p(k)^*}{\|x(k)\|_{\Sigma_2}^2}\right)$$

$$\hbar_M(k) + \frac{G(k-1)\mathcal{P}(k)x(k)e_a(k)^*}{\|x(k)\|_{\Sigma_2}^2} = \hbar_M(k-1) + \frac{G(k-1)\mathcal{P}(k)x(k)e_p(k)^*}{\|x(k)\|_{\Sigma_2}^2} \tag{21}$$

To obtain weighted energy relation, take weighted Euclidean norm of equation (21) on both side with respect to Σ

$$\left\|\hbar_M(k) + \frac{G(k-1)\mathcal{P}(k)x(k)e_a(k)^*}{\|x(k)\|_{\Sigma_2}^2}\right\|_{\Sigma}^2 = \left\|\hbar_M(k-1) + \frac{G(k-1)\mathcal{P}(k)x(k)e_p(k)^*}{\|x(k)\|_{\Sigma_2}^2}\right\|_{\Sigma}^2 \tag{22}$$

Note: Norm is a real number because weighted Euclidean norm $\|.\|_{\Sigma}^2$ is the distance with respect to Σ. So, equation (22) can be written as

$$\|\hbar_M(k)\|_{\Sigma}^2 + \frac{|e_a{}^*(k)|^2}{\|x(k)\|_{\Sigma_2}^2} = \|\hbar_M(k-1)\|_{\Sigma}^2 + \frac{|e_p{}^*(k)|^2}{\|x(k)\|_{\Sigma_2}^2} \tag{23}$$

Substituting $e_p(k)$ from equation (20) in (23)

$$\|\hbar_M(k)\|_{\Sigma}^2 + \frac{|e_a{}^*(k)|^2}{\|x(k)\|_{\Sigma_2}^2} = \|\hbar_M(k-1)\|_{\Sigma}^2 + \frac{\left|e_a(k)^* - \|x(k)\|_{\Sigma_2}^2 e^*(k|k-1)\right|^2}{\|x(k)\|_{\Sigma_2}^2}$$

$$\|\hbar_M(k)\|_{\Sigma}^2 + \frac{|e_a{}^*(k)|^2}{\|x(k)\|_{\Sigma_2}^2} = \|\hbar_M(k-1)\|_{\Sigma}^2 + \frac{|e_a{}^*(k)|^2}{\|x(k)\|_{\Sigma_2}^2} - \frac{\|x(k)\|_{\Sigma_2}^2\left(e_a{}^*(k)e(k|k-1) + e_a(k)e^*(k|k-1)\right)}{\|x(k)\|_{\Sigma_2}^2} + \frac{\left(\|x(k)\|_{\Sigma_2}^2\right)^2 |e(k|k-1)|^2}{\|x(k)\|_{\Sigma_2}^2}$$

$$\|\hbar_M(k)\|_{\Sigma}^2 = \|\hbar_M(k-1)\|_{\Sigma}^2 - e(k|k-1)e_a{}^*(k) - e_a(k)e^*(k|k-1) + \|x(k)\|_{\Sigma_2}^2 |e(k|k-1)|^2 \tag{24}$$

Now, using equation (1), (3) and (11) a priori error can be rewritten as

$$e(k|k-1) = \omega_M^H x(k) + \vartheta(k) - \varpi_M^H(k-1)x(k) = [\omega_M^H - \varpi_M^H(k-1)]x(k) + \vartheta(k)$$

$$e(k|k-1) = \hbar_M^H(k-1)x(k) + \vartheta(k) \tag{25}$$

Putting $e(k|k-1)$ from equation (25) and $e_a(k)$ from (18) into (24) the 2nd ,3rd and 4th term in right hand side of equation (24) can be written as

$$\|\hbar_M(k)\|_{\Sigma}^2 = \|\hbar_M(k-1)\|_{\Sigma}^2 - [\hbar_M^H(k-1)\Sigma_1 x(k)]^*[\hbar_M^H(k-1)x(k)+\vartheta(k)] - [\hbar_M^H(k-1)\Sigma_1 x(k)][\hbar_M^H(k-1)x(k)+\vartheta(k)]^* + \|x(k)\|_{\Sigma_2}^2|\hbar_M^H(k-1)x(k)+\vartheta(k)|^2$$

$$= \|\hbar_M(k-1)\|_{\Sigma}^2 - [\hbar_M^H(k-1)\Sigma_1 x(k)]^*\hbar_M^H(k-1)x(k) - [\hbar_M^H(k-1)\Sigma_1 x(k)]^*\vartheta(k) - \hbar_M^H(k-1)\Sigma_1 x(k)[\hbar_M^H(k-1)x(k)]^* - \hbar_M^H(k-1)\Sigma_1 x(k)\vartheta^*(k) + \|x(k)\|_{\Sigma_2}^2|\hbar_M^H(k-1)x(k)|^2 + \|x(k)\|_{\Sigma_2}^2|\vartheta(k)|^2 + \|x(k)\|_{\Sigma_2}^2[2|\vartheta(k)||x^H(k)\hbar_M(k-1)|]$$

$$\begin{aligned}\|\hbar_M(k)\|_{\Sigma}^2 = \|\hbar_M(k-1)\|_{\Sigma}^2 &- \hbar_M^H(k-1)x(k)x^H(k){\Sigma_1}^H\hbar_M(k-1) \\ &- \vartheta(k)x^H(k){\Sigma_1}^H\hbar_M(k-1) - \hbar_M^H(k-1)\Sigma_1 x(k)x^H(k)\hbar_M(k-1) \\ &- \hbar_M^H(k-1)\Sigma_1 x(k)\vartheta^*(k) \\ &+\hbar_M^H(k-1)\|x(k)\|_{\Sigma_2}^2 x(k)x^H(k)\hbar_M(k-1) + \|x(k)\|_{\Sigma_2}^2|\vartheta(k)|^2 \\ &+\|x(k)\|_{\Sigma_2}^2[2|\vartheta(k)||x^H(k)\hbar_M(k-1)|] \end{aligned} \quad (26)$$

By applying statistical expectations $E[.]$ both side on equation (26) and considering the following assumptions.

"(Assumption 1: The noise $\vartheta(k)$ is independent to input vector $x(m)$ for $\forall m$. it can be defined as a gaussian sequence with zero mean and variance σ_ϑ^2 Which is independent and identically distributed.

Assumption 2: Under slow adaptions the regressor $x_M(k)$ and the filter weight vector $\varpi(k-1)$ are uncorrelated.)" [4]

$$\begin{aligned}E[\|\hbar_M(k)\|_{\Sigma}^2] = E[\|\hbar_M(k-1)\|_{\Sigma}^2] &- E\left[\hbar_M^H(k-1)x(k)x^H(k){\Sigma_1}^H\hbar_M(k-1)\right] - \\ E\left[\vartheta(k)x^H(k){\Sigma_1}^H\hbar_M(k-1)\right] &- E[\hbar_M^H(k-1)\Sigma_1 x(k)x^H(k)\hbar_M(k-1)] - \\ E[\hbar_M^H(k-1)\Sigma_1 x(k)\vartheta^*(k)] &+ E\left[\hbar_M^H(k-1)\|x(k)\|_{\Sigma_2}^2 x(k)x^H(k)\hbar_M(k-1)\right] + \\ E\left[\|x(k)\|_{\Sigma_2}^2|\vartheta(k)|^2\right] &+ E\left[2|\vartheta(k)|\|x(k)\|_{\Sigma_2}^2|x^H(k)\hbar_M(k-1)|\right]\end{aligned}$$

From assumption one and two, $E[x(k)\vartheta(k)] = 0$, $E[|\vartheta(k)|^2] = \sigma_\vartheta^2$, and $E[x(k)\hbar_M(k-1)] = 0$ therefore,

$$E[\|\hbar_M(k)\|_{\Sigma}^2] = E[\|\hbar_M(k-1)\|_{\Sigma}^2] - E\left[\|\hbar_M(k-1)\|_{x(k)x^H(k)\Sigma_1^H}^2\right] - E\left[\|\hbar_M(k-1)\|_{\Sigma_1 x(k)x^H(k)}^2\right] + E\left[\|\hbar_M(k-1)\|_{\|x(k)\|_{\Sigma_2}^2 x(k)x^H(k)}^2\right] + \sigma_\vartheta^2 E\left[\|x(k)\|_{\Sigma_2}^2\right]$$

$$E[\|\hbar_M(k)\|_{\Sigma}^2] = E\left[\|\hbar_M(k-1)\|_{\Sigma - x(k)x^H(k)\Sigma_1^H - \Sigma_1 x(k)x^H(k) + \|x(k)\|_{\Sigma_2}^2 x(k)x^H(k)}^2\right] + \sigma_\vartheta^2 E\left[\|x(k)\|_{\Sigma_2}^2\right]$$

$$E[\|\hbar_M(k)\|_{\Sigma}^2] = E\left[\|\hbar_M(k-1)\|_{\Sigma_3}^2\right] + \sigma_\vartheta^2 E\left[\|x(k)\|_{\Sigma_2}^2\right] \quad (27)$$

Where $\Sigma_3 = \Sigma - x(k)x^H(k)\Sigma_1^H - \Sigma_1 x(k)x^H(k) + \|x(k)\|_{\Sigma_2}^2 x(k)x^H(k)$ (28)

Replacing Σ_1 in equation (28) from (15)

$$\Sigma_3 = \Sigma + \|x(k)\|_{\Sigma_2}^2 x(k)x^H(k) - x(k)x^H(k)\mathcal{P}(k)G(k-1)\Sigma - \Sigma G(k-1)\mathcal{P}(k)x(k)x^H(k) \quad (29)$$

The weight error vector $\hbar_M(k)$ and proportionate matrix $G(k-1)$ depends on estimated impulse response vector $\varpi_M(k)$ and degree of sparseness. If long system impulse response is considered and/or the SC-PRLS algorithm gets converge then for time index k to $k+1$, proportionate matrix $G(k-1)$ does not changes significantly because degree of sparseness also converges to optimum value ξ^*, which is the degree of sparseness for the system to be identified (whose impulse response vector is ω_M) [18-19]. Therefore, from equation (29) and

(27), it can be assumed that weight error vector $\hbar_M(k-1)$ is nearly independent of Σ_3. Hence, $E\left[\|\hbar_M(k-1)\|^2_{\Sigma_3}\right] \approx \|\hbar_M(k-1)\|^2_{E[\Sigma_3]}$ can be used in equation (27).

$$E[\|\hbar_M(k)\|^2_{\Sigma}] \approx \|\hbar_M(k-1)\|^2_{E[\Sigma_3]} + \sigma^2_{\vartheta} E\left[\|x(k)\|^2_{\Sigma_2}\right] \quad (30)$$

Further, for theoretical analysis, the approximations of standard RLS can be followed here also. As Auto covariance matrix is approximated to steady state mean value. Similarly, $E[G(k-1)]$can be replaced to steady state value of proportionate matrix G. This G can be constructed by calculated its elements by formula given below

$$g_n = \frac{\mu}{2M}\left(\frac{(1-\alpha_{sc})}{M}\left(1-\frac{\xi^*}{2}\right) + (2+\xi^*)\frac{(1+\alpha_{sc})|\omega_n|}{\|\omega_M\|_1}\right) \quad (31)$$

This equation can be used to calculate theoretical value of proportionate matrix and excess mean square error value.

5. Derivation of excess mean square error (EMSE)

Based on the above approximations and assumption 1, the term $E[\Sigma_3]$ and $E\left[\|x(k)\|^2_{\Sigma_2}\right]$ can be simplified as

$$\mathrm{E}[\Sigma_3] \approx \Sigma + E[\|x(k)\|^2_{\Sigma_2} x(k)x^H(k)] - \mathrm{E}[x(k)x^H(k)](1-\lambda)\mathcal{R}^{-1}G\Sigma$$
$$-\Sigma G(1-\lambda)\mathcal{R}^{-1}E[x(k)x^H(k)]$$

$$\mathrm{E}[\Sigma_3] \approx \Sigma + E[\|x(k)\|^2_{(1-\lambda)\mathcal{R}^{-1}G\Sigma G(1-\lambda)\mathcal{R}^{-1}} x(k)x^H(k)] - (1-\lambda)G\Sigma - \Sigma G(1-\lambda) \quad (32)$$

$$E\left[\|x(k)\|^2_{\Sigma_2}\right] \approx (1-\lambda)^2 tr(G\mathcal{R}^{-1}G\Sigma) \quad (33)$$

Now equation (30) can be written as

$$E[\|\hbar_M(k)\|^2_{\Sigma}] \approx \|\hbar_M(k-1)\|^2_{E[\Sigma_3]} + \sigma^2_{\vartheta}(1-\lambda)^2 tr(G\mathcal{R}^{-1}G\Sigma) \quad (34)$$

As $k \to \infty$, steady state error or excess mean square error (EMSE) will be given by 2nd term. This expression is quite complicated still we can understand that EMSE will depend only on noise power, forgetting factor and optimum value of proportionate matrix G. The optimum value proportionate matrix can be calculated by equation (30). Therefore, it depends on convergence control parameter and degree of sparseness of system to be identified.

Further, equation (20) can be written as

$$e_p(k)^* = e_a(k)^* - x^H(k)G(k-1)\mathcal{P}(k)x(k)e^*(k|k-1) \quad (35)$$

equation (21) can be written as

$$\hbar_M(k) + \frac{G(k-1)\mathcal{P}(k)x(k)e_a(k)^*}{x^H(k)G(k-1)\mathcal{P}(k)x(k)} = \hbar_M(k-1) + \frac{G(k-1)\mathcal{P}(k)x(k)e_p(k)^*}{x^H(k)G(k-1)\mathcal{P}(k)x(k)} \quad (36)$$

Taking weighted Euclidean norm with respect to $\Sigma_4(G(k-1)\mathcal{P}(k))^{-1}$ on both side of equation (36)

$$\|\hbar_M(k)\|^2_{\Sigma_4} + \frac{|e^*_a(k)|^2}{x^H(k)\Sigma_4{}^{-1}x(k)} = \|\hbar_M(k-1)\|^2_{\Sigma_4} + \frac{|e^*_p(k)|^2}{x^H(k)\Sigma_4{}^{-1}x(k)} \quad (37)$$

A priory error can also be written as

$$e_a(n) = \hbar^H_M(k-1)x(k) \quad (39)$$

$$e_p(n) = \hbar^H_M(k)x(k) \quad (40)$$

Using equation (25), (35) and (40) in (37) to eliminate e_p

$$\|\hbar_M(k)\|^2_{\Sigma_4} + \frac{|e^*_a(k)|^2}{x^H(k){\Sigma_4}^{-1}x(k)}$$
$$= \|\hbar_M(k-1)\|^2_{\Sigma_4} + \frac{|e^*_a(k) - x^H(k)G(k-1)\mathcal{P}(k)x(k)(\hbar^H_M(k-1)x(k)+\vartheta(k))^*|^2}{x^H(k){\Sigma_4}^{-1}x(k)}$$

$$\|\hbar_M(k)\|^2_{\Sigma_4}$$
$$= \|\hbar_M(k-1)\|^2_{\Sigma_4} + \frac{|e^*_a(k)|^2}{x^H(k){\Sigma_4}^{-1}x(k)}$$
$$+ \frac{|e^*_a(k) - x^H(k)G(k-1)\mathcal{P}(k)x(k)\hbar_M(k-1)x^H(k) - x^H(k)G(k-1)\mathcal{P}(k)x(k)\vartheta^*(k)|^2}{x^H(k){\Sigma_4}^{-1}x(k)}$$
$$= \|\hbar_M(k-1)\|^2_{\Sigma_4} - \frac{|e^*_a(k)|^2}{x^H(k){\Sigma_4}^{-1}x(k)}$$
$$+ \frac{|e^*_a(k) - x^H(k)G(k-1)\mathcal{P}(k)x(k)e^*_a(k) - x^H(k)G(k-1)\mathcal{P}(k)x(k)\vartheta^*(k)|^2}{x^H(k){\Sigma_4}^{-1}x(k)}$$

$$\|\hbar_M(k)\|^2_{\Sigma_4} + \left(2 - x^H(k){\Sigma_4}^{-1}x(k)\right)|e^*_a(k)|^2$$
$$= \|\hbar_M(k-1)\|^2_{\Sigma_4} + x^H(k){\Sigma_4}^{-1}x(k)|\vartheta(k)|^2 - 2(1 - x^H(k){\Sigma_4}^{-1}x(k)e^*_a(k)\vartheta^*_a(k))$$

$$\|\hbar_M(k)\|^2_{\Sigma_4} + \left(2 - x^H(k){\Sigma_4}^{-1}x(k)\right)|e^*_a(k)|^2 = \|\hbar_M(k-1)\|^2_{\Sigma_4} + x^H(k){\Sigma_4}^{-1}x(k)|\vartheta(k)|^2 - 2(1 - x^H(k){\Sigma_4}^{-1}x(k))e^*_a(k)\vartheta^*_a(k) \quad (41)$$

Taking statistical expectations operation on both side and taking assumption 1by considering input and the system noise are independent. Then it can be written as

$$E[\|\hbar_M(k)\|^2_{\Sigma_4}] + E[\left(2 - x^H(k){\Sigma_4}^{-1}x(k)\right)|e^*_a(k)|^2] = E[\|\hbar_M(k-1)\|^2_{\Sigma_4}] + E[x^H(k){\Sigma_4}^{-1}x(k)|\vartheta(k)|^2] \quad (42)$$

At the steady state stage, i.e. $k \to \infty$, $E[\|\hbar_M(k)\|^2_{\Sigma_4}] = E[\|\hbar_M(k-1)\|^2_{\Sigma_4}]$so the following equation holds

$$\lim_{k\to\infty} E[\left(2 - x^H(k){\Sigma_4}^{-1}x(k)\right)|e^*_a(k)|^2] = \lim_{k\to\infty} E[x^H(k){\Sigma_4}^{-1}x(k)|\vartheta(k)|^2]$$

$$\lim_{k\to\infty} E[\left(2 - x^H(k)G(k-1)\mathcal{P}(k)x(k)\right)|e^*_a(k)|^2] = \lim_{k\to\infty} E[x^H(k)G(k-1)\mathcal{P}(k)x(k)|\vartheta(k)|^2] \quad (43)$$

Given any matrix $W \in R^{nxnx}$, and any random $u \in R^n$ with $E[uu^T] = I$, we have $E[u^T W u] = tr(W)$. (Proof: expand the expectation directly.)

$$\lim_{k\to\infty}[2 - E[tr[x^H(k)G(k-1)\mathcal{P}(k)x(k)]]|e^*_a(k)|^2]$$
$$= \lim_{k\to\infty} E[tr[x^H(k)G(k-1)\mathcal{P}(k)x(k)]|\vartheta(k)|^2]$$
$$(2 - tr[G\mathcal{P}\mathcal{R}]) \lim_{k\to\infty} E[|e^*_a(k)|^2] = \lim_{k\to\infty} E[tr[G\mathcal{P}\mathcal{R}]|\vartheta(k)|^2]$$
$$(2 - tr[G\mathcal{P}\mathcal{R}]) \lim_{k\to\infty} E[|e^*_a(k)|^2] = tr[G\mathcal{P}\mathcal{R}] \lim_{k\to\infty} E[|\vartheta(k)|^2]$$
$$(2 - tr[G\mathcal{P}\mathcal{R}]) \lim_{k\to\infty} E[|\hbar^H_M(k-1)x(k)|^2] = tr[G\mathcal{P}\mathcal{R}] \lim_{k\to\infty} E[|\vartheta(k)|^2]$$

During the steady state, $\mathcal{P} = (1-\lambda)\mathcal{R}^{-1}$ and $\lim_{k\to\infty} E[|\hbar^H_M(k-1)x(k)|^2] = EMSE_{\ SC-PRLS}$

$$EMSE_{\ SC-PRLS} = \frac{(1-\lambda)\ tr[G]\ \sigma^2_\vartheta}{2-(1-\lambda)\ tr[G]} \quad (44)$$

As per the definition trace is defined for square matrix only and it can be calculated as sum of its main diagonal elements. The proportionate matrix G has only diagonal elements. So, It's the trace of proportionate matrix G can be given as

$$tr[G] = \mu$$

$$EMSE_{SC-PRLS} = \frac{(1-\lambda)\ \sigma_{\vartheta}^{2}\mu}{2-(1-\lambda)\mu} \quad (45)$$

Theoretically EMSE will be given by equation (44) that can be verified by simulated results.

6. Simulated results

The sparse system identification model was considered for simulation. The proposed SC-PRLS was implemented for sparse system identification. In SC-PRLS, each filter coefficient converges individually, depending on the gain factor. The gain factors are the elements of the proportionate metrics. The gain distributing factors depends on the degree of sparseness and convergence controlling factor μ. Simulated results verify the proven convergence of SC-PRLS. Convergence controlling factor μ can control the performance of the algorithm. The EMSE is derived for the algorithm, and simulations for the SC-PRLS algorithm verify the theoretical findings. By changing the value of μ from 15 to 90, simulation tests were carried out in order to verify the theoretically calculated EMSE for the SC-PRLS algorithm.

For the SC-PRLS algorithm's theoretical EMSE evaluation, the simulation experiments were conducted by varying the value of μ from 15 to 90. The desired sparse system to be identified is considered as $\omega_M = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1]^T$. The input signal calculated by a function x=awgn(x,SNR,'measured');Where SNR =40, and x is generated by random function and desired signal was generated by d=conv(x, ω_M); d=awgn(d1,SNR,'measured'). Figure 1 shows that algorithm is able to estimate the unknown system very fast. Quickly it is able to track the impulse response of the sparse system to be identified.

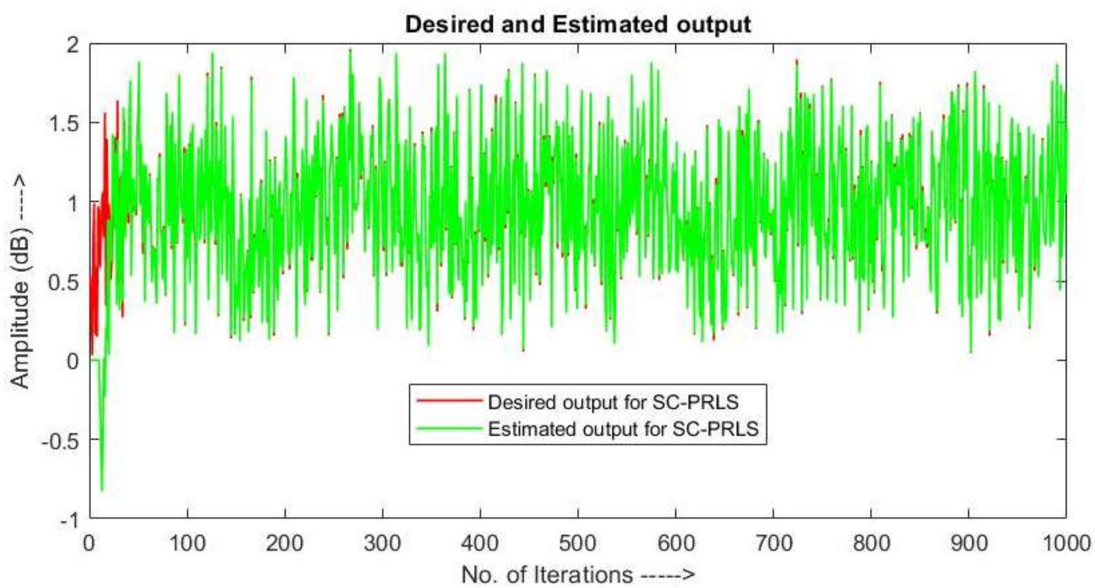

Figure 1. The Impulse response of unknown system and Estimated response by SC-PRLS algorithm.

The SC-PRLS algorithm is fast converging as seen in figure 2, which shown how MSE is converging. The convergence starts at nearly 160 iterations. Steady State Error is about 40 db. Further, the effect of Convergence controlling parameter μ is shown in figure 3, where it can be observed that steady state error is decrease when μ is reduced but after a particular range of μ the rate of convergence starts increasing.

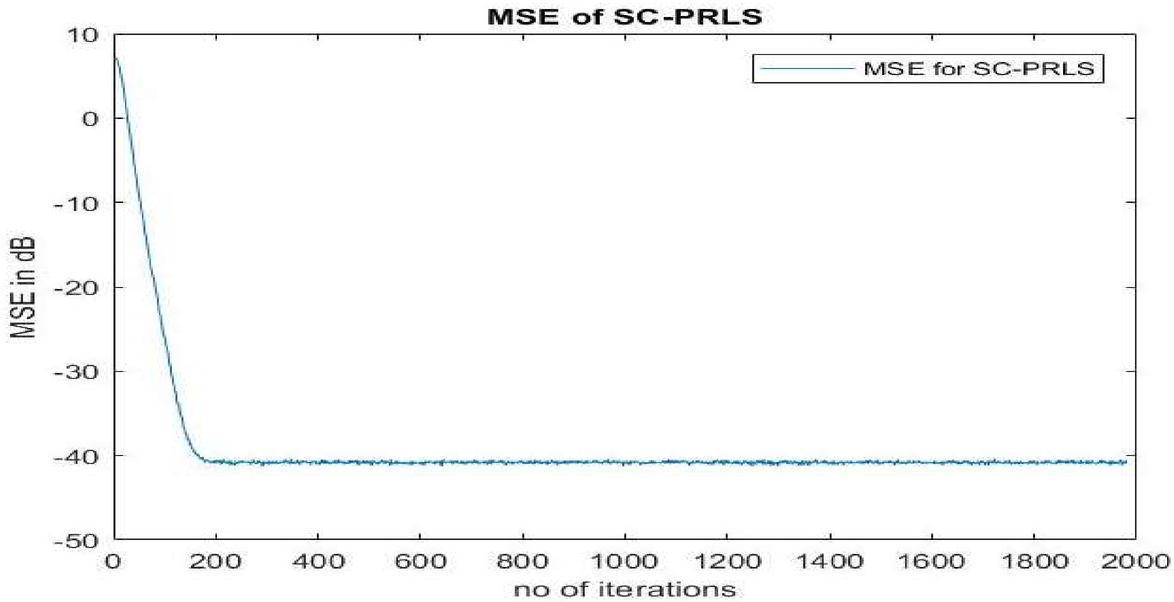

Figure 2. Mean Square Error of SC-PRLS v/s no. of iterations

To select the convergence controlling factor μ there is needs to trade-off between SSE and rate of convergence. By observing figure 3 the optimum range of μ need to select. The range of μ is decided as 25-95, where steady state error is in the range of 35-45 db and convergence starts at 200-400 iterations.

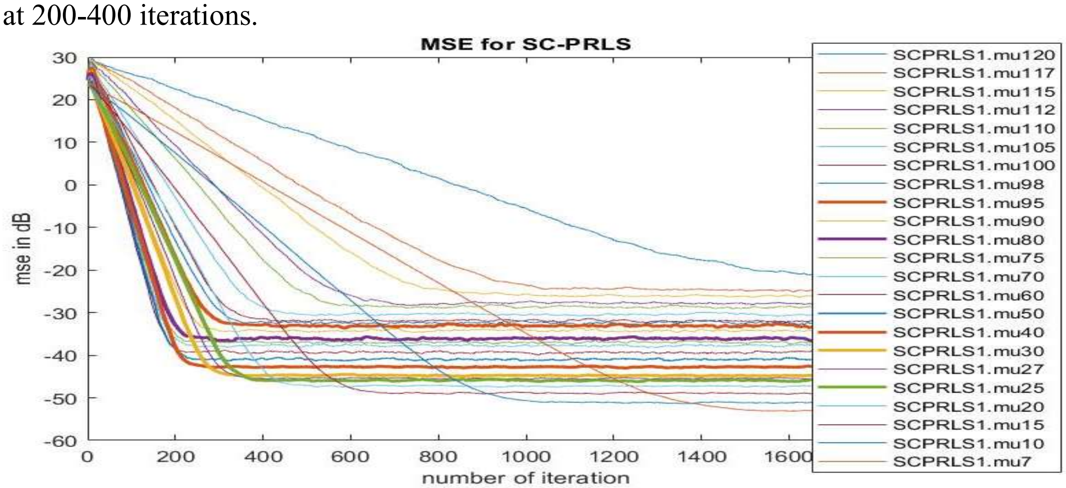

Figure 3. The effect of convergence controlling factor μ on Mean Square Error (MSE) for SC-PRLS

As no.of taps increase in adaptive filter the previously developed PRLS algorithms do not perform well. So observe the limitations simi.ar study is dene for SC-PRLS also and the observations are recorded in table1.

No of filter weights	Value of μ	Convergence starts at	Steady state value in dB

20	50	400	-42
50	50	400	-35
100	50	600	-30
100	40	500	-30
200	38	1000	-20
200	35	800	-20
200	30	750	-22
200	27	800	-22
200	25	1000	-24
200	20	1600	-25
300	30	1000	-18
300	25	1200	-20
300	35	1800	-18
300	32	1200	-18
400	32	1400	-12
400	30	1500	-14
400	25	1300	-15
400	20	2500	-18
500	25	1800	-12
500	27	1200	-12
500	30	2200	-10

Table1: effect of μ on convergence and steady state error when no of filter weights are changed from 20 to 500 by fixing the value of convergence controlling factor at $\mu = 50$.

When the number of filter weights are increased from 20 to 500, the starting point of convergence changed from 400 to 2200 iteration and steady state error changes from -42db to -10db. So, it can be concluded that the SC-PRLS algorithm converges for increased number of filter weights. If convergence controlling factor μ if fixed then performance is decreased for large no of filter taps. Further, the SC-PRLS algorithm performance can be improved by increasing the convergence controlling factor μ when no. of taps increases significantly.

S.No	Value of α_{sc} for SC-PRLS while maintain α_{sc}=0.015 for PRLS	Remark
1	α_{sc}>0.925	Performance of SC-PRLS is worse than RLS
2	α_{sc}=0.925	Performance of SC-PRLS is similar to RLS
3	α_{sc}=0.85	convergence of SC-PRLS is not better than PRLS but Steady state error is near to RLS
4	α_{sc}=0.75	Convergence and steady state error of SC-PRLS is better than PRLS, but steady state error of SC-PRLS is not better than RLS
5	Optimum value $\alpha_{sc} = 0.65$	Convergence and steady state error both are better than PRLS
6	α_{sc}=0.4	Performance of SC-PRLS is similar to PRLS
7	α_{sc}<0.4	performance of SC-PRLS is worse than PRLS

Table 2: performance comparation of SC-PRLS and PRLS for different values of α_{sc}.

Further the performance of SC-PRLS algorithm was compared with PRLS and RLS for the SSI. The summary of the comparison is given in the table 2. The performance of the algorithms depends on α_{sc} and convergence controlling factor μ. The performance changes when the value of α_{sc} changes. For comparison optimum values are chosen for PRLS algorithm and α_{sc} values changes for SC-PRLS.

Optimum value is fund to be $\alpha_{sc} = 0.65$ for SC-PRLS. If α_{sc} is increased from optimal value then performance of SC-PRLS degrades and at $\alpha_{sc} = 0.925$, it performs similar to RLS. While of α_{sc} is decreased below optimum then also performance degrades and at $\alpha_{sc} = 0.4$, it performs similar to PRLS. The convergence controlling factor also affects the performance of PRLS and SC-PRLS, the observations are reported in table 3. This is still topic of detailed discussion as by changing the value convergence controlling factor μ the performance for both the algorithms is changing. But comparison, convergence controlling factor at $\mu = 0.015$ is chosen for PRLS where it was giving good results and for SC-PRLS, μ is varied. It can be observed from the table that if μ<75 the performance of SC-PRLS supersede the PRLS. At μ=75 both the algorithms have similar performance in both SSE and rate of convergence. Below μ<28 rate convergence supersedes for PRLS. While for μ<75 performance of SC-PERLS is better than PRLS. the SSE is improving for smaller values of μ but after some point rater of convergence increases. The optimum value of μ=35 for SC-PRLS. Although, this optimum point changes with number of filter taps.

S.No	Value of μ for SC-PRLS while maintain μ=0.015 for PRLS	Remark
1	μ>75	Performance of PRLS is better than SC-PRLS
2	μ=75	Performance of SC-PRLS is similar to PRLS
3	μ<75	Performance of SC-PRLS is better than PRLS
4	μ=40	Performance of SC-PRLS is better than PRLS
5	μ<28	Convergence of PRLS is better than SC-PRLS but steady state error is better for SC-PRLS
6	Optimum value $\mu = 35$	Convergence and steady state error both are better than PRLS

Table 3: performance comparation of SC-PRLS and PRLS with respect to μ

The performance of SC-PRLS were also compared with PRLS algorithms for SSI in terms of μ and α_{sc} . The optimum values of μ is in found in the rage of 30-75 it depends on number of filter taps. The optimum α_{sc} is 0.65 for the SC-PRLS. Further, theoretically findings also verified by simulation results.

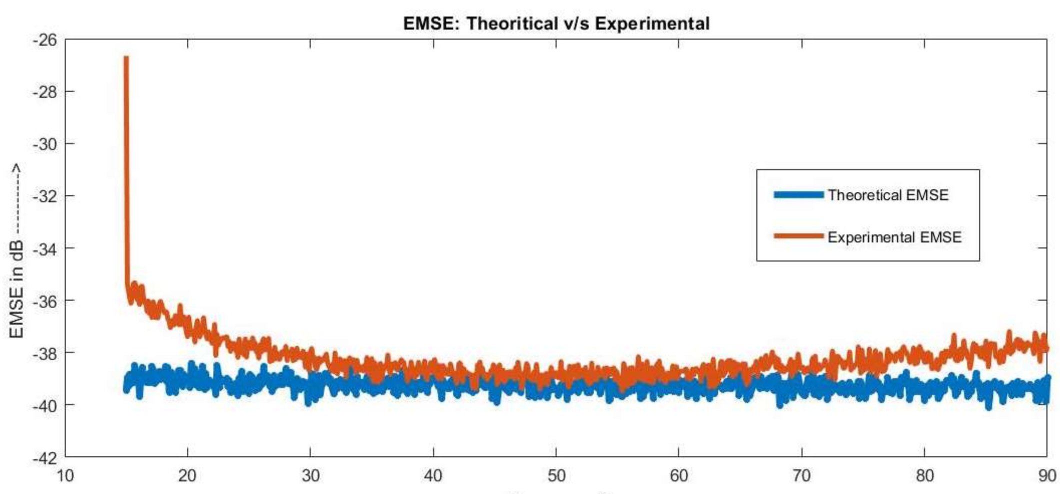

Figure 4. Comparison of theoretical EMSE with Experimental EMSE for different values of μ

The Figure 4 shown the comparison between theoretical findings and practical observations of Excess Mean Square Error (EMSE) for different values of the convergence controlling factor μ. It can be observed from the simulated results that practical performance of the algorithm matches with theoretical findings for a range of $\mu = 30 - 75$ more precisely $\mu = 40 - 60$ is optimum range.

7. Conclusion

The SC-PRLS algorithm is designed to perform better than PRLS and robust performance in variable sparsity. The convergence of the algorithms is also proved in the paper and validated by simulated results. The SSE behavior and transient response of SC-PRLS algorithm is calculated by the principle of energy conservation. Further, the expression for EMSE was also derived and validated by experimental results. For simulation, the algorithm was implemented in sparse systems identification. The algorithm quickly tracks the unknown sparse system's impulse response. The performance of the algorithms depends on the α_{sc} and convergence controlling parameter μ. The optimum value of α_{sc} is 0.65. The performance varies with μ, which needs to change whenever there is a change in the number of filter weights used in the algorithms. For a particular range of convergence controlling parameter μ, the performance of the SC-PRLS algorithm meets the theoretical analysis and performs better than PRLS. This range is decided by comparing the performance in terms of MSE, convergence rate, and EMSE for different values of μ. Therefore, μ should be chosen from (30-70) for better performance. Further, there is future scope to formulate a relationship between the number of filter weights and convergence controlling parameter μ for optimum performance.

References

1. Aurelio Uncini," Fundamentals of Adaptive Signal Processing",springer, 2015.
2. Simon Haykin, "Adaptive Filter Theory", Prentice-hall, 4th edition, 2002.
3. Liu, Jianming, "Adaptive filters for sparse system identification", Doctoral Dissertations, 2483, 2016. https://scholarsmine.mst.edu/doctoral_dissertations/2483
4. Z. Qin, J. Tao and Y. Xia, "A Proportionate Recursive Least Squares Algorithm and Its Performance Analysis," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 1, pp. 506-510, Jan. 2021, doi: 10.1109/TCSII.2020.3004466.
5. Z. Qin, J. Tao, Y. Xia and L. Yang, " Proportionate RLS with L_1 norm regularization: Performance analysis and fast implementation ", Digital Signal Process., vol. 122, pp. 103366, Apr. 2022. https://doi.org/10.1016/j.dsp.2021.103366
6. Y. Wang, Z. Qin, J. Tao and L. Yang, "Performance Analysis of PRLS-based Time-Varying Sparse System Identification," 2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM), Trondheim, Norway, pp. 251-255, 2022, doi: 10.1109/SAM53842.2022.9827876.
7. K. T. Wagner and M. I. Doroslovacki, "Analytical analysis of transient and steady-state properties of the proportionate NLMS algorithm", Proc. Asilomar Conf., pp. 256-260, 2008. DOI: 10.1109/ACSSC.2008.5074403
8. W. Gao, J. Chen, C. Richard, W. Shi and Q. Zhang, "Transient Performance Analysis of the ℓ_1-RLS," in IEEE Signal Processing Letters, vol. 29, pp. 90-94, 2022, doi: 10.1109/LSP.2021.3127467.

9. P. Loganathan, A. W. H. Khong and P. A. Naylor, "A Class of Sparseness-Controlled Algorithms for Echo Cancellation," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 8, pp. 1591-1601, 2009, doi: 10.1109/TASL.2009.2025903.

10. Meenal Mahajan, ranjeetkaur, “A hybrid of SC-PNLMS and SC-MPNLMS adaptive algorithms for acoustic echo cancellation,” International journal of signal processing, 2017.

11. Die Gan, Zhixin Liu, “Performance analysis of the compressed distributed least squares algorithm”, Systems & Control Letters, Volume 164, pp.105228, 2022.

https://doi.org/10.1016/j.sysconle.2022.105228

12. P. S. R. Diniz, Adaptive Filtering: Algorithms and Practical Implementation, New York, NY, USA:Springer, pp. 209-246, 2013.

13. R. L. Das and M. Chakraborty, "On convergence of proportionate-type normalized least mean square algorithms", IEEE Trans. Circuits Syst. II Exp. Briefs, vol. 62, no. 5, pp. 491-495, Dec. 2015.

14. A. Tedjani and A. Benallal, "Performance study of three different sparse adaptive filtering algorithms for echo cancellation in long acoustic impulse responses," 2017, 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B), Boumerdes, Algeria, pp. 1-7, 2017, doi: 10.1109/ICEE-B.2017.8192212.

15. Zhen Qin, Jun Tao, Le Yang, Ming Jiang, “Proportionate recursive maximum correntropy criterion adaptive filtering algorithms and their performance analysis”, Digital Signal Processing, Volume 139, 2023.

https://doi.org/10.1016/j.dsp.2023.104073

16. B. Widrow and S. D. Stearns, Adaptive Signal Processing,Prentice Hall, 1985.

17. X. Hong, J. Gao and S. Chen, "Zero-attracting recursive least squares algorithms", IEEE Trans. Veh. Technol., vol. 66, no. 1, pp. 213-221, Jan. 2017.

18. Patrik O. Hoyer, “Non-negative Matrix Factorization with Sparseness Constraints”, The Journal of Machine Learning Research, Volume 5, pp 1457–1469, 2004.

https://doi.org/10.48550/arXiv.cs/0408058

19. Richard K. Martin, William A. Sethares, Robert C. Williamson and Richard Johanson, “Exploiting Sparsity in Adaptive Filters,” IEEE Transactions on signal processing, vol. 50, no.8, pp 1883-1894, 2002.