L1-Norm Penalized Bias Compensated Linear Constrained Affine Projection Algorithm

Rajni Yadav, Chandra Shekhar Rai

Abstract*: *This paper presents an l1-norm penalized bias compensated linear constrained affine projection (l1-BC-CAP) algorithm for sparse system identification having linear phase aspectin the presence of noisy colored input. The motivation behind the development of the proposed algorithm is formulated on the concept of reusing the previous projections of input signal in affine projection algorithm (APA) that makes it suitable for colored input. At First, l1-CAP algorithm is derived by adding zero attraction based on l1-norm into constrained affine projection (CAP) algorithm. Then, the proposed l1-BC-CAP algorithm is derived by addinga bias compensator into the filter coefficient update equation of l1-norm constrained affine projection (l1-CAP) algorithm to alleviate the adverse consequence of input noise on the estimation performance. Hence, the resulting l1-BC-CAP algorithm excels the estimation performance when applied to linear phase sparse system in the existence of noisy colored input. Further, this work also examines the stability concept of the proposed algorithm

Keywords*: *Affine projection, bias compensator, linear constraint, sparsity.

I. INTRODUCTION

Use of linear constrained adaptive filtering in many digital signal processing applications has been on a steady rise owing to their utility in considering the prior knowledge about the framework to be estimated. The estimation of the frameworkrelays on some linear constraints which are available in advance. Some examples of linear constrained adaptive filtering are adaptive beam forming, linear phase system identification, code division multiple access and many more. [1-3].These applications have powered deep interest in developing linear constrained adaptive filters. The constrained least mean squares (CLMS) algorithm has gained a lot of attention in linear constrained adaptive filtering due to ease and simplicity in implementation [4]. But CLMS algorithm has poor performance in the presence of colored inputs. Moreover, the constrained affine projection algorithm (CAPA) is developed to consider the colored input [5]. As CAPA reusesthe previous projections of input signal, hence it has better performance for colored input. However, CAP algorithm does not take into account the sparsity of the system. Later several sparsity aware affine projection algorithms have been developed to consider the sparsity of the system [12-13]. These algorithms append a zero attraction in conventional affine projection algorithms. These algorithms do not consider linear constraint of applications in development. This paper first develops l1-norm constrained affine projection (l1-CAP) algorithm that appends the zero attraction based on l1-norm to consider the sparsity of the system. The above mentioned algorithm performs well for constrained applications in the presence of noiseless colored input. However, the performance of l1-CAP algorithm is deteriorated in the presence of input noise. Moreover, the input noise adds a bias in numerator as well as in denominator of l1-CAP algorithm. Also it is to be noted that in case of l1-CAP algorithm, both the estimation performance and the stability of the algorithm are influenced by input noise. However, the deterioration of the performance of the proposed algorithm is caused largely by the bias in the numerator [10]. To solve the problem of input noise, bias compensation criterion has been developed [6-10]. Some of the bias compensation criterion based adaptive algorithms proposed in past are: bias compensated normalized least mean square (BC-NLMS), bias-compensated robust set-membership NLMS (BC-SM-NLMS), bias-compensated normalized sub-band adaptive filter (BC-NSAF), bias compensated affine projection like (APL), bias compensated affine projection (APA) [6-10].These algorithms add a bias compensator to make the estimation unbiased for noisy input. However, these algorithms do not deal with sparsity and linear constraint of the system simultaneously. Based on the above concept of bias compensation, this paper presents l1-norm penalized bias compensated constrained affine projection (l1-BC-CAP) algorithm that takes into account the input noise. The proposed work adds a bias compensator into update equation of l1-CAP algorithm to mitigate the unfavorable impact of input noise on the estimation performance in constrained applications against colored input. The rest of the paper is divided as follows. In section II, the l1-norm penalized constrained affine projection (l1-CAP) algorithm is derived.

Revised Manuscript Received on February 15, 2020.
* Correspondence Author
Rajni Yadav*, Department of Electronics and Communication, Maharaja Agrasen Institute of Technology, Guru Gobind Singh Indraprastha University, Delhi, India,. Email: rajni@mait.ac.in
Chandra Shekhar Rai,, University School of Information, &Communication Technology, Guru Gobind Singh Indraprastha University, Delhi, India,. Email: csrai@ipu.ac.in.

Section IIIdemonstrates the derivation of l1-BC-CAP algorithm. Section IVillustrates the convergence behavior of the proposed algorithm. Simulations and results are discussed in Section V, and hence it is concluded in Section VI.

II. L1-NORM PENALIZED LINEAR CONSTRAINED AFFINE PROJECTION (L1-CAP) ALGORITHM

This section derives l1-norm penalized linear constrained affine projection (l1-CAP) algorithm for linear constrained filtering problem.

Consider the desired output d(k) of an unknown systemas $d(k) = \mathbf{w}_0^T\mathbf{x}(k) + \mathbf{z}(k)$,where $\boldsymbol{w}_0$ is unknown system coefficients vector of dimension NX1, $\mathbf{x}(k) \in \mathbb{R}^{NX1}$is the noise free input vector and z(k) is the observation noise of channel.

Considering$\mathbf{w}(k)$ as the adaptive filter coefficient vector of length N, the cost function of l1-norm penalizedlinear constrained affine projection (l1-CAP) algorithm can be drafted as:

$$\mathbf{w}(k+1) = \arg\min \|\mathbf{w} - \mathbf{w}(k)\|^2 + \beta\|\mathbf{w}\|_1$$

Subject to $\mathbf{d}(k) - \mathbf{X}^T(k)\mathbf{w} = \mathbf{0}$, and $\boldsymbol{\Theta}^T\mathbf{w} = \mathbf{h}$(1)where $\mathbf{X}(k) \in \mathbb{R}^{NXL}$ is input matrix consist of previous L-1 projection and current input vector

$\mathbf{d}(k)$is desired output vector of dimension Lx1consisting of previous L-1 and current output of the unknown system

andL is the projection order of l1-norm penalizedconstrained affine projection (l1-CAP) algorithm; β is sparsity regularizer.

The parameter $\boldsymbol{\Theta}$represents **N**x**P** constraint matrix while **f** represents a vector which comprises of the **P** constrained output values.

With the help of Langrage multiplier approach, the unconstrained cost function of l1-CAP algorithm can be drafted as:

$$J(\mathbf{w}) = \|\mathbf{w} - \mathbf{w}(k)\|^2 + \beta\|\mathbf{w}\|_1 + \Lambda_1^T(\boldsymbol{\Theta}^T\mathbf{w} - \mathbf{h}) + \Lambda_2^T(\mathbf{d}(k) - \mathbf{X}^T(k)\,\mathbf{w}) \quad (2)$$

where

Λ_1 and Λ_2 are Lagrange multipliers, and β is sparsity regularizer.

Taking the gradient of cost function in (2), we have

$$g(w) = \frac{\partial J(\mathbf{w})}{\partial(\mathbf{w})} = 2(\mathbf{w} - \mathbf{w}(k)) + \beta sign(\mathbf{w}) + \boldsymbol{\Theta}\,\Lambda_1 - \mathbf{X}(k)\Lambda_2 \quad (3)$$

Setting derivate equal to zero in (3), the coefficient recursive equationof l1-CAP algorithm can be written as:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \frac{\beta}{2}sign(\mathbf{w}(\mathbf{k})) - \frac{\boldsymbol{\Theta}}{2}\Lambda_1 + \frac{\mathbf{X}(k)}{2}\Lambda_2 \quad (4)$$

The value of Langrage multiplier $\Lambda_1$can be computed as:

$$\boldsymbol{\Theta}^T(\mathbf{w}(k)) - \mathbf{h} - \frac{\beta\boldsymbol{\Theta}^T}{2}sign(\mathbf{w}(k)) - \frac{\boldsymbol{\Theta}^T\boldsymbol{\Theta}}{2}\Lambda_1 + \frac{\boldsymbol{\Theta}^T\mathbf{X}(k)}{2}\Lambda_2 = \mathbf{0} \quad (5)$$

$$\Lambda_1 = [\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T\mathbf{X}(k)\Lambda_2 - \beta[\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T sign(\mathbf{w}(k)) \quad (6)$$

Now the value of $\Lambda_2$is calculated as:

$$\mathbf{d}(k) - \mathbf{X}^T(k)\Big[\mathbf{w}(k) - \frac{\beta}{2}sign(\mathbf{w}(\mathbf{k})) - \frac{\boldsymbol{\Theta}}{2}\{[\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T\mathbf{X}(k)\Lambda_2 - \beta[\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T sign(\mathbf{w}(\mathbf{k}))\}\Big] + \frac{\mathbf{X}(k)}{2}\Lambda_2)) = 0 \quad (7)$$

Taking into account the error signal$e(k) = \mathbf{x}^T(k)\mathbf{w}_0 + z(k) - \mathbf{x}^T(k)\mathbf{w}(k)$ and error vector $\mathbf{e}(k) = \mathbf{X}^T(k)\mathbf{w}_0 + \mathbf{z}(k) - \mathbf{X}^T(k)\mathbf{w}(k)$, we can find $\Lambda_2$as

$$\Lambda_2 = \mathbf{2}(\mathbf{X}^T(k)P\mathbf{X}(k))^{-1}\left[\mathbf{e}(k) + \frac{\beta}{2}\mathbf{X}^T(k)\mathbf{Q}sign(\mathbf{w}(k))\right] \quad (8)$$

Hence

$$\Lambda_1 = \mathbf{2}[\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T\mathbf{X}(k)(\mathbf{X}^T(k)\mathbf{Q}\mathbf{X}(k))^{-1}\mathbf{e}(k) + \beta\mathbf{X}^T(k)(\mathbf{X}^T(k)\mathbf{Q}\,\mathbf{X}(k))^{-1}[\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T\mathbf{X}^T(k)\mathbf{Q}sign(\mathbf{w}(k)) - \beta[\boldsymbol{\Theta}^T\boldsymbol{\Theta}]^{-1}\boldsymbol{\Theta}^T sign(\mathbf{w}(k)) \quad (9)$$

Therefore; the coefficient recursive equation of l1-CAP algorithm becomes:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \frac{\beta}{2}\left[\mathbf{I} - \{\mathbf{X}(k)(\mathbf{X}^T(k)\mathbf{Q}\mathbf{X}(k))^{-1}\mathbf{Q}\mathbf{X}^T(k)\}\right]\mathbf{Q}sign(\mathbf{w}(k)) + \mathbf{X}(k)(\mathbf{X}^T(k)\mathbf{Q}\,\mathbf{X}(k))^{-1}\mathbf{Q}\mathbf{e}(k) + \frac{\beta}{2}\{\mathbf{X}(k)(\mathbf{X}^T(k)\mathbf{Q}\mathbf{X}(k))^{-1}\mathbf{Q}\mathbf{X}^T(k)\mathbf{Q}sign(\mathbf{w}(k))\} \quad (10)$$

where

$$\mathbf{Q} = \mathbf{I} - \boldsymbol{\Theta}(\boldsymbol{\Theta}^T\boldsymbol{\Theta})^{-1}\boldsymbol{\Theta}^T \quad (11)$$

$$\mathbf{H} = \boldsymbol{\Theta}(\boldsymbol{\Theta}^T\boldsymbol{\Theta})^{-1}\mathbf{h} \quad (12)$$

By considering,

$$\mathbf{I} - \{\mathbf{X}(k)(\mathbf{X}^T(k)\mathbf{Q}\mathbf{X}(k))^{-1}\mathbf{Q}\mathbf{X}^T(k)\} = \mathbf{I} \quad (13)$$

Consider 0<μ<1 be the step size for the stability of algorithm [11,14].

Therefore, the coefficient recursive equation of l1-norm penalized linear constrained affine projection (l1-CAP) algorithm becomes:

$$\mathbf{w}(k+1) = \mathbf{Q}\left[\mathbf{w}(k) - \frac{\mu\beta}{2}sign(\mathbf{w}(k)) + \mu\mathbf{X}(k)(\mathbf{X}^T(k)\mathbf{Q}\mathbf{X}(k))^{-1}\mathbf{e}(k)\right] + \mathbf{H} \quad (14)$$

III. L1-NORM PENALIZED BIAS COMPENSATED LINEAR CONSTRAINED AFFINE PROJECTION ALGORITHM

This section will consider noisy input,$\tilde{\mathbf{x}}(k) = \mathbf{x}(k) + \mathbf{v}(k)$, where $\mathbf{v}(k)$is input noise with variance $\sigma_v^2$and zero mean.Fig. 1 shows the system identification problem in the presence of input noise.

Retrieval Number: C4815029320/2020©BEIESP
DOI: 10.35940/ijeat.C4815.029320
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Hence, the noisy input matrix, $\tilde{\mathbf{X}}(k) = \mathbf{X}(k) + \mathbf{V}(k)$ consists of previous (L-1) projection of input signal and input noise. The input noise matrix, $\mathbf{V}(k) = [\,\mathbf{v}(k), \mathbf{v}(k-1), \dots, \mathbf{v}(k-L+1)]$, a priori error $\tilde{e}(k) = \mathbf{x}^T(k)\mathbf{w}_0 + \mathbf{z}(k) - \tilde{\mathbf{x}}^T(k)\mathbf{w}(k) = e(k) - \mathbf{v}^T(k)\mathbf{w}(k)$, a priori error vector $\tilde{\mathbf{e}}(k) = [\tilde{e}(k)\tilde{e}(k-1)\ \dots\ \tilde{e}(k-L+1)]^T$ are considered for derivation. The equation (14) can be rewritten for noisy input as:

$$\mathbf{w}(k+1) = \mathbf{Q}\left[\mathbf{w}(k) - \frac{\mu\beta}{2} sign(\mathbf{w}(k)) + \mu\tilde{\mathbf{X}}(k)\left(\tilde{\mathbf{X}}^T(k)\mathbf{Q}\tilde{\mathbf{X}}(k)\right)^{-1}\tilde{\mathbf{e}}(k)\right] + \mathbf{H} \quad (15)$$

$$\mathbf{w}(k+1) = \mathbf{Q}\left[\mathbf{w}(k) - \frac{\mu\beta}{2} sign(\mathbf{w}(k)) + \mu\mathbf{X}(k)\left(\tilde{\mathbf{X}}^T(k)\mathbf{Q}\tilde{\mathbf{X}}(k)\right)^{-1}\mathbf{e}(k)\right] + \mu\mathbf{V}(k)\mathrm{Q}\left(\tilde{\mathbf{X}}^T(k)\mathbf{Q}\tilde{\mathbf{X}}(k)\right)^{-1}\mathbf{e}(k) + \mathbf{H} \quad (16)$$

The extra term in (16) shows the bias in the numerator as well as in the denominator. This will hamper both the estimation performance and stability aspect of the algorithm. However, the deterioration of the performance is caused largely by the adverse impact on the numerator term. Hence, the proposed work compensates the bias in numerator to excel the estimation performance in the presence of input noise.

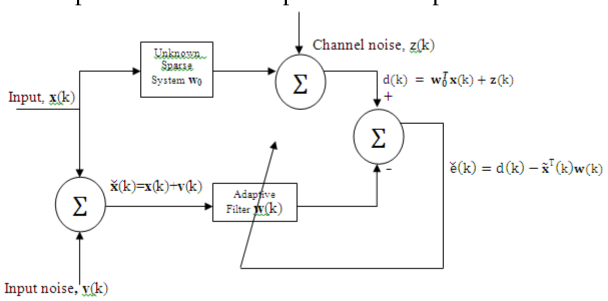

Fig. 1. Adaptive System Identification in the presence of input noise

Hence, to overcome the effect of bias generated by noisy input, a bias compensator $\mathbf{D}(k)$ is added in the above weight update equation which is given as:

$$\mathbf{w}(k+1) = \mathbf{Q}\left[\mathbf{w}(k) - \frac{\mu\beta}{2} sign(\mathbf{w}(k)) + \frac{\mu[\tilde{\mathbf{X}}(k)(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}\right] + \mathbf{H} + \mathbf{D}(k) \quad (17)$$

In order to find the value of bias compensator $\mathbf{D}(k)$, the sparsity of the system is not taken into account.

Hence, the weight update equation (17) becomes:

$$\mathbf{w}(k+1) = \mathbf{Q}\left[\mathbf{w}(k) + \frac{\mu[\tilde{\mathbf{X}}(k)(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}\right] + \mathbf{H} + \mathbf{D}(k) \quad (18)$$

Defining weight misalignment vector as:

$$\Delta\mathbf{w}(k+1) = \mathbf{w}(k+1) - \mathbf{w}_{opt} \quad (19)$$

where $\mathbf{w}_{opt}$ is optimum weight vector.

Hence,

$$\Delta\mathbf{w}(k+1) = \mathbf{Q}\Delta\mathbf{w}(k) + \frac{\mu[\tilde{\mathbf{X}}(k)\mathrm{Q}(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)} + \mathbf{H} + \mathbf{D}(k) + \mathbf{Q}\mathbf{w}_{opt} - \mathbf{w}_{opt} \quad (20)$$

Using $\mathbf{Q}\mathbf{w}_{opt} - \mathbf{w}_{opt} + \mathrm{H} = (\mathbf{I} - \boldsymbol{\Theta}(\boldsymbol{\Theta}^T\boldsymbol{\Theta})^{-1}\boldsymbol{\Theta}^T)\mathbf{w}_{opt} - \mathbf{w}_{opt} + \boldsymbol{\Theta}(\boldsymbol{\Theta}^T\boldsymbol{\Theta})^{-1}\boldsymbol{\Theta}^T\mathbf{w}_{opt} = \mathbf{0}$, we can write (20) as:

$$\Delta\mathbf{w}(k+1) = \mathbf{Q}\Delta\mathbf{w}(k) + \frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)} + \mathbf{D}(k) \quad (21)$$

Taking expectation of (21) on both sides while considering the availability of the matrix $\tilde{\mathbf{X}}(k)$, the recursion equation of the weight-misalignment vector becomes:

$$\mathrm{E}[\Delta\mathbf{w}(k+1)|\tilde{\mathbf{X}}(k)] = \mathbf{Q}\,\mathrm{E}[\Delta\mathbf{w}(k)|\tilde{\mathbf{X}}(k)] + \mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] + \mathrm{E}\left[\mathbf{D}(k)|\tilde{\mathbf{X}}(k)\right] \quad (22)$$

To achieve unbiased estimation, the criterion [10] is applied as:

$$\mathrm{E}[\Delta\mathbf{w}(k+1)|\tilde{\mathbf{X}}(k)] = \mathrm{E}[\Delta\mathbf{w}(k)|\tilde{\mathbf{X}}(k)] = 0 \quad (23)$$

Hence

$$\mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] = -\mathrm{E}\left[\mathbf{D}(k)|\tilde{\mathbf{X}}(k)\right] \quad (24)$$

and

$$\mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] = \mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\mathbf{e}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] - \mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\mathbf{v}^T(k)\mathbf{w}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] \quad (25)$$

The bias compensator is derived on basis of given below assumptions:

Assumption 1: Input noise $\mathbf{v}(k)$, input signal $\mathbf{x}(k)$ and measurement noise z(k) are considered to be white Gaussian process having zero mean and variance σ_v^2, σ_x^2 and σ_z^2.respectively.

Assumption 2: The signals $\mathbf{v}(k)$, z(k) and $\mathbf{x}(k)$ and $\boldsymbol{w}(k)$ are statistically independent.

Considering the above assumptions, the first term on right side of (25) reduces to,

$$\mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\mathbf{e}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] = \mathrm{E}\left[\frac{\mu[(\mathbf{X}(k)+\mathbf{V}(k))\mathrm{Q}(-\mathbf{X}^T(k)\Delta\mathbf{w}(k+1)+\mathbf{z}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] = 0 \quad (26)$$

And the second term on right side of (25) leads to:

$$\mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\mathbf{v}^T(k)\mathbf{w}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] = \mathrm{E}\left[\frac{\mu[\mathbf{X}(k)\mathbf{Q}(\mathbf{v}^T(k)\mathbf{w}(k))+\mathbf{V}(k)\mathbf{Q}(\mathbf{v}^T(k)\mathbf{w}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] \quad (27)$$

Using the above assumptions,

$$\mathrm{E}\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Q}(\mathbf{v}^T(k)\mathbf{w}(k))]}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] = \mathrm{E}\left[\frac{\mu L\sigma_v^2\mathbf{Q}\,\mathbf{w}(k)}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] \quad (28)$$

Therefore,

$$\mathrm{E}\left[\mathbf{D}(k)|\tilde{\mathbf{X}}(k)\right] = \mathrm{E}\left[\frac{\mu L\sigma_v^2\mathbf{Q}\,\mathbf{w}(k)}{\tilde{\mathbf{X}}^T(k)\,\mathbf{Q}\,\tilde{\mathbf{X}}(k)}|\tilde{\mathbf{X}}(k)\right] \quad (29)$$

Retrieval Number: C4815029320/2020©BEIESP
DOI: 10.35940/ijeat.C4815.029320
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$\mathbf{D}(\mathrm{k}) = \frac{\mu L\, \sigma_v^2\, \mathbf{Q}\mathbf{w}(k)}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}$ (30)

Hence, the coefficient recursion equation of the proposed l1-norm penalized bias compensator constrained affine projection algorithm becomes:

$$\mathbf{w}(k+1) = \mathbf{Q}\left[\mathbf{w}(k) - \frac{\mu\beta}{2} sign(\mathbf{w}(k)) + \frac{\mu[\tilde{\mathbf{X}}(k)(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + \mathbf{H} + \frac{\mu L\, \sigma_v^2\, \mathbf{Q}\mathbf{w}(k)}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)} \quad (31)$$

Since (31) requires variance σ_v^2 of the input noise**,** an estimation of the same should be calculated as it is not available in practice. In this paper, the method of estimation of σ_v^2proposed by Haiquan Zhao and ZongshengZheng [9] is used.

Consider measurement noise free error $\tilde{e}_{nf}(k)$ as:

$\tilde{e}_{\mathbf{nf}}(k) = \tilde{e}(k) - z(k)$ (32)

$\tilde{e}_{\mathbf{nf}}(k) = -\Delta\mathbf{w}^T(k)\mathbf{x}(k) + \mathbf{w}^T(k)v(k)$ (33)

where

$\Delta\mathbf{w} = \mathbf{w}(k) - \mathbf{w}_0$ (34)

Taking the expectation of square of (33) and considering the above assumptions, we have

$\sigma_{\tilde{e}_{\mathbf{nf}}}^2 = \sigma_v^2(k)\, E[\mathbf{w}^T(k)\mathbf{w}(k)]$ (35) $\sigma_v^2(k) = \frac{\sigma_{\tilde{e}_{\mathbf{nf}}}^2(k)}{E[\mathbf{w}^T(k)\mathbf{w}(k)]}$

$\sigma_v^2(k) = \frac{\sigma_{\tilde{e}_{\mathbf{nf}}}^2(k)}{\sigma_{\mathbf{w}}^2(k)}$ (36) Hence, an estimate $\hat{\sigma}_v^2(k)$ can be written as:

$\hat{\sigma}_v^2(k) = \frac{\sigma_{\tilde{e}_{\mathbf{nf}}}^2(k)}{\sigma_{\mathbf{w}}^2(k)}$ (37)

where

$\sigma_{\tilde{e}_{\mathbf{nf}}}^2(k) = a\, \sigma_{\tilde{e}_{\mathbf{nf}}}^2(k-1) + (1-a)\, \tilde{\mathbf{e}}_{\mathbf{nf}}^2(k)$

(38)$\sigma_{\mathbf{w}}^2(k) = b\, \sigma_{\mathbf{w}}^2(k-1) + (1-b)[\mathbf{w}^T(k)\mathbf{w}(k)]$ (39)

and parameters a and b are close to unity.

IV. CONVERGENCE ANALYSIS

For the convergence analysis, we are considering the jointly Gaussian distribution of any two elements of weight misalignment vector, $\Delta\boldsymbol{w}(k)$.

Let $\Delta w_i(k)$ and $\Delta w_j(k)$ be the two elements of $\Delta\boldsymbol{w}(k)$. Hence we can define the jointly Gaussian distribution as:

$$(\Delta w_i(k), \Delta w_j(k)) \sim N(\mu_i, \mu_j, \sigma_i^2, \sigma_j^2, \rho_{i,j})$$

where

$\mu_i = E[\Delta w_i(k)]$ (40)

$\mu_j = E[\Delta w_j(k)]$ (41)

$\sigma_i^2 = E[[\Delta w_i^2(k)] - E[\Delta w_i(k)]^2$ (42)

$\sigma_j^2 = E[[\Delta w_j^2(k)] - E[\Delta w_j(k)]^2$ (43)

$\rho_{i,j} = E[\Delta w_i(k)\Delta w_j(k)] - E[\Delta w_i(k)]\, E[\Delta w_j(k)]$ (44)

The optimum weight vector $\mathbf{w}_{opt}$for the constrained APA can be defined as:

$\boldsymbol{w_{opt}} = \boldsymbol{w_0} + \boldsymbol{R^{-1}}\boldsymbol{\Theta}\, (\boldsymbol{\Theta}^T \boldsymbol{R^{-1}}\, \boldsymbol{\Theta})^{-1}(\boldsymbol{h} - \boldsymbol{\Theta}^T \boldsymbol{R^{-1}}\, \boldsymbol{P})$ (45)

where

$\boldsymbol{w_0} = \boldsymbol{R^{-1}P}$ (46)

$\boldsymbol{P} = \boldsymbol{E}[\mathbf{x}(k)\boldsymbol{d}(k)]$and (47)

$\boldsymbol{R} = \boldsymbol{E}[\mathbf{x}(k)\mathbf{x}^T(k)]$(48)

Subtracting$\mathbf{w}_{opt}$ both sides from (31), we have

$$\Delta\boldsymbol{w}(k+1) = \mathbf{Q}\left[\Delta\boldsymbol{w}(k) - \frac{\mu\beta}{2} sign(\mathbf{w}(\mathbf{k})) + \frac{\mu[\tilde{\mathbf{X}}(k)(\tilde{\mathbf{e}}(k))]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + \frac{\mu L\, \sigma_v^2\, \mathbf{Q}\mathbf{w}(k)}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)} \quad (49)$$

$\tilde{\mathbf{e}}(k) = \mathbf{X}^T(k)\mathbf{w}_{opt} + \mathbf{Z}(k) - \tilde{\mathbf{X}}^T(k)\mathbf{w}(k) = (\tilde{\mathbf{X}}^T(k) - \mathbf{V}^T(k))\mathbf{w}_{opt} + \mathbf{Z}(k) - \tilde{\mathbf{X}}^T(k)\mathbf{w}(k) = -\tilde{\mathbf{X}}^T(k)\tilde{\boldsymbol{w}}(k) - \mathbf{V}^T(k)\mathbf{w}_{opt} + \mathbf{Z}(k)$ (50)

Therefore,

$$\Delta\boldsymbol{w}(k+1) = \mathbf{Q}\left[[\Delta\boldsymbol{w}(k)] - \frac{\mu\beta}{2}[sign(\mathbf{w}(k))] - \left[\frac{\mu[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\Delta\boldsymbol{w}(k)]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] - \left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{V}^T(k)\mathbf{w}_{opt}]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + \left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Z}(k)]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right]\right] + \left[\frac{\mu L\, \sigma_v^2\, \mathbf{Q}\Delta\boldsymbol{w}(\boldsymbol{k})}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + \left[\frac{\mu L\, \sigma_v^2\, \mathbf{Q}\mathbf{w}_{opt}}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] \quad (51)$$

Taking Expectation of (51) yields

$$E[\Delta\boldsymbol{w}(k+1)] = \mathbf{Q}\left[E[\Delta\boldsymbol{w}(k)] - \frac{\mu\beta}{2}E[sign(\mathbf{w}(k))] - E\left[\frac{\mu[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\Delta\boldsymbol{w}(k)]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] - E\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{V}^T(k)\mathbf{w}_{opt}]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + E\left[\frac{\mu[\tilde{\mathbf{X}}(k)\mathbf{Z}(k)]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right]\right] + E\left[\frac{\mu L\, \sigma_v^2\, \mathbf{Q}\Delta\boldsymbol{w}(\boldsymbol{k})}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + E\left[\frac{\mu L\, \sigma_v^2\, \mathbf{Q}\mathbf{w}_{opt}}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] \quad (52)$$

Considering the above assumptions, (52) reduces to

$$E[\Delta\boldsymbol{w}(k+1)] = \mathbf{Q}\left[E[\Delta\boldsymbol{w}(k)] - \frac{\mu\beta}{2}E[sign(\mathbf{w}(k))] - \mu E\left[\frac{[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] E[\Delta\boldsymbol{w}(k)]\right] + \mu L\, \sigma_v^2\, \mathbf{Q}E\left[\frac{\Delta\boldsymbol{w}(k)}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] \quad (53)$$

We can find the value of $E[sign(\mathbf{w}(k))]$as [15]:

$E[sign(w_i(k))] = E[sign(\Delta w_i(k) + w_{opt,i})] = 1 - \varphi(-\frac{\mu_{total}(k)}{\sigma_i(k)})$ (54)

where$\mu_{total}(k) = \mu_i + w_{opt,i}$ (55)

and

$\varphi(x)$is CDF of Normal distribution.

As $E[sign(\mathbf{w}(k))]$ is bounded, hence we are not considering this term in the stability analysis of the proposed algorithm.

Hence

$$E[\Delta\mathbf{w}(k+1)] = \mathbf{Q}\left[I - \mu E\left[\frac{[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)]}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + \mu L\, \sigma_v^2 E\left[\frac{I}{\tilde{\mathbf{X}}^T(k)\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right]\right] E[\Delta\mathbf{w}(k)] \quad (56)$$

The square of symmetric matrix Q is also Q i.e.

$\boldsymbol{Q^2} = \boldsymbol{Q}$ (57)

and

$\boldsymbol{Q}\Delta\boldsymbol{w}(k+1) = \Delta\boldsymbol{w}(k+1)$ (58)

Considering (57) and (58), we can write (56) as:

$$E[\Delta\boldsymbol{w}(k+1)] = \left[I - \mu E\left[\frac{[\boldsymbol{Q}\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\boldsymbol{Q}^T]}{\tilde{\mathbf{X}}^T(k)\boldsymbol{Q}^T\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right] + \mu L\, \sigma_v^2 E\left[\frac{I}{\tilde{\mathbf{X}}^T(k)\boldsymbol{Q}^T\, \mathbf{Q}\, \tilde{\mathbf{X}}(k)}\right]\right] E[\Delta\boldsymbol{w}(k)] \quad (59)$$

Considering

$\boldsymbol{Q}\breve{\mathbf{X}}(\mathrm{k}) = \breve{\mathbf{X}}(\mathrm{k})$ (60)

Rewriting (60) using (59) yields

$$E[\Delta\boldsymbol{w}(\mathrm{k}+1)] = \left[I - \mu\mathrm{E}\left[\frac{[\breve{\mathbf{X}}(\mathrm{k})\breve{\mathbf{X}}^T(\mathrm{k})]}{\breve{\mathbf{X}}^T(\mathrm{k})\breve{\mathbf{X}}(\mathrm{k})}\right] + \mu\mathrm{L}\,\sigma_{\mathrm{v}}^2 E\left[\frac{I}{\breve{\mathbf{X}}^T(\mathrm{k})\breve{\mathbf{X}}(\mathrm{k})}\right]\right] E[\Delta\boldsymbol{w}(\mathrm{k})] \quad (61)$$

$$E[\Delta\boldsymbol{w}(\mathrm{k}+1)] = \left[I - \mu\mathrm{E}\left[\frac{[\breve{\mathbf{X}}(\mathrm{k})\breve{\mathbf{X}}^T(\mathrm{k})]}{\breve{\mathbf{X}}^T(\mathrm{k})\breve{\mathbf{X}}(\mathrm{k})}\right] + \mu\mathrm{L}\,\sigma_{\mathrm{v}}^2 \frac{I}{E[\breve{\mathbf{X}}^T(\mathrm{k})\breve{\mathbf{X}}(\mathrm{k})]}\right] E[\Delta\boldsymbol{w}(\mathrm{k})] \quad (62)$$

Considering

$$\mathrm{F}(\mathrm{k}) = \left[I - \mu\left[\frac{[\breve{\mathbf{X}}(\mathrm{k})\breve{\mathbf{X}}^T(\mathrm{k})]}{\breve{\mathbf{X}}^T(\mathrm{k})\breve{\mathbf{X}}(\mathrm{k})}\right] + \mu\mathrm{L}\,\sigma_{\mathrm{v}}^2 \frac{I}{[\breve{\mathbf{X}}^T(\mathrm{k})\breve{\mathbf{X}}(\mathrm{k})]}\right] \quad (63)$$

Hence

$$E[\Delta\boldsymbol{w}(\mathrm{k}+1)] = \mathrm{E}[\mathrm{F}(\mathrm{k})\Delta\boldsymbol{w}(\mathrm{k})] \quad (64)$$

$$E[\|\Delta\boldsymbol{w}(\mathrm{k}+1)\|^2] = E[\|\mathrm{F}(\mathrm{k})\Delta\boldsymbol{w}(\mathrm{k})\|^2] \le E[\|\mathrm{F}(\mathrm{k})\|^2]E[\|\Delta\boldsymbol{w}(\mathrm{k})\|^2] \quad (65)$$

The term $E[\mathrm{F}(\mathrm{k})]$ should be less than equal to one so that the proposed algorithm will converge.

Considering SVD decomposition for transformed matrix $\boldsymbol{Q}\breve{\mathbf{X}}(\mathrm{k}) = \breve{\mathbf{X}}(\mathrm{k})$ as

$$\breve{\mathbf{X}}(\mathrm{k}) = \mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k}) \quad (66)$$

where $\emptyset(\mathrm{k}) \in \mathbb{R}^{\mathrm{NXL}}$ is a rectangular diagonal matrix having singular values of $\breve{\mathbf{X}}(\mathrm{k})$ on its main diagonal , $\mathbf{U}(\mathrm{k}) \in \mathbb{R}^{\mathrm{NXN}}$ and $\mathbf{V}(\mathrm{k}) \in \mathbb{R}^{\mathrm{LXL}}$ are unitary matricessuch that the columns of $\mathbf{U}(\mathrm{k})$ and the columns of $\mathbf{V}(\mathrm{k})$ show the left-singular vectors and right-singular vectors of $\breve{\mathbf{X}}(\mathrm{k})$, respectively.

Hence, we can write (63) as:

$$\mathrm{F}(\mathrm{k}) = \left[I - \mu\left[\frac{[\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k})\mathbf{V}(\mathrm{k})\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})]}{\mathbf{V}(\mathrm{k})\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k})}\right] + \mu\mathrm{L}\,\sigma_{\mathrm{v}}^2\left[\frac{I}{\mathbf{V}(\mathrm{k})\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k})}\right]\right] \quad (67)$$

As $\mathbf{V}^T(\mathrm{k})\mathbf{V}(\mathrm{k}) = \mathbf{I}_{\mathbf{LxL}}$ and $\mathbf{U}^T(\mathrm{k})\mathbf{U}(\mathrm{k}) = \mathbf{I}_{\mathbf{NxN}}$

Therefore,

$$\mathrm{F}(\mathrm{k}) = \left[I - \mu\left[\frac{[\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})]}{\mathbf{V}(\mathrm{k})\emptyset^T(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k})}\right] + \mu\mathrm{L}\,\sigma_{\mathrm{v}}^2\left[\frac{I}{\mathbf{V}(\mathrm{k})\emptyset^T(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k})}\right]\right] \quad (68)$$

$$\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})\left(\mathbf{V}(\mathrm{k})\emptyset^T(\mathrm{k})\emptyset(\mathrm{k})\mathbf{V}^T(\mathrm{k})\right)^{-1}\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k}) = \mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k}) \quad (69)$$

Assuming $(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}$ exists, then the term $\mu\mathrm{L}\,\sigma_{\mathrm{v}}^2(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}$ will be bounded and $\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k}) \in \mathbb{R}^{\mathrm{NXN}}$ will be a diagonal matrix such that

$$\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k}) = \begin{bmatrix} I_{LXL} & 0_{LX(N-L)} \\ 0_{(N-L)XL} & 0_{(N-L)X(N-L)} \end{bmatrix} \quad (70)$$

$$E[\|\Delta\boldsymbol{w}(\mathrm{k}+1)\|^2] = E[\Delta\boldsymbol{w}^T(\mathrm{k})\mathrm{F}^T(\mathrm{k})\mathrm{F}(\mathrm{k})\Delta\boldsymbol{w}(\mathrm{k})] = E[\Delta\boldsymbol{w}^T(\mathrm{k})[I-\mu]\mathrm{F}(\mathrm{k})\Delta\boldsymbol{w}(\mathrm{k})] \quad (71)$$

$$\mathrm{E}\{(\Delta\boldsymbol{w}^T(\mathrm{k})[I - \mu\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})]^T[I - \mu\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})]\Delta\boldsymbol{w}(\mathrm{k})\} \quad (72)$$

$$\mathrm{E}\left\{(\Delta\boldsymbol{w}^T(\mathrm{k})\left(I - \mu(2-\mu)\mathbf{U}(\mathrm{k})\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k})\mathbf{U}^T(\mathrm{k})\right)\Delta\boldsymbol{w}(\mathrm{k})\right\}$$

$$E[\|\Delta\boldsymbol{w}(\mathrm{k}+1)\|^2] = E\left[\|\Delta\boldsymbol{w}(\mathrm{k})\|^2 + \left((2-\mu)\emptyset(\mathrm{k})(\emptyset^T(\mathrm{k})\emptyset(\mathrm{k}))^{-1}\emptyset^T(\mathrm{k})\right)\|\ddot{\boldsymbol{w}}(\mathrm{k})\|^2\right] \quad (73)$$

where $\ddot{\boldsymbol{w}}(\mathrm{k}) = \mathbf{U}^T(\mathrm{k})\Delta\boldsymbol{w}(\mathrm{k})$

For acuate stability of the proposed algorithm, μ should be zero or two, but for practical purpose $0 \le \mu \le 1$ should be adopted [14].

V. RESULT AND DISCUSSION

In this section, simulations are taken in MATLAB software to justify the estimation behavior of the proposed algorithm in linear constrained sparse system identification against noisy input. The unknown finite impulse response (FIR) system and adaptive filter have same dimension N . Here we have considered N=163. The system coefficients are considered symmetric. The unknown sparse system coefficients are Gaussian distributed with linear phase aspect. For symmetric and odd condition, the linear phase constraint is defined as[2]:

$$\boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -1 & \dots & 0 \\ -1 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{\frac{N-1}{2}} \\ \mathbf{0}^T \\ -\boldsymbol{J}_{(N-1)/2} \end{bmatrix} \quad (74)$$

$$\mathbf{h} = [\mathbf{0}\ \mathbf{0} \dots \mathbf{0}]^T \quad (75)$$

Here $\mathbf{I}_{\frac{N-1}{2}}$ is an identity matrix of order $\frac{N-1}{2}$ and $\mathbf{J}$ is the identity matrix having all rows turned around. The colored input is produced by passing thewhite input through a system, $H(z) = \frac{1}{1-0.9z^{-1}}$. The measurement noise, z(k) is considered white. The assessment criterions for estimation performance are taken as: normalized mean square deviation (NMSD) and convergence speed. Here K is the sparsity constant that tells the number of non-zero coefficients among others.

First the proposed algorithm is compared with constrained affine projection (CAP) [5], affine projection algorithm (APA) [11] , bias compensator APA (BC-APA) [10] and l1-norm penalized constrained affine projection (l1-CAP) algorithms against Gaussian input noise and Laplace input noise for different values of sparsity constant, K={3,13,53} in figs. (2) and (3). The other parameters taken in this experiment are: step size μ =0.05, Input noise variance σ_{v}^2=0.08, Channel SNR=20dB; sparsityregularizer $\beta = 5\mathrm{X}10^{-5}$ and projection order L=6.

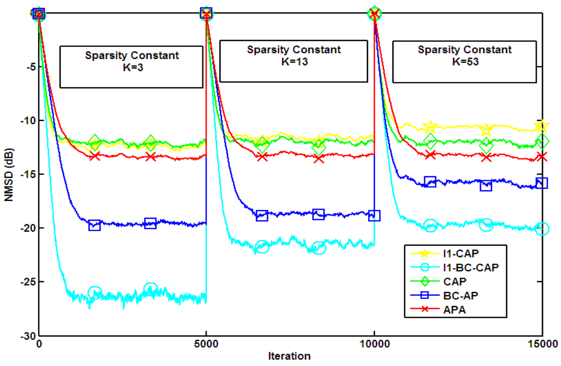

Fig. 2.NMSD of the proposed l1-BC-CAP, CAP, BC-APA, CAP, APA and l1-CAP algorithms with different K={3,13,53} in the presence of Gaussian inputnoise (μ =0.05, σ_v^2=0.08, SNR=20dB, $\beta = 5X10^{-5}, L = 6$)

Table- I: Comparison of transient NMSD of proposed algorithm with other algorithms for K=3 from fig. 2

Algorithm	Transient NMSD (dB)	Iteration Number
l1-BC-CAP	-21.78	746
BC-AP	-13.19	746
CAP	-11.53	746
l1-CAP	-12.08	746
APA	-10.47	746

Table- II: Comparison of steady state NMSD of proposed algorithm with other algorithms for K=3 from fig. 2

Algorithm	Steady State NMSD (dB)	Iteration number
l1-BC-CAP	-27.3	1120
BC-AP	-15.86	1602
CAP	-12.29	1404
l1-CAP	-12.43	1361
APA	-18.9	2338

Table I and II compare the transient NMSD and steady state NMSD of the proposed l1-BC-CAP algorithm with other APA variants for sparsity constant K=3 and Gaussian input noise. Table I confirms the lowest transient NMSD of the proposed algorithm. Table II, it is evident that the proposed achieves the lowest steady state NMSD with largest convergence rate among others. Similar types of results can be obtained for Laplace input and other values of K. However, for higher sparsity level the estimation performance of the proposed algorithm degrades but still it is better than other algorithms.

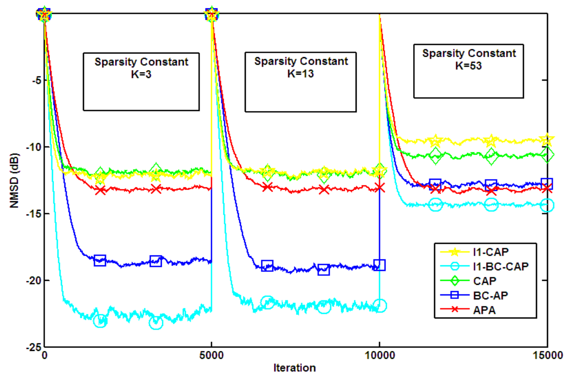

Fig. 3.NMSD of the proposed l1-BC-CAP, CAP, BC-APA, CAP, APA and l1-CAP algorithms with different K={3,13,53} in the presence of Laplace input noise (μ =0.05, σ_v^2=0.08, SNR=20dB, $\beta = 5X10^{-5}, L = 6$)

Table- III: Comparison of simulation time of proposed algorithm with existing variants

Algorithms	Simulation Time Per Run (Second)
l1-BC-CAP	0.141553
BC-APA	0.117227
CAP	0.056583
l1-CAP	0.086993
APA	0.035812

Table III compares the simulations time of proposed algorithm with its variants. Although the proposed algorithm needs the highest simulation time, yet the estimation performance is superior to other algorithms.

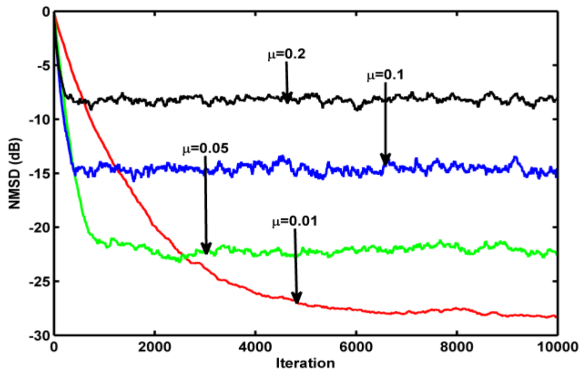

Fig. 4.Performance of l1-BC-CAP algorithm under various step sizes(Gaussian input noise, σ_v^2=0.08, SNR=20dB, K=13, $\beta = 5X10^{-5}, L = 6$)

Next experiment shows the step size's impact on the performance of the proposed algorithm. The convergence speed and NMSD error are related inversely to step size as shown in fig. 4.Therefore, the step size should be selected very carefully to make a balance between convergence speed and NMSD error. Here, the step size μ=0.05 is selected to take care of both.

The minimum MSD error is obtained for μ=0.01 and the highest convergence speed is for μ=0.2.

Figure 5 shows the performance of the proposed l1-BC-CAP algorithm under different input noise variance. Here Gaussian input noise is considered for simulation. Four values of input noise variance, σ_v^2=[0.05 0.08 0.2 0.4] are considered in this experiment. As the variance of the input noise increases, the performance of the proposed algorithm l1-BC-CAP is degraded slightly since the proposed algorithm has unbiased the bias produced due to the input noise, therefore its performance is not significantly degraded by increasing the value of input noise variance.

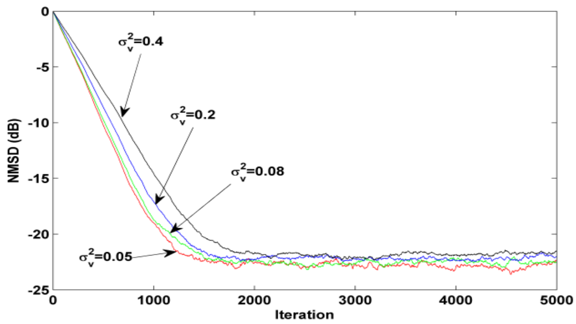

Fig. 5.Performance of l1-BC-CAP algorithm under various input noise variance (Gaussian input noise, μ =0.05, SNR=20dB, K=13, $\beta = 5X10^{-5}, L = 6$)

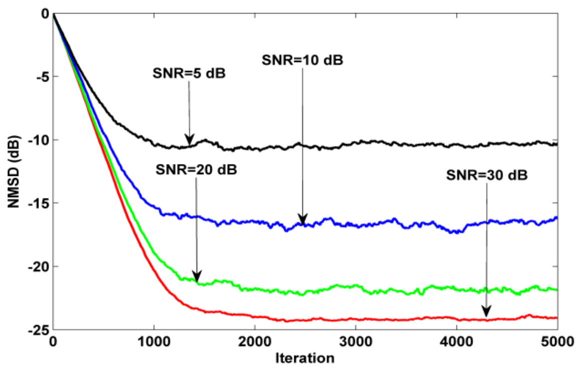

Fig. 6.Performance of the proposed algorithm under various channel SNR on the estimation (Gaussian input noise, σ_v^2=0.08, μ =0.05, K=13, $\beta = 5X10^{-5}, L = 6$)

In fig. 6, the performance of the proposed algorithm is tested for varying SNR values of channel noise. The channel noise and input noise are considered as white Gaussian noise. The other parameters are taken as: step size μ =0.05, input noise variance σ_v^2 =0.08, sparsity constant K=13; sparsity regularizer $\beta = 5X10^{-5}$. Here four different values of SNR of channel noise, SNR= [30dB 20dB 10 dB 5dB] are adopted to check the performance of the proposed algorithm.

As the value of SNR rises up, the performance improves and as it goes down, the performance degrades. So the channel SNR has severe impact on the performance of the proposed algorithm.

In next simulation, the impact of sparsity regularizer β on the performance of the proposed algorithm is tested. The other parameters are taken as: step size μ =0.05, input noise variance σ_v^2=0.08, sparsity constant K=13; channel SNR= 20 dB.

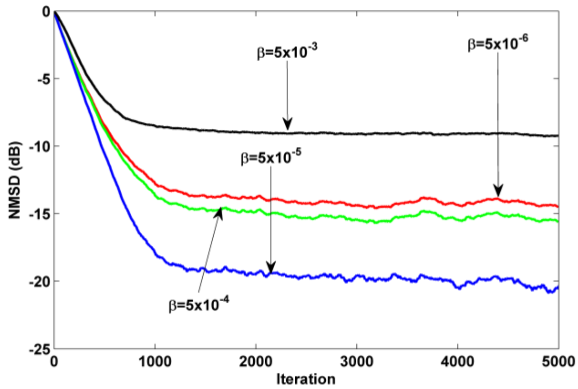

Fig. 7.Performance ofl1-BC-CAP algorithm under various sparsity regularizerβ(Gaussian input noise, σ_v^2=0.08, μ =0.05, $SNR = 20dB, K = 13, \beta = 5X10^{-5}, L = 6$)

From fig. 7, it is clear that when the value of the parameter β decreases from $5 * 10^{-3}$ to $5 * 10^{-5}$, the performance improves considerably. However, as it goes below $5 * 10^{-5}$ to $5 * 10^{-6}$, the performance starts to deteriorate. Therefore, the parameter β should therefore be chosen wisely as it directly affects the amount of zero attraction on the system coefficients.

VI. CONCLUSION

This paper presents an l1-norm penalized bias compensated linear constrained affine projection (l1-BC-CAP) algorithm. The proposed algorithm is used to identify a sparse system with linear phase aspect in the presence of colored input, corrupted by the additive input noise and channel noise. From table 1, it is clear that the l1-BC-CAP achieves -27.3 dB NMSD errors in just 1120 iterations than other APA based algorithms. Thus, the convergence speed is higher than other algorithms as shown in table 1. The simulation time is slightly higher than other algorithms as shown in table 2. However, the increase in simulation time is mainly due to addition of bias compensator term. Further, the performance is also tested for several sparsity levels and different input noise variances. The impact of several other parameters like step size μ, output SNR, sparsity regularizer β is also illustrated. The proposed algorithm outperforms for different sparsity levels and input noise variances. Thus the l1-BC-CAP algorithm has the ability to replace the existing adaptive algorithms in many practical implementations which involve combined effect of linear constrained adaptive filtering, sparseness characteristics, and colored input corrupted by input noise

REFERENCES

1. M. T. Schiavoni and M. G. Amin, "A linearly constrained minimization approach to adaptive linear phase and notch filters," *Proceedings. The Twentieth Southeastern Symposium on System Theory*, Charlotte, NC, USA, 1988, pp. 682-685.
2. R. Yadav,CS Rai. "Linear phase sparse system identification in the presence of impulsive noise", International Journal of Electronics Letters.2018.
3. J.A. Apolinario, S. Werner, P.S.R. Diniz, T. Laakso, "Constrained normalized adaptive filters for CDMA mobile communications", in Proceedings EUSIPCO, vol. IV, 1998, pp.2053–2056.
4. O. L. Frost, "An algorithm for linearly constrained adaptive array processing," in *Proceedings of the IEEE*, vol. 60, no. 8, pp. 926-935, Aug. 1972.
5. MLRD Campos,JA Apolinrio, "The constrained affine projection algorithm – Development and convergence issues," the Conf. Signal Processing, Communications, Circuits, Systems, Istanbul, Turkey,pp. 1-4, May 2000.
6. B Kang, JYoo, P. Park, "Bias-compensated Normalised LMS Algorithm with Noisy Input", IEEE Electronics Letters, 2013; 49(8) ,pp. 538–539.
7. ZZheng, Z Liu, L Lu, "Bias-compensated robust set-membership NLMS algorithm against impulsive noises and noisy inputs",IEEE Electronics Letters,2017, 53(16), pp. 1100–1102.
8. Z Zheng,HZhao, "Bias-compensatednormalizedsubbandadaptivefilteralgorithm", IEEESignalProcess Lett.2016, 23(6), pp. 809–813.
9. H Zhao, Z Zheng, "Bias-compensated affine-projection-like algorithms with noisy input", IEEE Electronics Letters, 2016, 52(9), pp. 712–714.
10. SM Jung, NK Kwon, P Park., "A bias-compensated affine projection algorithm for noisy input data", In: 9thAsianControlConference(ASCC),2013, pp. 1–5.
11. SG Sankaran, AAL Beex,"Convergence behavior of affine projection algorithms", IEEE Transactions on Signal Processing.2000,48(6), pp.1086–1096.
12. R Meng,RCLamare,VH Nascimento, "Sparsity-awareaffineprojectionadaptivealgorithmsforsystem identification,. In: Sensor Signal Processing for Defence, 2011, pp. 1–5.
13. MVS Lima, WA Martins, PSR Diniz, "Affine projection algorithms for sparse system identification", In: IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, 5666–5670.
14. S Werner, JA Apolinario, MLRD Campos, PSR Diniz, "Low-complexity constrained affine-projection algorithms", IEEE Transactions on Signal Processing.2005, 53(12),pp. 4545–4555.
15. W Wang, H Zhao, B Chen,"Bias compensated zero attracting normalized least mean square adaptive filter and its performance analysis" Signal Processing.2018,143, pp. 94–105.

AUTHORS PROFILE

Rajni Yadav, is an Assistant Professor in Department of Electroics and Communication Engineering, Maharaja Agrasen Institute of Technology, GGSIPU, Delhi. She obtained her M. Tech degree from Indira Gandhi Delhi Technical University for Woman in 2012. Her teaching and research interests include: Signal Processing, Adaptive filtering and Estimation Theory.

Dr. Chandra Shekhar Rai, is a Professor with the University School of Information & Communication Technology. He obtained his M.E. degree in Computer Engineering from SGS Institute of Technology & Science, Indore in 1994 and completed Ph.D. in area of Neural Network from Guru Gobind Singh Indraprastha University in 2003. His teaching and research interests include: Artificial Neural Systems, Computer Networks and, Signal Processing.