

A New Penalized Improved Proportionate Affine Projection Algorithm for Block Sparse Systems

Senthil Murugan Boopalan¹, Krithiga Sukumaran¹, Ashokaraju Devarajan², Sarojini Raju³, Kalphana Ilangovan⁴, Janani Munisamy¹

¹*Department of Electronics and Communication Engineering, Thanthai Periyar Government Institute of Technology, Vellore, Tamil Nadu, India.*

²*Department of Electrical and Electronics Engineering, Government College of Engineering, Salem, Tamil Nadu, India.*

³*Department of Electronics and Communication Engineering, Government College of Engineering, Bodinayakanur, Tamil Nadu, India.*

⁴*Department of Electronics and Communication Engineering, Government College of Engineering, Dharmapuri, Tamil Nadu, India.*

Abstract

Adaptive signal processing is a growing area of research, and there is a constant drive to improve adaptive algorithms for sparse system identification, such as echo cancellation. A new improved proportionate affine projection algorithm penalized with the l_0 -norm, is proposed for system identification of sparse systems whose impulse response coefficients are grouped in blocks, or block-sparse system identification. The l_0 -norm penalty is integrated into the cost function of the Block sparse Improved proportionate affine projection algorithm (BS-IPAPA) to derive the proposed algorithm l_0 -norm BS-IPAPA (l_0 -BS-IPAPA). The mean stability condition is derived. Simulation results reveal that the l_0 -BS-IPAPA outperforms its competing algorithms, namely IPAPA, l_0 -IPAPA, BS-IPAPA, and l_0 -APA with regard to misalignment, tracking, and convergence rate performance for identification of block-sparse systems. For a colored input single cluster channel, an improvement of about 31% in the initial convergence rate is achieved by the l_0 -BS-IPAPA over the BS-IPAPA. Similarly, for a colored input double cluster channel, the proposed algorithm is 14% better than the BS-IPAPA. Studies comparing the performance for higher projection order and dispersive systems are illustrated. To enhance the performance of the l_0 -BS-IPAPA further, its variable step-size version is derived and studied. A transient study of the proposed l_0 -BS-IPAPA against varying SNRs is conducted. An average 3 dB improvement in echo return loss enhancement is achieved by the proposed algorithm over the BS-IPAPA for real-world datasets. A comparative study is made between the computational overload of the proposed and its competing algorithms.

Keywords: affine projection algorithm, adaptive filter, echo cancellation, block sparse systems, system identification, signal processing.

1. Introduction

Sparse system identification is an active domain in the field of signal processing. Most of the practical systems like echo cancellation, channel equalization, and channel estimation are sparse or block-sparse. A sparse system is one in which most of the energy is concentrated in only a

few coefficients (significant) and others (insignificant) have little or zero energy [1], [2]. In block-sparse, the significant coefficients are bundled in clusters or blocks [3]. The normalized least mean square (NLMS) algorithm has limited use in echo cancellation because of its inability to exploit the sparse nature of the echo-path

and a poor convergence for the colored input [4], [5]. The affine projection algorithm (APA) shows better colored input performance by utilising a set of ‘p’ most recent regressor vectors [6]. The improved proportionate normalized least mean square (IPNLMS) algorithm utilises the sparse nature of the echo-path by individual allotment of the gain factors that are proportional to the filter coefficients magnitude [7].

The pursuit of efficient block-sparse system identification has led to several landmark developments in adaptive filtering. Liu and Grant [8] initially integrated the mixed $l_{2,1}$ norm into the cost function of the IPNLMS algorithm to target block-sparse structures. Parallel to this, advancements in Compressed Sensing (CS) inspired the integration of an approximated l_0 -norm penalty into the APA to more aggressively exploit system sparsity [9]–[10]. Over the years, many researchers have exploited sparsity [11]–[15]. By attracting both significant and insignificant coefficients toward their true values, the l_0 -norm penalty significantly accelerates the convergence rate of the adaptive filter. Drawing inspiration from both l_0 -NLMS and IPNLMS, hybrid approaches such as l_0 -IPNLMS and l_0 -norm improved proportionate affine projection algorithm (l_0 -IPAPA) were subsequently proposed [16], with [17] focusing on the development of computationally efficient implementations. Recently, the literature has seen a proliferation of affine-based algorithms specifically tailored for block-sparse environments [18]–[24], setting the stage for more specialized regularization techniques that balance convergence speed with steady-state accuracy. Table 1a and Table 1b compare the six adaptive algorithms based on their limitations and performance metrics.

Table 1a. Comparison of existing works in terms of limitations.

Existing work	Limitation
Least mean square (LMS) & NLMS	Slow convergence rate and poor performance for the colored input
APA	High Computational complexity
IPNLMS	Higher complexity and parameter tuning
Block sparse IPNLMS(BS-IPNLMS)	Better performance only when the insignificant coefficients appear in blocks
l_0-APA	Higher Computational complexity and poor performance for dense or non-sparse systems
Variable step-size APA (VSS-APA)	Higher computational complexity and poor performance for sparse systems

Table 1b. Comparison of existing works in terms of performance metrics.

Existing work	Convergence rate	Steady state NM (dB)	Tracking	Computational cost
LMS & NLMS	slow	-20 to -30	poor	low
APA	fast	-30 to -40	good	high
IPNLMS	fast	-35 to -45	good	moderate
BS-IPNLMS	very fast	-40 to -50	very good	moderate to high
l_0-APA	very fast	-45 to -55	very good	very high
VSS-APA	fast	-40 to -50	very good	high

In [16], the l_0 -norm is integrated into block sparse proportionate affine projection algorithm (BS-PAPA) to improve the convergence performance for block-sparse systems. Although block sparse improved PAPA (BS-IPAPA) shows good performance for block sparsity, its convergence speed is still limited and can be improved further. Motivated by the BS-IPAPA and [16], we propose the l_0 -norm block-sparse improved proportionate affine projection algorithm (l_0 -BS-IPAPA) by adding an approximated l_0 -norm to the BS-IPAPA cost function. The l_0 -norm in the update equation of the proposed algorithm acts as a zero attractor and exerts attraction to all coefficients. The zero attractor makes l_0 -BS-IPAPA show a higher convergence rate and improved steady-state performance than the BS-IPAPA and l_0 -IPAPA. Simulation results confirm the improved performance of the proposed algorithm over its competing algorithms.

The primary contributions of this research are:

1. **Algorithmic Innovation:** Inspired by the BS-IPAPA, a novel algorithm named l_0 -BS-IPAPA is proposed by incorporating an approximated l_0 -norm into the cost function, specifically designed to exploit the structural sparsity of block-clustered impulse responses.
2. **Theoretical Analysis:** The steady-state mean stability condition is derived for the proposed algorithm. The analysis reveals that the condition indeed is dependent on the input signal power level.
3. **Comprehensive Validation:** Many simulations are carried out to prove the efficacy of the l_0 -BS-

IPAPA over the other related algorithms in the literature. Experimental validation with real-world datasets is provided.

4. Complexity and Performance Profiling: In terms of computational complexity, the proposed algorithm is evaluated against its related algorithms by considering the additions, multiplications, divisions, comparisons, memory spaces, and execution time. In addition, a study on its transient performance is carried out.

The remaining manuscript is structured as follows. The following section briefs the algorithms like l_0 -IPAPA and BS-IPAPA, and section 3 introduces the proposed l_0 -BS-IPAPA. Section 4 concentrates on the mean stability condition for the l_0 -BS-IPAPA. Simulation experiments and the results obtained are elaborated in section 5. The study on transient performance, computational complexity and experimental validation with real-world datasets are presented in section 6. With emphasis on the future work, section 7 concludes the research undertaken.

Notation

Throughout this manuscript, the following conventions are adopted to distinguish between different types of mathematical quantities. Boldfaced lower case symbols (e.g., $\mathbf{x}$) denote column vectors and boldfaced uppercase symbols (e.g., $\mathbf{X}$) denote matrices. Additionally, lowercase letters in normal font are used to represent scalar quantities. The notation $e(l)$ and $\mathbf{e}(l)$ are utilized to reflect the time-varying nature of scalars and vectors. The following notations are adopted in this article.

$(\cdot)^T$	Transposed vector
$\mathbf{E}(\cdot)$	Statistical Expectation or mean
$\ \cdot\ $	Vector norm in Euclidean space
$\lceil \cdot \rceil$	Ceiling function
$\ \mathbf{x}\ _A^2$	Weighted inner product
$\mathbf{I}_p$	Identity matrix of size $p \times p$

2. Related Work

The theoretical framework of this study is grounded in the principles of sparse system identification, specifically targeting the challenge of echo cancellation. The practical significance of the adaptive algorithm can be highlighted in echo cancellation, channel estimation, and channel equalization. Adaptive filters are highly important in real-world applications like Voice over Internet Protocol (VoIP) networks to minimize the line echo caused by impedance mismatch and delay distortion. For frequency selective fading, orthogonal frequency division multiplexing (OFDM) systems heavily depend on channel estimation, and subsequent channel equalization.

In an echo cancellation scenario, the primary objective is to accurately model the characteristic of a dynamic echo path. A typical modelling of an echo canceller is shown in Figure 1. The echo path to be identified is given by $\mathbf{h}$, whose length is L . The far-end input signal vector at time instant i is defined by the following snapshot $\mathbf{y}(i) = [y(i), y(i-1), y(i-2), \dots, y(i-L+1)]^T$. The desired signal is then stated, as the sum of the echo path's response and the additive noise or near-end $a(i)$ [2],

$$\mathbf{x}(i) = \mathbf{y}^T(i)\mathbf{h} + a(i) \quad (1)$$

The estimation of the echo-generating path is carried out by the adaptive filter as $\mathbf{s}(i-1) = [s_0(i-1), s_1(i-1), \dots, s_{L-1}(i-1)]^T$. The adaptive filter's output is

$$z(i) = \mathbf{y}^T(i)\mathbf{s}(i-1) \quad (2)$$

Then, the estimation error for the unknown path is

$$e(i) = \mathbf{x}(i) - z(i) \quad (3)$$

2.1. l_0 -IPAPA

To enhance the convergence speed and more effectively exploit system sparsity, the APA and its variants, such as IPAPA, utilize a "data recycling" mechanism. Instead of updating the filter based on a single input vector, these algorithms use a collection of recent input vectors, forming an input data matrix.

Following the notation in [25], let the input data matrix at time instant i be defined as:

$$\mathbf{Y}(i)=[\mathbf{y}(i),\mathbf{y}(i-1),\dots,\mathbf{y}(i-p+1)] \quad (4)$$

with ‘ p ’ denotes the projection order associated with IPAPA and $p < L$. The output $\mathbf{z}$, desired response $\mathbf{x}$, and error vectors $\mathbf{e}$ associated with IPAPA are expressed as

$$\mathbf{x}(i) = [\mathbf{x}(i), \mathbf{x}(i-1), \dots, \mathbf{x}(i-p+1)]^T \quad (5)$$

$$\mathbf{e}(i) = \mathbf{x}(i) - \mathbf{z}(i) \quad (6)$$

APA is derived by projection onto affine subspaces. The l_0 -IPAPA is obtained by adding an l_0 -norm penalty function to the IPAPA's cost function.

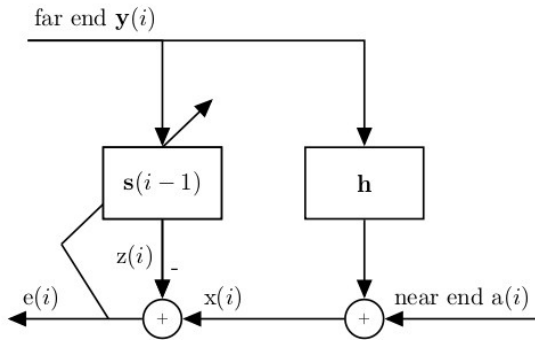


Figure 1. An Echo canceller model.

The update equation for the l_0 -IPAPA is given by [10]

$$\begin{aligned} \mathbf{s}(i) = & \mathbf{s}(i-1) + \mu \mathbf{T}_{\text{IPAPA}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{e}(i) \\ & + \frac{\kappa\beta}{2} \mathbf{T}_{\text{IPAPA}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \text{sgn}(\mathbf{s}(i-1)) e^{-\beta \mathbf{s}(i-1)} \\ & - \frac{\kappa\beta}{2} \text{sgn}(\mathbf{s}(i-1)) e^{-\beta \mathbf{s}(i-1)} \end{aligned} \quad (7)$$

where $\mathbf{Y}^+(i) = (\mathbf{Y}^T(i) \mathbf{T}_{\text{IPAPA}}(i-1) \mathbf{Y}(i) + \delta_p \mathbf{I}_p)^{-1}$, δ is a small regularization parameter to prevent numerical instability, $\kappa = \mu\gamma$ wherein μ is the step-size and γ is a factor for zero attraction, $\text{sgn}(\cdot)$ is the signum function and β should always be chosen greater than zero.

The diagonal matrix $\mathbf{T}_{\text{IPAPA}}(i-1)$ is the gain distribution matrix of the IPAPA and is given by

$$\mathbf{T}_{\text{IPAPA}}(i-1) = \text{diag}[t_0(i-1)\mathbf{1}_Q, t_1(i-1)\mathbf{1}_Q, \dots, t_{M-1}(i-1)\mathbf{1}_Q] \quad (8)$$

where $\mathbf{1}_Q$ is a row vector having ‘ Q ’ ones.

The m^{th} block or cluster has the gain element $t_m(i-1)$ and is given by

$$t_m(i-1) = \frac{(1-\alpha)}{2} + \frac{(1+\alpha)|s_m(i-1)|_2}{2 \sum_{n=0}^{M-1} |s_n(i-1)|_2}, 0 \leq m \leq M-1 \quad (9)$$

where α is a constant and is chosen between -1 and 1. If $\alpha = -1$, IPAPA is equal to APA. By setting $\alpha = 1$, the IPAPA behaves like PAPA. A good choice for α is either 0 or -0.5[7].

2.2. BS-IPAPA

BS-IPAPA was developed by introducing a mixed-norm $l_{2,1}$ penalty into the IPAPA framework, specifically to exploit the block-sparse characteristics of echo channels. Within the BS-IPAPA architecture, the *a priori* output error vector is $\mathbf{e}(i)$ defined as $\mathbf{e}(i) = \mathbf{x}(i) - \mathbf{Y}^T(i)\mathbf{s}(i-1)$.

The weight update equation of the BS-IPAPA is given by [8]

$$\begin{aligned} \mathbf{s}(i) = & \mathbf{s}(i-1) \\ & + \mu \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) (\mathbf{Y}^T(i) \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) + \delta_p \mathbf{I}_p)^{-1} \mathbf{e}(i) \end{aligned} \quad (10)$$

The diagonal matrix $\mathbf{T}_{\text{BI}}(i-1)$ is

$$\mathbf{T}_{\text{BI}}(i-1) = \text{diag}[t_0(i-1)\mathbf{1}_Q, t_1(i-1)\mathbf{1}_Q, \dots, t_{M-1}(i-1)\mathbf{1}_Q] \quad (11)$$

with $\mathbf{1}_Q$, a row vector having Q ones. The m^{th} cluster has the gain element $t_m(i-1)$ and is defined by (9).

While the BS-IPAPA effectively handles block-sparse structures, its performance can be further optimized by incorporating a more aggressive sparsity-promoting penalty. In the following section, we propose an improved proportionate affine projection algorithm that integrates an approximated l_0 -norm penalty to further accelerate convergence and reduce steady-state misalignment in block-sparse systems.

3. Proposed l_0 -BS-IPAPA

To more effectively leverage the inherent sparsity of block-structured systems, we introduce an approximated l_0 -norm penalty into the BS-IPAPA cost function. This modification aims to enhance the algorithm's ability to

distinguish between active clusters and zero-valued coefficients, thereby accelerating the convergence rate and improving steady-state accuracy.

The proposed l_0 -BS-IPAPA looks for an optimum solution for the following minimization problem:

$$\min_{\mathbf{s}(i)} \|\mathbf{s}(i) - \mathbf{s}(i-1)\|_{\mathbf{T}_{\text{BI}}^{-1}(i-1)}^2 \quad (12)$$

$$+\gamma \|\mathbf{T}_{\text{BI}}^{-1}(i-1)\mathbf{s}(i)\|_0 \text{ s.t. } \mathbf{d}(i) - \mathbf{Y}^T(i)\mathbf{s}(i) = 0$$

where $\mathbf{T}_{\text{BI}}^{-1}(i-1)$ is defined by (11) and γ is a zero attractor factor. Except for the gain distribution matrix, the derivation of the update equation of the l_0 -BS-IPAPA is same as that of l_0 -BS-PAPA [21] and is given by

$$\mathbf{s}(i) = \mathbf{s}(i-1) + \mu \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{e}(i) \quad (13)$$

$$+ \frac{\kappa}{2} \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \mathbf{f}_{\beta}(\mathbf{s}(i-1)) - \frac{\kappa}{2} \mathbf{f}_{\beta}(\mathbf{s}(i-1))$$

where $\mathbf{Y}^+(i) = (\mathbf{Y}^T(i) \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) + \delta_p \mathbf{I}_p)^{-1}$

$$\text{and } \mathbf{f}_{\beta}(\cdot) = \begin{cases} \beta^2 t + \beta & \text{if } -\frac{1}{\beta} \leq t < 0 \\ \beta^2 t - \beta & \text{if } 0 < t \leq \frac{1}{\beta} \end{cases}$$

$$\mathbf{s}(i) = \mathbf{s}(i-1) + \mu \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{e}(i) \quad (14)$$

$$+ \frac{\kappa}{2} \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \mathbf{f}_{\beta}(\mathbf{s}(\lceil \frac{i-1}{B} \rceil))$$

$$- \frac{\kappa}{2} \mathbf{f}_{\beta}(\mathbf{s}(\lceil \frac{i-1}{B} \rceil))$$

In (13), the approximation $\mathbf{s}(i) \approx \mathbf{s}(i-1)$ is utilized. The step-size μ is introduced to reduce the tradeoff between convergence rate and steady-state error. The $\delta_p \mathbf{I}_p$ regularization is for ensuring numerical stability. Additionally, $\kappa = \mu\gamma$ is acting as a zero attractor parameter, speeding up the convergence of near zero/inactive coefficients to zero.

4. Convergence Analysis of l_0 -BS-IPAPA

In this section, we investigate the mean stability of the proposed l_0 -BS-IPAPA. The step-size μ is a critical parameter that governs the convergence behavior of

adaptive algorithms; while a larger step-size can accelerate initial convergence, it also increases the risk of divergence from the optimal solution. Consequently, it is essential to determine the theoretical range of μ that ensures convergence in the mean, defined as $E(\tilde{\mathbf{s}}(i)) \rightarrow 0$ as $i \rightarrow \infty$. To facilitate this analysis, we define the weight-error vector as $\tilde{\mathbf{s}}(i) = \mathbf{h} - \mathbf{s}(i)$, where $\mathbf{h}$ represents the optimal impulse response. By substituting this definition into the update equation of the l_0 -BS-IPAPA, the error transition can be expressed as:

$$\tilde{\mathbf{s}}(i) = \tilde{\mathbf{s}}(i-1) - \mu \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{e}(i) \quad (15)$$

$$- \frac{\kappa}{2} \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \mathbf{f}_{\beta}(\mathbf{s}(i-1)) + \frac{\kappa}{2} \mathbf{f}_{\beta}(\mathbf{s}(i-1))$$

$\mathbf{x}(i) = \mathbf{a}(i) + \mathbf{Y}^T(i) \mathbf{h}$, where the near end noise vector $\mathbf{a}(i) = [\mathbf{a}(i), \mathbf{a}(i-1), \dots, \mathbf{a}(i-p+1)]^T$, can be substituted in $\mathbf{e}(i) = \mathbf{x}(i) - \mathbf{Y}^T(i) \mathbf{s}(i-1)$ to get

$\mathbf{e}(i) = \mathbf{a}(i) + \mathbf{Y}^T(i) \tilde{\mathbf{s}}(i-1)$, which in (15) yields

$$\tilde{\mathbf{s}}(i) = \tilde{\mathbf{s}}(i-1) - \mu \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) (\mathbf{a}(i) + \mathbf{Y}^T(i) \tilde{\mathbf{s}}(i-1)) \quad (16)$$

$$- \frac{\kappa}{2} \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \mathbf{f}_{\beta}(\mathbf{s}(i-1)) + \frac{\kappa}{2} \mathbf{f}_{\beta}(\mathbf{s}(i-1))$$

where $\mathbf{a}(i) = [\mathbf{a}(i), \mathbf{a}(i-1), \dots, \mathbf{a}(i-p+1)]^T$.

Taking expectation on both sides to get

$$E(\tilde{\mathbf{s}}(i)) = E(\tilde{\mathbf{s}}(i-1)) - \mu E(\mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) (\mathbf{a}(i) + \mathbf{Y}^T(i) \tilde{\mathbf{s}}(i-1))) \quad (17)$$

$$- \frac{\kappa}{2} E(\mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \mathbf{f}_{\beta}(\mathbf{s}(i-1))) + \frac{\kappa}{2} E(\mathbf{f}_{\beta}(\mathbf{s}(i-1)))$$

To ensure the tractability of the mean stability and convergence analysis, the following standard assumptions are adopted, consistent with the established literature [26]–[29]:

- Assumption 1: The additive near-end noise $\mathbf{a}(i)$ is a zero-mean white Gaussian noise.
- Assumption 2: The input vector $\mathbf{y}(i)$, the weight-error vector $\tilde{\mathbf{s}}(i)$ and the noise $\mathbf{a}(i)$ are mutually statistically independent. This implies that the current weight-error vector is independent of the current input data.

The above assumptions are needed not because they are true in practice but because they make the expectation analysis easily tractable while giving results that closely match the practical behaviour.

Then, (17) can be rewritten as

$$\begin{aligned} E(\tilde{\mathbf{s}}(i)) = & (\mathbf{I} - \mu E(\mathbf{T}_{\text{BI}}(i-1)\mathbf{Y}(i)\mathbf{Y}^+(i)\mathbf{Y}^T(i))E(\tilde{\mathbf{s}}(i-1)) \\ & - \frac{\kappa}{2}E(\mathbf{T}_{\text{BI}}(i-1)\mathbf{Y}(i)\mathbf{Y}^+(i)\mathbf{Y}^T(i)\mathbf{f}_\beta(\mathbf{s}(i-1))) + \frac{\kappa}{2}E(\mathbf{f}_\beta(\mathbf{s}(i-1))) \end{aligned} \quad (18)$$

Here, the contribution of $\mathbf{f}_\beta(\mathbf{s}(i-1))$ is negligible. Consequently, the last two terms can be omitted without loss of generality.

The update equation becomes

$$E(\tilde{\mathbf{s}}(i)) = (\mathbf{I} - \mu \mathbf{W}_{\text{BI}})E(\tilde{\mathbf{s}}(i-1)) \quad (19)$$

where $\mathbf{W}_{\text{BI}} = E(\mathbf{T}_{\text{BI}}(i-1)\mathbf{Y}(i)\mathbf{Y}^+(i)\mathbf{Y}^T(i))$.

Thus, the convergence in the mean is guaranteed for the range of μ

$$0 < \mu < \frac{2}{\Lambda_{\max}(E(\mathbf{W}_{\text{BI}}))} \quad (20)$$

The inequality derived in (20) establishes the necessary and sufficient condition for the stability of the l_0 -BS-IPAPA. Notably, the analysis reveals that the upper bound for the step-size μ is not a fixed constant; rather, it is dynamic and directly dependent on the statistical properties and power level of the input signal. Consequently, to ensure convergence in the mean, the selection of μ must be carefully tuned to the specific characteristics of the far-end excitation signal.

5. Simulation Results and Discussion

A series of simulation experiments are performed to evaluate the proposed l_0 -BS-IPAPA. The simulation setup includes: an adaptive filter's length $N = 1024$, matching the unknown system's length. The proposed algorithm's performance is evaluated in comparison with the prevailing algorithms such as the IPAPA, l_0 -IPAPA, l_0 -APA, and BS-IPAPA, in terms of the convergence speed, normalized misalignment (NM), and tracking capability. Two different classes of clusters, as in Figure 2, are utilised. A single-cluster is displayed in Figure 2a, which has 32 active non-zero coefficients in the [281, 312] range and a double-cluster is shown in Figure 2b,

which has 64 non-zero coefficients, 32 each in the ranges [281, 312] and [793, 824].

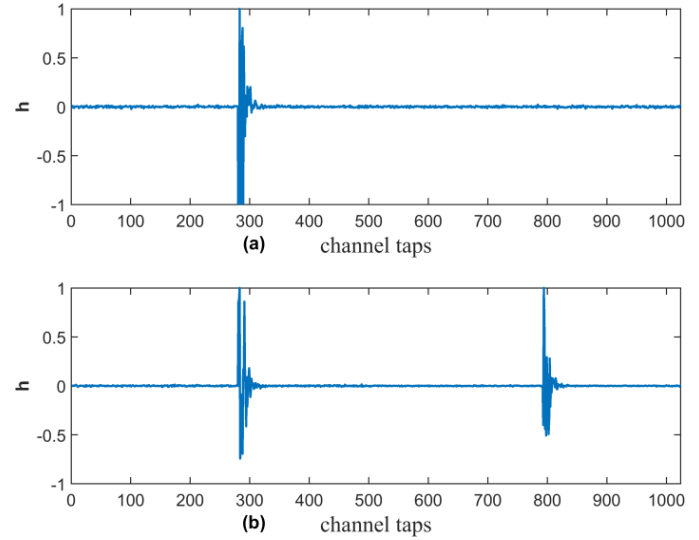


Figure 2. Cluster-sparse channel with its types: (a) a single cluster; (b) a double cluster.

In this study, we evaluate the performance of the proposed l_0 -BS-IPAPA using two distinct echo path scenarios: a network echo path and a satellite echo path. These are modeled as single-cluster and double-cluster sparse systems, respectively, to reflect real-world impulse response characteristics. To rigorously assess the tracking capability of the algorithm, the simulation environment is designed with an abrupt structural change. The experiment begins by identifying a single-cluster sparse system. At a predefined midpoint of the simulation, the system undergoes an instantaneous transition to a double-cluster configuration. This transition simulates a sudden change in the echo path, such as a shift in the acoustic environment or a circuit reconfiguration allowing us to measure how quickly and accurately the l_0 -BS-IPAPA readapts to new block-sparse parameters compared to existing benchmarks.

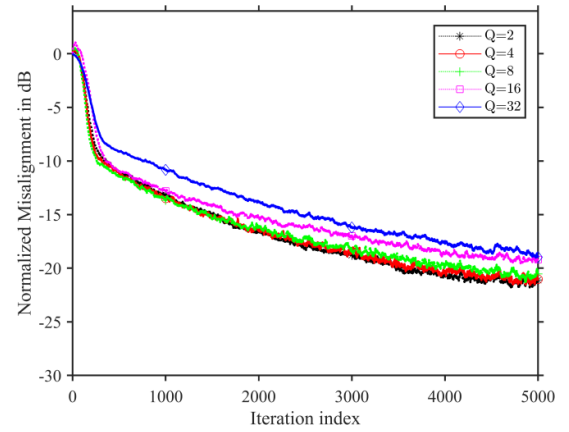
The performance of the l_0 -BS-IPAPA is validated using three distinct excitation signals: White Gaussian Noise (WGN), colored noise, and real-time speech. The colored noise is generated by processing WGN through a first-order system with a pole located at 0.8 to simulate spectral correlation. The speech signal is a real-world recording sampled at a frequency of 8 kHz. The simulation parameters are chosen as in [21]. The adaptive filter $\mathbf{s}$ is initialized as a zero vector. To simulate practical operating environments, the unknown

system is corrupted by independent additive WGN, resulting in a Signal-to-Noise Ratio (SNR) of 30 dB. To quantitatively assess the performance of the adaptive algorithms, we employ the Normalized Misalignment (NM) metric [24], defined in decibels (dB) as: $10\lg\{(\|\mathbf{h}-\mathbf{s}\|)^2/(\|\mathbf{h}\|)^2\}$.

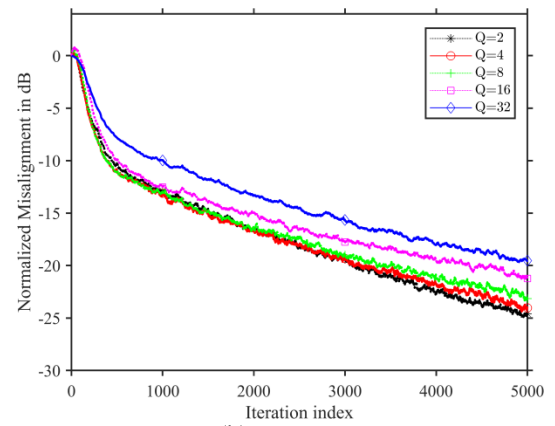
For all simulation scenarios, the identification process is divided into two phases: the first half of the iterations focuses on a single-cluster sparse system, while the other half evaluates the tracking response following a transition to a double-cluster configuration. The efficacy of the l_0 -BS-IPAPA is established through a comparative analysis against state-of-the-art algorithms found in current literature. In all simulations, the accuracy of the plots in this section is improved by averaging over 20 independent Monte Carlo runs. Excepting the special cases, the following parameters are used: $\mu = 0.9$, $\beta = 5$, $Q = 4$, $p = 4$. In Figure 2, choice of active filter coefficient locations and choice of cluster sizes are chosen as in [11].

5.1. Effect of Varying the Block Size Q on l_0 -BS-IPAPA

The block size Q is a pivotal parameter in the l_0 -BS-IPAPA, as it dictates the granularity of the proportionate gains. To evaluate its influence on the NM, we conducted a parametric study using block sizes of $Q = \{2, 4, 8, 16, 32\}$. The resulting performance curves are illustrated in Figures 3a, 3b, and 4b for colored noise, WGN, and speech inputs, respectively. An analysis of these figures reveals a clear inverse relationship between the block size and identification accuracy: as Q increases, the convergence rate tends to decelerate and the steady-state NM floor rises. This behavior suggests that larger block sizes may over-average the proportionate gains, thereby reducing the algorithm's sensitivity to the specific sparsity pattern of the impulse response. Across all three input signal types, the optimal balance between convergence speed and steady-state precision was consistently achieved at $Q=4$.



(a) colored



(b) WGN

Figure 3. Effect of varying block-size Q on the proposed l_0 -BS-IPAPA.

5.2. Performance Measurement of l_0 -BS-IPAPA with Varying Zero Attractor Parameter

The performance of the l_0 -BS-IPAPA is significantly influenced by the zero-attractor parameter, κ , which governs the strength of the penalty applied to the coefficients. Figure 4b illustrates the impact of varying κ on the algorithm's NM. The results indicate that as κ increases across the tested values of $[1, 3, 6, 9] \times 10^{-7}$, the convergence rate accelerates. This enhancement is attributed to the heightened attraction intensity, which more aggressively forces insignificant coefficients toward zero. However, this gain in speed comes at the cost of steady-state accuracy. Consequently, the selection of the zero-attractor parameter requires a judicious compromise between rapid initial convergence and the desired level of steady-state precision. Based on the observed curves, a balanced value of $\kappa (=4 \times 10^{-7})$ should be chosen for rest of the simulations to optimize this tracking-accuracy trade-off.

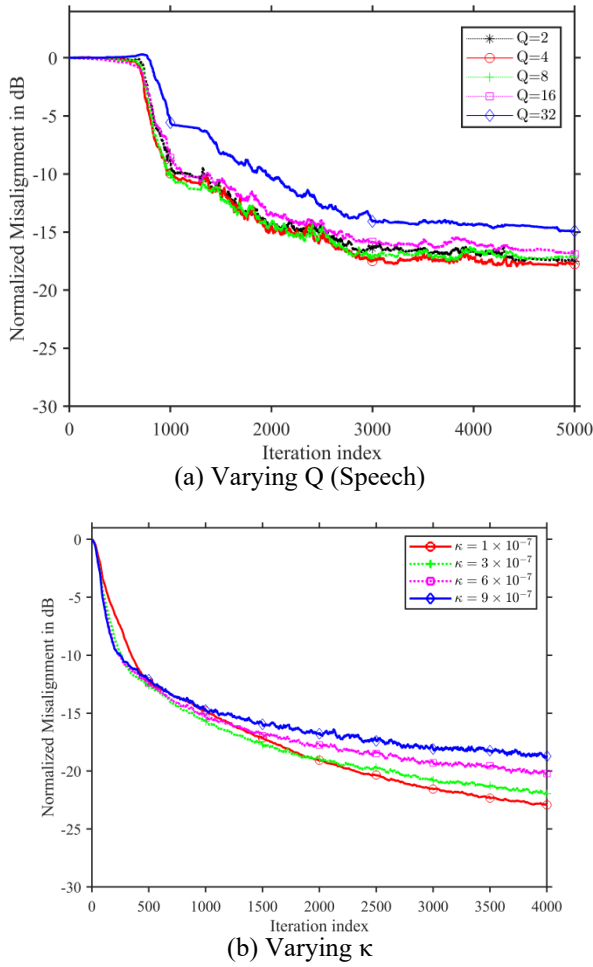


Figure 4. Effect of varying block-size Q and zero attractor κ on the proposed l_0 -BS-IPAPA.

5.3. Performance Comparison of l_0 -BS-IPAPA and the Competing Algorithms

Figures 5a and 5b illustrate the performance of the proposed l_0 -BS-IPAPA compared to existing benchmarks using WGN and speech signal inputs, respectively. Furthermore, the algorithms' robustness against highly correlated signals is evaluated in Figures 6a and 6b, where colored noise inputs are generated using poles at 0.8 and 0.95. To ensure a fair comparison, the regularization parameter δ is set to 0.01 for the l_0 -APA, while for all other algorithms, it is normalized to $\frac{0.01}{N}$. For the sparsity-aware variants—including the l_0 -APA, l_0 -IPAPA, and the proposed l_0 -BS-IPAPA—the zero-attractor parameter is consistently set to 4×10^{-7} . Across all tested input signals (WGN, colored noise, and speech), the simulation results consistently demonstrate that the l_0 -BS-IPAPA outperforms its counterparts in terms of superior convergence rate, a lower steady-state

NM floor, and enhanced tracking ability, particularly when identifying the structural transitions in block-sparse systems.

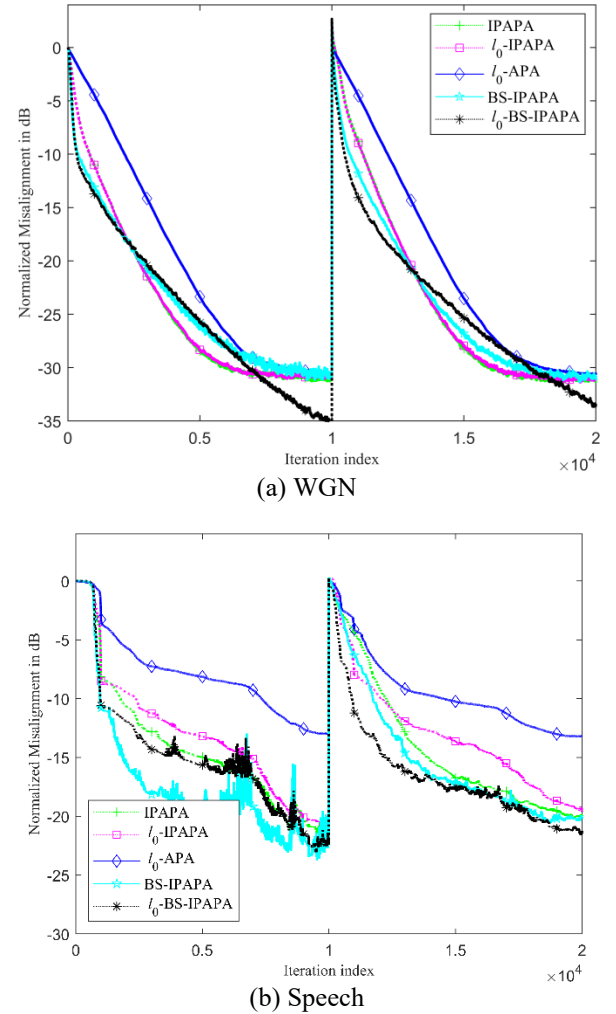


Figure 5. Performance evaluation of l_0 -BS-IPAPA and existing algorithms.

In Figure 6b, in case of the colored input (pole at 0.95) single cluster channel, for the iteration index in the range [1000, 1666], the NM shown by the BS-IPAPA is in the range [-12.01, -14.40] dB. On the other hand, the proposed algorithm yields NM in the range [-13.95, -17.11] dB, which clearly indicates an improvement of about 31% (rounded off to the nearest integer) in the initial convergence rate achieved by the l_0 -BS-IPAPA over the BS-IPAPA. For a colored input double cluster channel with the iteration index in the range [10666, 11332], the NM shown by the BS-IPAPA is in [-12.02, -14.03] dB whereas the misalignment shown by the l_0 -BS-IPAPA is in [-13.01, -15.3] dB which is 14% better than the NM shown by the BS-IPAPA. Significant improvement is achieved in NM for the other two inputs.

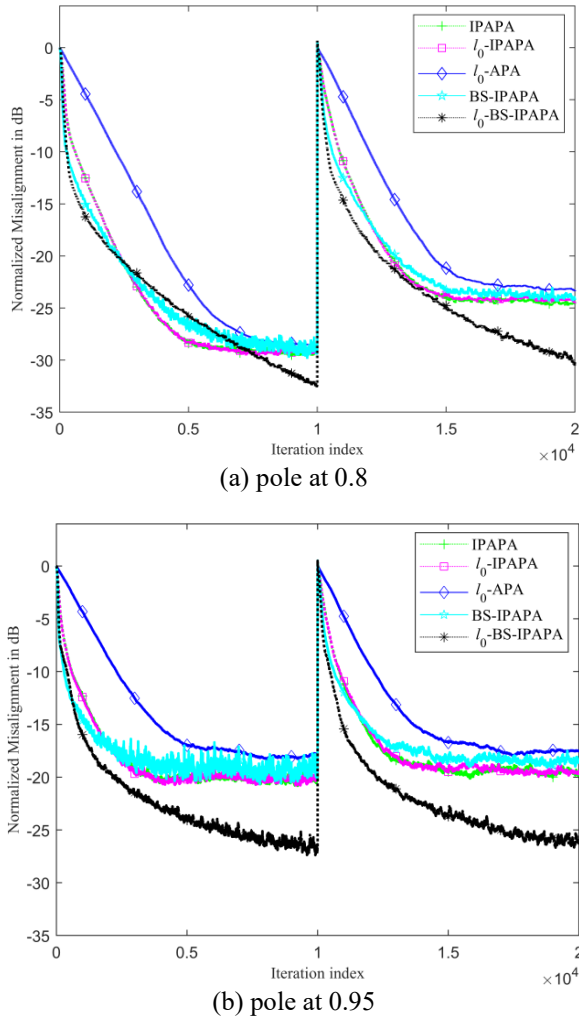


Figure 6. Performance evaluation of l_0 -BS-IPAPA and existing algorithms with colored input.

To validate the performance improvement statistically, the mean, standard deviation, and 95% confidence intervals were obtained using Student's t-distribution with 19 degrees of freedom. At 95% confidence interval (C.I), 31% and 14 % arrived at [30.2, 32.2] % and [13.3, 14.9] %, respectively. Further, a paired t-test was conducted between the proposed and BS-IPAPA. Pairs were formed by evaluating both algorithms across 20 independent simulation runs. The test resulted in a test statistic of $t(19)=2.5$ with an exact p-value of 0.0213 ($p<0.05$), confirming that the claimed improvements are statistically significant.

5.4. Performance Comparison for Higher Projection Order and Dispersive systems

Figures 7a and 7b evaluate the performance of the l_0 -BS-IPAPA and the BS-IPAPA under higher projection

orders and varying degrees of dispersiveness. Since the APA with a low projection order behaves similarly to the NLMS algorithm, it is essential to assess the proposed method's behavior as the projection order p increases. As shown in Figure 7a, higher projection orders accelerate the convergence rate for both algorithms due to the improved decorrelation of the colored input signal. However, this speed is accompanied by an increase in the steady-state NM floor, reflecting the classic trade-off in APA-based filtering.

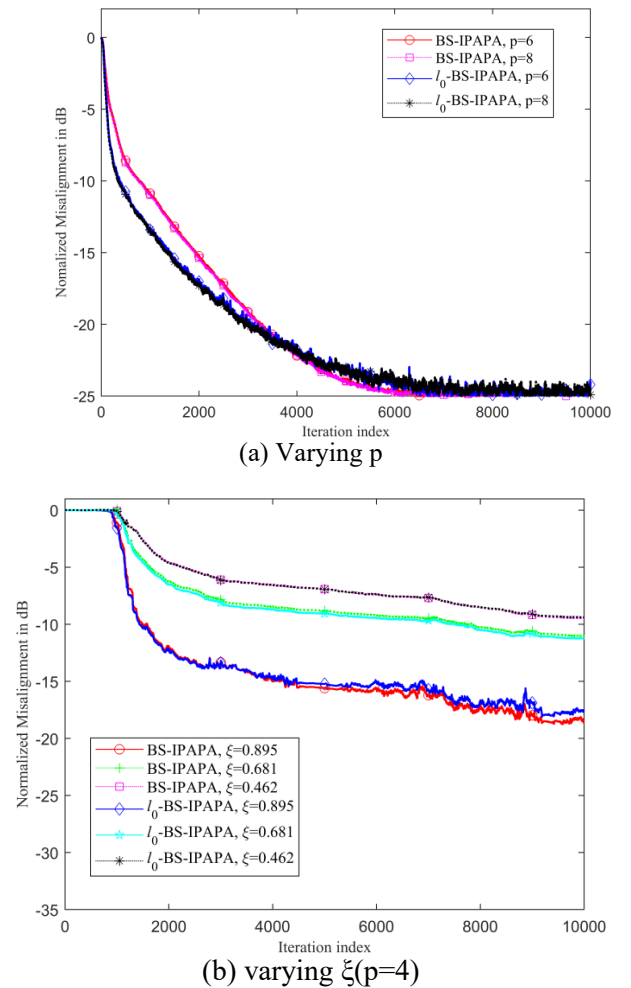


Figure 7. Performance evaluation of l_0 -BS-IPAPA and existing algorithms with colored input for higher projection order and dispersive systems.

Furthermore, since the exact sparsity of an unknown system is rarely known *a priori*, we evaluate the l_0 -BS-IPAPA across three levels of sparsity: high ($\xi = 0.895$), medium ($\xi = 0.681$), and low ($\xi = 0.462$). As illustrated in Figure 7b, the l_0 -BS-IPAPA consistently yields a faster convergence rate and a lower steady-state NM floor than the BS-IPAPA across all levels of dispersiveness. These results confirm that the proposed

algorithm outperforms even in less sparse (dispersive) environments and under high-order projection settings.

5.5. Performance Comparison of l_0 -BS-IPAPA and Existing Algorithms with Varying Block Sizes

To further investigate the sensitivity of the l_0 -BS-IPAPA to its structural parameters, we conducted a dynamic simulation where the block size Q is adjusted sequentially. Four distinct block sizes $Q = \{4, 8, 16, 32\}$ were evaluated using a colored noise input to test the sensitivity analysis for Q . As shown in Figure 8a, the block size was incremented every 10,000 iterations, while all other algorithmic parameters were maintained as described in Section 5.3. The results demonstrate a consistent trend across all evaluated algorithms: as Q increases from 4 to 32 (from left to right in Figure 8a), there is a noticeable deceleration in the convergence rate and a corresponding rise in the steady-state NM floor. This confirms that the choice of block size significantly impacts the resolution of the proportionate update. However, across all four intervals, the l_0 -BS-IPAPA consistently outperforms the IPAPA, l_0 -IPAPA, l_0 -APA, and BS-IPAPA. Even at larger block sizes, the proposed algorithm maintains a faster convergence trajectory and a superior steady-state misalignment floor, highlighting its robustness to variations in the block-sparse configuration.

5.6. Variable step-size l_0 -BS-IPAPA

The use of a fixed step-size μ in the l_0 -BS-IPAPA inevitably involves a fundamental trade-off: a larger μ facilitates rapid initial convergence but results in a higher steady-state normalized misalignment, whereas a smaller μ yields higher precision at the cost of convergence speed. To circumvent this limitation and further enhance the performance of the proposed algorithm, we introduce a time-varying step-size, $\mu(i)$, leading to the development of the Variable step-size l_0 -BS-IPAPA (VSS- l_0 -BS-IPAPA).

By incorporating $\mu(i)$ into the weight update equation (13) and neglecting the regularization parameter δ in the matrix inversion for the purpose of derivation, the update law is reformulated as:

$$\mathbf{s}(i) = \mathbf{s}(i-1) + \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \boldsymbol{\mu}(i) \mathbf{e}(i) + \frac{\kappa}{2} \mathbf{T}_{\text{BI}}(i-1) \mathbf{Y}(i) \mathbf{Y}^+(i) \mathbf{Y}^T(i) \mathbf{f}_{\text{p}}(\mathbf{s}(i-1)) - \frac{\kappa}{2} \mathbf{f}_{\text{p}}(\mathbf{s}(i-1)) \quad (21)$$

where the diagonal matrix

$$\boldsymbol{\mu}(i) = \text{diag} [\mu_0(i), \mu_1(i), \dots, \mu_{p-1}(i)].$$

The *a posteriori* error vector is given by

$$\boldsymbol{\varepsilon}(i) = \mathbf{x}(i) - \mathbf{Y}^T(i) \mathbf{s}(i) \quad (22)$$

Substituting (21) in (22) and making use of the expression for $\mathbf{e}(i) = \mathbf{x}(i) - \mathbf{Y}^T(i) \mathbf{s}(i-1)$, it gives that

$$\boldsymbol{\varepsilon}(i) = [\mathbf{I}_p - \boldsymbol{\mu}(i)] \mathbf{e}(i) \quad (23)$$

In system identification or echo cancellation, the idea is to make the adaptive filter output cancel the unknown echo path output so that the *a posteriori* error is equal to the near end $\mathbf{a}(i)$. So, it is reasonable to have a condition $\boldsymbol{\varepsilon}(i) = \mathbf{a}(i)$.

Using the condition in (23), it results that

$$\varepsilon_{k+1}(i) = [1 - \mu_k(i)] e_{k+1}(i) = a(i-k) \quad (24)$$

with $k=0,1,\dots,p-1$. In matrix form, the condition is given by

$$\boldsymbol{\varepsilon}(i) = \text{diag}[1 - \mu_0(i), 1 - \mu_1(i), \dots, 1 - \mu_{p-1}(i)] \begin{bmatrix} e_1(i) \\ e_2(i) \\ \vdots \\ e_p(i) \end{bmatrix} \quad (25)$$

$$= \begin{bmatrix} a(i) \\ a(i-1) \\ \vdots \\ a(i-p+1) \end{bmatrix} \quad (26)$$

where $\boldsymbol{\varepsilon}(i) = [\varepsilon_1(i), \varepsilon_2(i), \dots, \varepsilon_p(i)]^T$

The objective is to find an expression for the variable step-size $\mu(i)$ such that $E(\varepsilon_{k+1}^2(i)) = E(v_{k+1}^2(i))$. Squaring (24) and taking expectation on both sides to get

$$[1 - \mu_k(i)]^2 E(e_{k+1}^2(i)) = E(a^2(i-k)) \quad (27)$$

Solving the quadratic Equation (27) gives two solutions for $\mu(k)$. From the analysis [25], it is reasonable to choose a smaller value to obtain a lower steady-state error without sacrificing the convergence rate.

$$\mu_k(i) = 1 - \frac{\sqrt{E(a^2(i-k))}}{\sqrt{E(e_{k+1}^2(i))}} \quad (28)$$

$$\mu_k(i) = 1 - \frac{\sigma_a(i-k)}{\sigma_{e_{k+1}}(i)} \quad (29)$$

Introducing a parameter τ in the denominator to avoid divide-by-zero error, thereby

$$\mu_k(i) = 1 - \frac{\sigma_a(i-k)}{\tau + \sigma_{e_{k+1}}(i)} \quad (30)$$

In case, the power becomes zero, τ avoids division by zero.

Using the analysis in [30], the noise power can be estimated indirectly as

$$\sigma_a^2(i) = \sigma_x^2(i) - \sigma_d^2(i) \quad (31)$$

where d is the output of the system to be identified.

Like in VSS- l_0 -BS-PAPA, the unknown variances $\sigma_x^2(i)$, $\sigma_d^2(i)$, $\sigma_{e_{k+1}}^2(i)$ can be computed from the following recursive equations for the VSS- l_0 -BS-IPAPA.

$$\sigma_x^2(i) = \rho\sigma_x^2(i-1) + (1-\rho)x^2(i) \quad (32)$$

$$\sigma_d^2(i) = \rho\sigma_d^2(i-1) + (1-\rho)d^2(i) \quad (33)$$

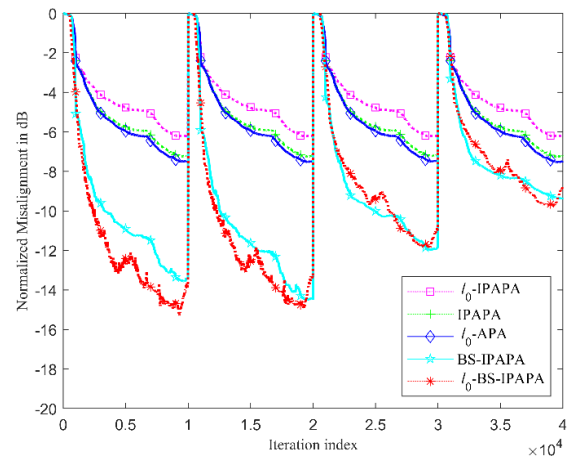
$$\sigma_{e_{k+1}}^2(i) = \rho\sigma_{e_{k+1}}^2(i-1) + (1-\rho)e_{k+1}^2(i) \quad (34)$$

where the positive parameter ρ is called the forgetting factor. It is always chosen close to 1 to retain a long history of past observations and to reduce the variance of parameter estimates.

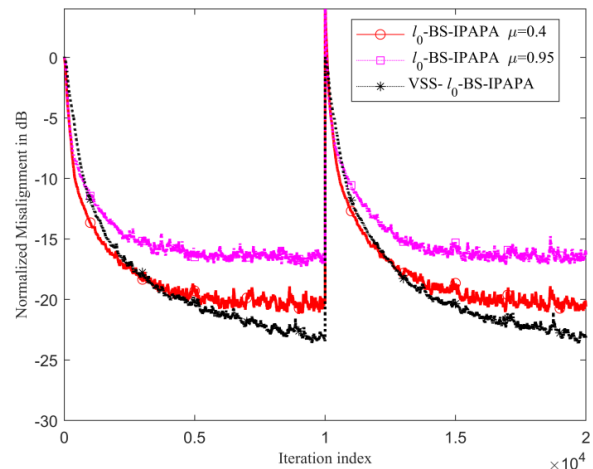
The performance advantage of the VSS- l_0 -BS-IPAPA over its fixed step-size counterpart is illustrated in Figure 8b. For this evaluation, the projection order p and the parameter τ are set to 2 and 1×10^{-4} , respectively,

while the forgetting factor is chosen as $\rho = 0.997$. All other simulation parameters remain consistent with those detailed in Section 5.3.

The results highlight the inherent limitations of the fixed step-size l_0 -BS-IPAPA: with $\mu = 0.4$, the algorithm achieves a low steady-state NM but suffers from a slower convergence rate. Conversely, when μ is increased to 0.95, the convergence speed improves significantly at the expense of a much higher steady-state misalignment floor. In contrast, the VSS- l_0 -BS-IPAPA successfully bridges this gap, providing a fast initial convergence comparable to the $\mu = 0.95$ case while maintaining a low steady-state NM similar to the $\mu = 0.4$ case. This confirms that the variable step-size mechanism effectively optimizes the algorithm's behavior throughout the different phases of the adaptation process.



(a) Varying Q



(b) Variable step-size

Figure 8. Performance evaluation of l_0 -BS-IPAPA for varying block sizes and step-sizes.

6. Transient Performance of the l_0 -BS-IPAPA against Different SNRs

The performance of the proposed l_0 -BS-IPAPA under varying noise conditions is illustrated in Figure 9. To evaluate the algorithm's robustness, the Mean Square Error (MSE), defined as $10\lg E|e(i)|^2$, is plotted for various SNRs. The excitation signal used is colored noise generated via a first-order Autoregressive AR(1) process with a pole at 0.8.

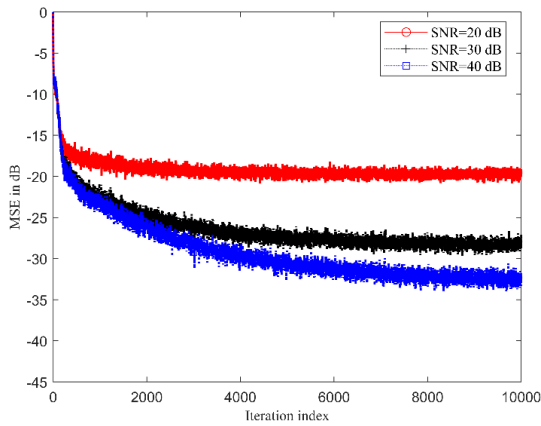


Figure 9. Learning curves of l_0 -BS-IPAPA for colored input ($\mu = 0.2$, $p=4$).

To simulate practical environments, the measurement noise $v(i)$ is modeled as zero-mean WGN with the variance σ_v^2 adjusted to achieve SNRs of 20, 30, and 40 dB. For these simulations, the step-size μ is fixed at 0.2, the projection order p is 4, and the regularization factor δ is 0.001. To ensure statistical reliability, the MSE curves are obtained by averaging over 500 independent Monte Carlo trials. As observed in Figure 9, the steady-state MSE of the l_0 -BS-IPAPA increases proportionally with the measurement noise level (i.e., as SNR decreases). Interestingly, the results also show that at lower SNR levels, the algorithm reaches its steady-state more rapidly, albeit at a significantly higher error floor. This behavior is consistent with adaptive filtering theory, where the presence of a higher noise floor effectively masks the smaller gradient updates during the final stages of convergence.

6.1. Computational Complexity

A comprehensive comparison of the computational requirements for the l_0 -BS-IPAPA and VSS- l_0 -BS-IPAPA and existing benchmarks is summarized in Table 2.

Table 2. Computational Complexity of l_0 -BS-IPAPA, VSS- l_0 -BS-IPAPA and Various Algorithms: Additions (A), Multiplications (MU), Divisions (D), Comparisons (C), Square Root (SQRT), Memory (ME) and Execution Time(E).

Algorithms	A	MU	D	C	SQRT	ME	E (ms)
IPAPA	$(p^2+p+1)L$	$(p^2+p+2)L + p^2$	2	0	0	L	29
l_0 -IPAPA	$(p^2+p+3)L-1$	$(p^2+p+5)L + p^2$	2	0	0	2L	30
BS-IPAPA	$(p^2+p)L+M-1$	$(p^2+p+3)L + p^2+M+2$	3	2M	M	L+M	33
l_0 -BS-IPAPA	$(p^2+p)L+2L + M-1$	$(p^2+p+6)L + p^2+M+2$	2	2M	M	2L+M	35
VSS- l_0 -BS-IPAPA	$(p^2+p)L+5L + M-1$	$(p^2+p+10)L + p^2+M+5$	3	2M	M+p	2L+M	36

The complexity is evaluated based on the number of arithmetic operations—additions, multiplications, divisions, comparisons, and square roots—as well as the required memory space per iteration to update the weight vector. At every iteration, to update $s(i)$ to $s(i+1)$, the proposed algorithm l_0 -BS-IPAPA with length L and projection order p requires $(p^2+p)L+2L+M-1$ additions, $(p^2+p+6)L+p^2+M+2$ multiplications, 2 divisions, 2M comparisons, M square roots, and 2L+M memory spaces. Compared to the BS-IPAPA, the l_0 -BS-IPAPA requires additional multiplications and additions for the computation of zero attraction term. It also needs extra N memory spaces for computing and storing the l_0 -norm of the filter coefficients. However, by using a ceiling function, the extra computations needed in l_0 -BS-IPAPA can be significantly minimized. The VSS- l_0 -BS-IPAPA introduces a marginal increase in complexity, requiring one additional division and p square roots to compute the time-varying step-size matrix. It should be noted that all affine-projection-based algorithms compared here require a direct matrix inversion of size $p \times p$, which is computationally manageable for the small projection orders (typically $p < 4$) used in echo cancellation. Projection order and block-size have a significant effect on computational complexity. $p=1$ would reduce the computational complexity significantly by behaving like an NLMS algorithm but lose the convergence rate for correlated input signals. Increasing p makes matrix inversion very expensive and memory usage costly. At the same time, increasing Q or decreasing M would

increase the number of multiplications within the norm calculation. If Q is made too large, the proposed algorithm behaves like an l_0 -PAPA and loses its ability to exploit block-sparsity. To measure the execution time in milliseconds, the algorithms were implemented in MATLAB R2020a using the tic-toc function, installed on a PC equipped with an Intel Core i7 Processor with 16 GB of RAM. The main limitations of the proposed algorithm are the increase in computational complexity as the projection order increases and an additional computational burden introduced by the l_0 -norm.

6.2. Experimental Validation of the l_0 -BS-IPAPA

To further prove the efficacy of the l_0 -BS-IPAPA, the algorithm was evaluated using acoustic signals obtained from the Microsoft Acoustic Echo Cancellation (AEC) Challenge dataset. The dataset provides realistic speech recordings and acoustic conditions ideal for evaluating the performance of adaptive algorithms under practical scenarios. Two WAV files namely, *farend_singletalk_lpb.wav* and *farend_singletalk_mic.wav* were used as the far_end signal and the echo signal, respectively.

The proposed algorithm was compared with the conventional BS-IPAPA under identical experimental conditions. The input signal y , echo x , algorithm parameters, and evaluation conditions were maintained the same for both algorithms to ensure a fair comparison.

The performance of the algorithms was evaluated using Echo Return Loss Enhancement (ERLE), which quantifies the amount of echo attenuation achieved by the adaptive echo canceller. A higher ERLE value indicates improved echo cancellation performance. On average, the BS-IPAPA achieves 16 dB while the proposed method attains 19.3 dB ERLE. An approximate 3dB improvement indicates that the proposed algorithm is better in identifying and exploiting the sparse characteristics of the acoustic echo path, thereby reducing the residual echo more effectively. The improved performance of the proposed algorithm can be attributed to the incorporation of the l_0 -norm, which promotes sparse coefficient adaptation and enhances the convergence behavior of the adaptive filter.

7. Conclusion and Future Work

In this paper, a novel l_0 -BS-IPAPA was proposed to enhance the identification of block-sparse systems. By integrating an l_0 -norm approximation into the block-sparse IPAPA framework, the algorithm leverages both the structural sparsity and the cluster-based nature of real-world impulse responses. A theoretical steady-state mean stability condition was derived. To further address the trade-off between convergence speed and steady-state precision, a variable step-size variant was developed. Simulation results across various excitation signals—including WGN, colored noise, and real-time speech—demonstrate that the proposed algorithm consistently outperforms state-of-the-art benchmarks such as IPAPA, l_0 -IPAPA, l_0 -APA, and BS-IPAPA. While the l_0 -norm term introduces a marginal increase in computational complexity, the resulting gains in convergence rate, steady-state misalignment, and tracking ability justify its use in high-performance applications.

Future work will involve a more exhaustive analysis of the l_0 -norm attraction intensity and its impact on the weight-error vector. Additionally, we aim to validate the algorithm in diverse real-world environments, such as acoustic echo cancellation, wireless channel estimation, and biomedical signal processing. Finally, the hardware feasibility of the l_0 -BS-IPAPA will be explored through Field-Programmable Gate Array (FPGA) or Application-Specific Integrated Circuit (ASIC) implementations to evaluate its suitability for real-time, low-power embedded systems.

Competing Interest Statement

The authors declare no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The MATLAB code, speech sample, WAV files and simulation data are available from the corresponding author on request (bsenthil24@gmail.com).

Author Contributions

Senthil Murugan Boopalan: conceptualization, methodology and design, software and coding, supervision, project administration, review, Krithiga Sukumaran: methodology and design, validation, formal analysis, review, Ashokaraju Devarajan: data curation, formal analysis, investigation, resources, review, Sarojini Raju: data curation, writing-original draft, writing - review and editing, formal analysis and interpretation of data, revising the draft, Kalphana Ilangoan: conceptualization, methodology and design, software and coding, data curation, formal analysis, investigation, resources, review, Janani Munisamy: software and coding, interpretation of data, revising the draft. All authors have read and approved the final draft of the manuscript.

References

- [1] K. Ozeki, *Theory of Affine Projection Algorithms for Adaptive Filtering*. New York, NY, USA: Springer, 2016.
- [2] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002.
- [3] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley, 2003.
- [4] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 132–135, Feb. 2004. <https://doi.org/10.1109/LSP.2003.821722>
- [5] D. L. Duttweiler, "Proportionate normalized least-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sep. 2000. <https://doi.org/10.1109/89.861368>
- [6] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electron. Commun. Jpn.*, vol. 67, no. 5, pp. 19–27, May 1984. <https://doi.org/10.1002/ecja.4400670503>
- [7] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE ICASSP*, Orlando, FL, USA, 2002, pp. 1881–1884. <https://doi.org/10.1109/ICASSP.2002.5744994>
- [8] J. Liu and S. L. Grant, "Proportionate adaptive filtering for block-sparse system identification," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 4, pp. 623–630, Apr. 2016. <https://doi.org/10.1109/TASLP.2015.2499602>
- [9] Y. Gu, J. Jin, and S. Mei, " ℓ_0 -norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009. <https://doi.org/10.1109/LSP.2009.2024736>
- [10] M. V. Lima, W. A. Martins, and P. S. R. Diniz, "Affine projection algorithms for sparse system identification," in *Proc. IEEE ICASSP*, Vancouver, BC, Canada, 2013, pp. 5666–5670. <https://doi.org/10.1109/ICASSP.2013.6638749>
- [11] A. W. H. Khong and P. A. Naylor, "Efficient use of sparse adaptive filters," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, 2006, pp. 1375–1379. <https://doi.org/10.1109/ACSSC.2006.354982>
- [12] J. Ni, Y. Xing, Z. Zhu, and J. Chen, "Proportionate total adaptive filtering algorithms for sparse system identification," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 54, no. 3, pp. 1840–1852, Mar. 2024. <https://doi.org/10.1109/TSMC.2023.3330125>
- [13] N. Chikyal, Vasundhara, C. Bhar, A. Kar, and M. G. Christensen, "Sparsity apprised logarithmic hyperbolic tan adaptive filters for nonlinear system identification and acoustic feedback cancellation," *IEEE Open J. Signal Process.*, vol. 6, pp. 1031–1047, 2025. <https://doi.org/10.1109/OJSP.2025.3600904>
- [14] M. Salman, A. A. F. Youssef, F. Elsayed, and M. Rashdan, "A sparse variable step-size least-mean-square algorithm for impulsive noise in a code-division multiple access system," *IEEE Access*, vol. 13, pp. 86026–86031, 2025. <https://doi.org/10.1109/ACCESS.2025.3570122>
- [15] M. Salman, H. Zayyani, F. A. P. De Figueiredo, H. Abuhilal, and M. Rashdan, "Dual-domain sparse adaptive filtering: Exploiting error memory for improved performance," *IEEE Access*, vol. 14, pp. 70188–70199, 2026. <https://doi.org/10.1109/ACCESS.2026.3686144>
- [16] K. Pelekanakis and M. Chitre, "New sparse adaptive algorithms based on the natural gradient and the ℓ_0 -norm," *IEEE J. Ocean. Eng.*, vol. 38, no. 2, pp. 323–332, Apr. 2013. <https://doi.org/10.1109/JOE.2012.2221811>
- [17] Z. Qu and Z. Zheng, "An efficient ℓ_0 -norm constraint memory improved proportionate affine projection algorithm," *Circuits Syst. Signal Process.*, vol. 36, no. 8, pp. 3448–3456, Aug. 2017. <https://doi.org/10.1007/s00034-016-0467-4>
- [18] J. Liu and S. L. Grant, "Proportionate affine projection algorithms for block-sparse system identification," in *Proc. IEEE ICASSP*, Shanghai, China, 2016, pp. 529–533. <https://doi.org/10.1109/ICASSP.2016.7471731>
- [19] Y. Li, Z. Jiang, O. M. O. Osman, X. Han, and J. Yin, "Mixed-norm constrained sparse APA algorithm for satellite and network echo channel estimation," *IEEE Access*, vol. 6, pp. 65901–65908, 2018. <https://doi.org/10.1109/ACCESS.2018.2878310>
- [20] S. M. Boopalan, S. Alagala, and A. Ramalingam, "A memory sparse proportionate affine projection algorithm for echo cancellation: Analysis and simulations," *Arab. J. Sci. Eng.*, vol. 47, no. 3, pp. 3367–3381, Mar. 2022. <https://doi.org/10.1007/s13369-021-06219-w>
- [21] S. M. Boopalan and S. Alagala, "A new affine projection algorithm with adaptive ℓ_0 -norm constraint for block-sparse system identification," *Circuits Syst. Signal Process.*, vol. 42, no. 3, pp. 1792–1807, Mar. 2023. <https://doi.org/10.1007/s00034-022-02197-y>

- [22] L. Luo, Y. Yu, T. Yang, et al., "Affine projection algorithms with novel schemes of variable projection order," *Circuits Syst. Signal Process.*, vol. 43, no. 12, pp. 8074–8090, Dec. 2024.
<https://doi.org/10.1007/s00034-024-02829-5>
- [23] Y. Zong and J. Ni, "Cluster-sparsity-induced affine projection algorithm and its variable step-size version," *Signal Process.*, vol. 195, Art. no. 108490, Jun. 2022.
<https://doi.org/10.1016/j.sigpro.2022.108490>
- [24] S. M. Boopalan, S. Raju, et al., "A memory improved proportionate affine projection algorithm for sparse system identification," *Int. J. Electr. Comput. Eng.*, vol. 15, no. 5, pp. 4605–4619, Oct. 2025.
<https://doi.org/10.11591/ijece.v15i5.pp4605-4619>
- [25] T. Günsler, J. Benesty, S. L. Gay, and M. M. Sondhi, "A robust proportionate affine projection algorithm for network echo cancellation," in *Proc. IEEE ICASSP*, Istanbul, Turkey, 2000, vol. 2, pp. 793–796.
<https://doi.org/10.1109/ICASSP.2000.859079>
- [26] S. G. Sankaran and A. A. L. Beex, "Convergence behavior of affine projection algorithms," *IEEE Trans. Signal Process.*, vol. 48, no. 4, pp. 1086–1096, Apr. 2000. <https://doi.org/10.1109/78.827542>
- [27] H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 90–102, Jan. 2004.
<https://doi.org/10.1109/TSP.2003.820077>
- [28] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data normalized adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, Mar. 2003.
<https://doi.org/10.1109/TSP.2002.808106>
- [29] C. Park, S. H. Ryu, and P. Park, "A variable step-size bias-compensated affine projection algorithm with noisy inputs," *J. Franklin Inst.*, vol. 362, no. 12, Art. no. 107792, Aug. 2025.
<https://doi.org/10.1016/j.jfranklin.2025.107792>
- [30] C. Paleologu, J. Benesty, and S. Ciochină, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp. 1466–1478, Oct. 2008.
<https://doi.org/10.1109/TASL.2008.2002980>