Spring 2016

Adaptive filters for sparse system identification

Jianming Liu

Department: Electrical and Computer Engineering

ADAPTIVE FILTERS FOR SPARSE SYSTEM IDENTIFICATION

by

JIANMING LIU

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

2016

Approved by

Steven L. Grant, Advisor
Daryl G. Beetner
Kurt Kosbar
Randy H. Moss
Matt Insall

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following six published or accepted papers, formatted in the style used by the Missouri University of Science and Technology, listed as follows:

Paper 1, J. Liu and S. L. Grant, "Proportionate adaptive filtering for block-sparse system identification," has been accepted by *IEEE/ACM Transactions on Audio, Speech and Language Processing*.

Paper 2, J. Liu and S. L. Grant, "Proportionate affine projection algorithms for block-sparse system identification," has been accepted by *IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016.

Paper 3, J. Liu and S. L. Grant, "Block sparse memory improved proportionate affine projection sign algorithm," has been published in *IET Electronics Letters*.

Paper 4, J. Liu, S. L. Grant and J. Benesty, "A low complexity reweighted proportionate affine projection algorithm with memory and row action projection," has been published in *EURASIP Journal on Advances in Signal Processing*.

Paper 5, J. Liu and S. L. Grant, "A new variable step-size zero-point attracting projection algorithm," has been published in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Nov. 2013.

Paper 6, J. Liu and S. L. Grant, "An improved variable step-size zero-point attracting projection algorithm," has been published in *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP15)*, Apr. 2015.

ABSTRACT

Sparse system identification has attracted much attention in the field of adaptive algorithms, and the adaptive filters for sparse system identification are studied.

Firstly, a new family of proportionate normalized least mean square (PNLMS) adaptive algorithms that improve the performance of identifying block-sparse systems is proposed. The main proposed algorithm, called block-sparse PNLMS (BS-PNLMS), is based on the optimization of a mixed $l_{2,1}$ norm of the adaptive filter's coefficients. A block-sparse improved PNLMS (BS-IPNLMS) is also derived for both sparse and dispersive impulse responses. Meanwhile, the proposed block-sparse proportionate idea has been extended to both the proportionate affine projection algorithm (PAPA) and the proportionate affine projection sign algorithm (PAPSA).

Secondly, a generalized scheme for a family of proportionate algorithms is also presented based on convex optimization. Then a novel low-complexity reweighted PAPA is derived from this generalized scheme which could achieve both better performance and lower complexity than previous ones. The sparseness of the channel is taken into account to improve the performance for dispersive system identification. Meanwhile, the memory of the filter's coefficients is combined with row action projections (RAP) to significantly reduce the computational complexity.

Finally, two variable step-size zero-point attracting projection (VSS-ZAP) algorithms for sparse system identification are proposed. The proposed VSS-ZAPs are based on the approximations of the difference between the sparseness measure of current filter coefficients and the real channel, which could gain lower steady-state misalignment and also track the change in the sparse system.

ACKNOWLEDGMENTS

I would like to thank Dr. Steven L. Grant, my advisor and the chairman of my doctoral committee. The present work would not have been completed without his great advice and insightful guidance. It has always been a pleasure working with him.

I also would like to express my sincere gratitude to the remainder of my doctoral committee, Drs. Daryl G. Beetner, Kurt Kosbar, Randy H. Moss and Matt Insall, for their precious time in examining this dissertation and their constructive suggestions. Without their fruitful discussions and excitations, this work could not be finished.

I truly appreciate the support of Missouri S&T Dissertation Completion Fellowship Program, which encourages timely completion of my doctoral degree.

I would like to thank my parents, Xueyu Liu and Yufen Li for their continuous encouragement and support. Particularly, this dissertation is dedicated to my wife, Hui Geng, for her great love and support, and our sweet little girl, Anya G. Liu, who gives us a lot fun after the hard work.

TABLE OF CONTENTS

Page

PUBLICATION DISSERTATION OPTION iii

ABSTRACT iv

ACKNOWLEDGMENTS v

LIST OF ILLUSTRATIONS x

LIST OF TABLES xii

SECTION 1

1. INTRODUCTION 1

1.1. BACKGROUND 1

1.2. PROBLEM STATEMENT 2

1.3. SUMMARY OF CONTRIBUTIONS 4

PAPER 5

I. PROPORTIONATE ADAPTIVE FILTERING FOR BLOCK-SPARSE SYSTEM IDENTIFICATION 5

Abstract 5

1. INTRODUCTION 6

2. REVIEW OF PNLMS 9

3. PROPOSED BS-PNLMS 11

3.1 MOTIVATION OF THE PROPOSED BS-PNLMS 11

3.2 THE PROPOSED BS-PNLMS ALGORITHM 13

3.3 EXTENSION TO THE BS-IPNLMS ALGORITHM 15

4. COMPUTATIONAL COMPLEXITY 17

5. SIMULATION RESULTS 18

5.1 EFFECT OF *P* ON THE PERFORMANCE OF BS-PNLMS 18

5.2 CONVERGENCE PERFORMANCE OF BS-PNLMS AND BS-IPNLMS FOR BLOCK-SPARSE SYSTEMS 19

5.3 CONVERGENCE PERFORMANCE OF BS-PNLMS AND BS-IPNLMS FOR THE ACOUSTIC ECHO PATH AND A RANDOM DISPERSIVE SYSTEM 23

6. CONCLUSION AND FUTURE WORK 27

7. ACKNOWLEDGEMENT 28

8. REFERENCES 29

II. PROPORTIONATE AFFINE PROJECTION ALGORITHMS FOR BLOCK-SPARSE SYSTEM IDENTIFICATION 32

Abstract 32

1. INTRODUCTION 33

2. REVIEW OF PAPA 34

3. PROPOSED BS-PAPA 36

3.1 THE PROPOSED BS-PAPA 36

3.2 EFFICIENT IMPLEMENTATION OF PROPOSED BS-PAPA 38

3.3 MEMORY BS-PAPA 39

4. SIMULATION RESULTS 40

5. CONCLUSION 43

6. REFERENCES 44

III. BLOCK SPARSE MEMORY IMPROVED PROPORTIONATE AFFINE PROJECTION SIGN ALGORITHM 45

Abstract 45

1. INTRODUCTION 46

2. REVIEW OF MIP-APSA 47

3. ALGORITHM DESIGN 49

4. COMPLEXITY 51

5. SIMULATION RESULTS 52

6. CONCLUSION 55

7. REFERENCES 56

IV. A LOW COMPLEXITY REWEIGHTED PROPORTIONATE AFFINE PROJECTION ALGORITHM WITH MEMORY AND ROW ACTION PROJECTION 57

Abstract 57

1. INTRODUCTION 58

2. REVIEW OF VARIOUS PAPAS 60

3. THE PROPOSED SC-RPAPA WITH MRAP 63

3.1 THE PROPOSED RPAPA 63

3.2 THE PROPOSED SC-RPAPA 65

3.3 THE PROPOSED SC-RPAPA WITH MRAP 67
4. COMPUTATIONAL COMPLEXITY 71
5. SIMULATION RESULTS 73
5.1 THE PERFORMANCE OF THE PROPOSED RPAPA 73
5.2 THE PERFORMANCE OF THE PROPOSED SC-RPAPA 75
5.3 THE PERFORMANCE OF THE PROPOSED SC-RPAPA WITH MRAP ... 77
6. CONCLUSION 80
7. REFERENCES 81
V. A NEW VARIABLE STEP-SIZE ZERO-POINT ATTRACTING PROJECTION ALGORITHM 85
Abstract 85
1. INTRODUCTION 86
2. REVIEW OF VSS ZAP 87
3. PROPOSED VSS ZAP 89
4. SIMULATION RESULTS 91
5. CONCLUSION 98
6. REFERENCES 99
VI. AN IMPROVED VARIABLE STEP-SIZE ZERO-POINT ATTRACTING PROJECTION ALGORITHM 101
Abstract 101
1. INTRODUCTION 102
2. REVIEW OF VSS ZAP 104
2.1 INTRODUCTION TO ZAP 104
2.2 REVIEW OF VARIABLE STEP-SIZE ZAP ALGORITHMS 104
3. PROPOSED VSS ZA-LMS 106
3.1 THE PROPOSED SCHEME OF VARIABLE STEP-SIZE ZAP 106
3.2 IMPROVED VARIABLE STEP-SIZE ZAP FOR BOTH SPARSE AND NON-SPARSE SYSTEM 107
4. SIMULATION RESULTS 109
5. CONCLUSION 112
6. REFERENCES 113
SECTION 115

2. CONCLUSIONS 115
3. PUBLICATIONS 116
BIBLIOGRAPHY 118
VITA 119

LIST OF ILLUSTRATIONS

Figure Page

PAPER I

1. Three types of sparse systems. 12
2. Comparison of the BS-PNLMS algorithms with different group sizes for block-sparse systems at SNR=30dB. 20
3. Comparison of NLMS, PNLMS, IPNLMS, BS-PNLMS and BS-IPNLMS algorithms for block-sparse systems at SNR=30dB . 22
4. Two impulse responses. 24
5. Comparison of NLMS, PNLMS, IPNLMS, BS-PNLMS and BS-IPNLMS algorithms for acoustic echo path and dispersive system in Figure 4 and SNR=30dB . 24

PAPER II

1. Block-sparse impulse systems. 41
2. Comparison of BS-PAPA with different group sizes for colored input with *SNR*=30dB. 42
3. Comparison of APA, PAPA, MPAPA, BS-PAPA and BS-MPAPA algorithms for colored noise with *SNR*=30dB. 42

PAPER III

1. Two block-sparse systems used in the simulations. 53
2. Normalized misalignment of APSA, MIP-APSA, and BS-MIP-APSA for colored input signal. 53
3. Normalized misalignment of APSA, MIP-APSA, and BS-MIP-APSA for speech input signal. 54

PAPER IV

1. Comparison of the different metrics. 65
2. Reweighted metric with different σ parameters. 66
3. Two impulse responses used in the simulation 74
4. Comparison of RPAPA with PAPA, l_0 PAPA and mu-law PAPA for WGN input, SNR=30 dB, $M=2$, $\mu=0.2$ 74
5. Comparison of RPAPA with PAPA, l_0 PAPA and mu-law PAPA for colored input, SNR=30 dB, $M=2$, $\mu=0.2$ 75
6. Comparison of SC-RPAPA with APA, PAPA, and RPAPA for WGN input, SNR=30 dB, $M=2$, $\mu=0.2$ 76

7. Comparison of SC-RPAPA with APA, PAPA, and RPAPA for colored input, SNR=30 dB, $M=2$, $\mu=0.2$ 76

8. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for WGN input, SNR=30 dB, $M=2$, $\mu=0.2$ 78

9. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for colored input, SNR=30 dB, $M=2$, $\mu=0.2$ 78

10. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for WGN input, SNR=30 dB, $M=32$, $\mu=0.2$ 79

11. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for colored input, SNR=30 dB, $M=32$, $\mu=0.2$ 79

PAPER V

1. Sparse impulse response. 92

2. Dispersive random impulse response. 92

3. Comparison of normalized misalignment for l_1 norm constraint ZAP under sparse system. 93

4. Comparison of variable step-size for l_1 norm constraint ZAP under sparse system. 93

5. Comparison of normalized misalignment for l_0 norm constraint ZAP under sparse system. 94

6. Comparison of variable step-size for l_0 norm constraint ZAP under sparse system. 94

7. Comparison of normalized misalignment for l_1 norm constraint ZAP under dispersive system. 95

8. Comparison of variable step-size for l_1 norm constraint ZAP under dispersive system. 96

9. Comparison of normalized misalignment for l_0 norm constraint ZAP under dispersive system. 96

10. Comparison of variable step-size for l_0 norm constraint ZAP under dispersive system. 97

PAPER VI

1. Performance demonstration of approximation $\mathrm{sgn}(\boldsymbol{h}(n)) = \mathrm{sgn}(\boldsymbol{w}(n))$ in (10). 110

2. Comparison of normalized misalignment for sparse system identification. 111

3. Comparison of normalized misalignment for dispersive system identification. 111

LIST OF TABLES

Table Page

PAPER I

1. The block-sparse algorithms 16
2. Computational complexity of the algorithms' coefficient updates – Addition (A), Multiplication (M), Division (D), Comparison (C), Square Root (Sqrt) and Memory Word (MW). 17

PAPER IV

1. The SC-RPAPA algorithm with MRAP 70
2. Computational complexity of the algorithms' coefficient updates. 72

PAPER VI

1. Sparseness measures in [12] 90

1. INTRODUCTION

1.1. BACKGROUND

Sparse system identification has attracted much attention in the field of adaptive algorithms. A sparse impulse response is that in which a large percentage of the energy is distributed to only a few coefficients of its impulse response [1]. In the last decade, sparse system identification has been widely applied in many signal processing applications: echo cancellation, radar imaging, wireless communication, etc.

To improve on the convergence performance of normalized least mean squares (NLMS) and affine projection algorithm (APA), the proportionate NLMS (PNLMS) and proportionate APA (PAPA) algorithms exploit the sparseness of a given system [2]-[3]. The idea behind proportionate algorithms is to update each coefficient of the filter independently by adjusting the adaptation step size in proportion to the estimated filter's coefficients. In comparison to NLMS and APA, the PNLMS and PAPA have very fast initial convergence and tracking when the echo path is sparse. Recently, it was shown that both PNLMS and PAPA can be deduced from a *basis pursuit* perspective [4]-[5].

A special family of sparse system, called the block-sparse system, is very common in the real applications, such as network echo cancellation (NEC) and satellite-linked communications etc. However, the traditional PNLMS and PAPA do not take this point into account. Considering the, it is necessary to further improve the proportionate algorithm by exploiting this special block-sparse characteristic of the sparse impulse response.

Besides to the family of proportionate algorithms, the family of zero-point attracting projection (ZAP) algorithms has been recently proposed to solve the sparse system identification problem [6]-[7]. When the solution is sparse, the gradient descent recursion will accelerate the convergence of near-zero coefficients of the sparse system.

The ZAP algorithm applied the sparseness constraint to the standard LMS cost function and when the solution is sparse, the gradient descent recursion will accelerate the convergence of near-zero coefficients of the sparse system. Analysis showed that the step-size of the ZAP term denotes the importance or the intensity of attraction. A large step-size for ZAP results in a faster convergence, but the steady-state misalignment also increases with the step-size [8]. So, the step-size of ZAP is also a trade-off between convergence rate and steady-state misalignment.

1.2. PROBLEM STATEMENT

The input signal $\mathbf{x}(n)$ is filtered through the unknown coefficients $\mathbf{h}(n)$ to get the observed output signal $d(n)$

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n), \tag{1.1}$$

where $v(n)$ is the measurement noise, and L is the length of the impulse response. We define the estimated error as

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \tag{1.2}$$

where $\hat{\mathbf{h}}(n)$ is the adaptive filter's coefficients. The NLMS algorithm updates the filter coefficients as below [1]:

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mu\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{x}(n)+\delta}, \tag{1.3}$$

in which μ is the step-size of adaption and δ is the regularization parameter. The family of PNLMS algorithm can be described as below [4]:

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mu\mathbf{G}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{G}(n-1)\mathbf{x}(n)+\delta}, \tag{1.4}$$

where

$$\mathbf{G}(n-1) = diag\left[g_1(n-1), g_2(n-1), \cdots, g_L(n-1)\right], \tag{1.5}$$

$$g_l(n-1) = \frac{\gamma_l(n-1)}{\frac{1}{L}\sum_{i=1}^{L}\gamma_i(n-1)}, \tag{1.6}$$

$$\gamma_l = \max\left\{\rho\max\left\{q, \left|\hat{h}_1\right|, \cdots, \left|\hat{h}_L\right|\right\}, \left|\hat{h}_l\right|\right\}, \tag{1.7}$$

q prevents the filter coefficients from stalling when $\hat{\mathbf{h}}(0)=\mathbf{0}_{L\times 1}$ at initialization, and ρ prevents the coefficients from stalling when they are much smaller than the largest coefficient.

Meanwhile, grouping the M most recent input vectors together gives the input signal matrix:

$$\mathbf{X}(n)=[\mathbf{x}(n),\mathbf{x}(n-1),\cdots,\mathbf{x}(n-M+1)].$$

Therefore, the estimated error vector is

$$\mathbf{e}(n)=\mathbf{d}(n)-\mathbf{X}^{T}(n)\hat{\mathbf{h}}(n-1), \tag{1.8}$$

in which

$$\mathbf{d}(n)=[d(n),d(n-1),\cdots,d(n-M+1)], \tag{1.9}$$

$$\mathbf{e}(n)=[e(n),e(n-1),\cdots e(n-M+1)], \tag{1.10}$$

and M is the projection order. The PAPA algorithm updates the filter coefficients as follows [5]:

$$\begin{aligned}\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\\ \mu\mathbf{G}(n-1)\mathbf{X}(n)\left(\mathbf{X}(n)\mathbf{G}(n-1)\mathbf{X}(n)+\delta\mathbf{I}_{M}\right)^{-1}\mathbf{e}(n),\end{aligned} \tag{1.11}$$

in which $\mathbf{I}_M$ is $M\times M$ identity matrix.

The ZA-LMS algorithm with l_1 norm constraint updates its coefficients as [6]

$$\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\mu\mathbf{x}e(n)-\kappa\,\mathrm{sgn}\left(\hat{\mathbf{h}}(n-1)\right), \tag{1.12}$$

in which κ is the step-size of zero attractor, and $\mathrm{sgn}(\cdot)$ is a component-wise sign function defined as

$$\mathrm{sgn}(x)=\begin{cases}\dfrac{x}{|x|}, & x\neq 0;\\ 0, & \textit{elsewhere.}\end{cases} \tag{1.13}$$

1.3. SUMMARY OF CONTRIBUTIONS

Finally, we will briefly summarize the contributions we have made in this dissertation as below.

A new family of proportionate adaptive algorithms that improve the performance of identifying block-sparse systems is proposed. The main proposed algorithm, called block-sparse PNLMS (BS-PNLMS), is based on the optimization of a mixed $l_{2,1}$ norm of the adaptive filter's coefficients. It is demonstrated that both NLMS and traditional PNLMS are special cases of BS-PNLMS. Meanwhile, this block-sparse idea has been applied to improved PNLMS (IPNLMS), PAPA and proportionate affine sign algorithm (PAPSA) too.

A general framework is proposed to derive proportionate adaptive algorithms for sparse system identification. The proposed algorithmic framework employs the convex optimization and covers many traditional proportionate algorithms. Meanwhile, based on this framework, a novel reweighted proportionate algorithm is derived to achieve both better performance and lower computational complexity.

Finally, an improved variable step-size (VSS) scheme for zero-point attracting projection (ZAP) algorithm is presented. The proposed VSS ZAP is proportional to the sparseness difference between filter coefficients and the true impulse response. Meanwhile, it works for both sparse and non-sparse system identification.

PAPER

I. PROPORTIONATE ADAPTIVE FILTERING FOR BLOCK-SPARSE SYSTEM IDENTIFICATION

Jianming Liu and Steven L. Grant

Abstract

In this paper, a new family of proportionate normalized least mean square (PNLMS) adaptive algorithms that improve the performance of identifying block-sparse systems is proposed. The main proposed algorithm, called block-sparse PNLMS (BS-PNLMS), is based on the optimization of a mixed $l_{2,1}$ norm of the adaptive filter's coefficients. It is demonstrated that both the NLMS and the traditional PNLMS are special cases of BS-PNLMS. Meanwhile, a block-sparse improved PNLMS (BS-IPNLMS) is also derived for both sparse and dispersive impulse responses. Simulation results demonstrate that the proposed BS-PNLMS and BS-IPNLMS algorithms outperformed the NLMS, PNLMS and IPNLMS algorithms with only a modest increase in computational complexity.

1. INTRODUCTION

Sparse system identification has attracted much attention in the field of adaptive algorithms. The family of proportionate algorithms exploits this sparseness of a given system to improve the convergence performance of normalized least mean square (NLMS) [1]-[13] and is widely used in network echo cancellation (NEC), etc.

The idea behind PNLMS is to update each coefficient of the filter independently by adjusting the adaptation step size in proportionate to the estimated filter's coefficient [2]. The proportionate NLMS (PNLMS), as compared to the NLMS, has very fast initial convergence and tracking when the echo path is sparse. However, large coefficients converge quickly (fast initial convergence) at the cost of dramatically slowing the convergence of the small coefficients (after the initial period) [3]-[4]. As the large taps adapt, the remaining small coefficients adapt at a rate slower than NLMS.

The mu-law PNLMS (MPNLMS) algorithm proposed in [3]-[4] addresses the issue of assigning too large of an update gain to the large coefficients. The total number of iterations for overall convergence is minimized when all of the coefficients reach the ε-vicinity of their true values simultaneously (where ε is some small positive number). The PNLMS (EPNLMS) algorithm is the second implementation of the same philosophy used to generate the MPNLMS algorithm [5]. The EPNLMS algorithm gives the minimum gain possible to all of the coefficients with a magnitude less than ε. This is based on the assumption that the impulse response is sparse and contains many small magnitude coefficients. However, the MPNLMS algorithm's performance is more robust than the EPNLMS algorithm regarding the choice of algorithm parameters, as well as input signal and unknown system characteristics [1]. Furthermore, the l_0 norm family algorithms have recently become popular for sparse system identification. A new PNLMS algorithm based on the l_0 norm was proposed to represent a better measure of sparseness than the l_1 norm in a PNLMS-type algorithm [6]. Benesty demonstrated that PNLMS could be deduced from a basis pursuit perspective [7]. A more general framework was further proposed to derive proportionate adaptive algorithms for sparse system identification, which employed convex optimization [8].

In many simulations, however, it seems that we fully benefit from PNLMS only when the impulse response is close to a delta function [9]. Indeed, PNLMS converges much slower than NLMS when the impulse response is dispersive. The PNLMS++ algorithm, which achieves improved convergence by alternating between NLMS and PNLMS each sample period, was proposed in an attempt to address this problem [9]. The improved PNLMS (IPNLMS) was proposed to exploit the “proportionate” idea by introducing a controlled mixture of proportionate (PNLMS) and non-proportionate (NLMS) adaptations [10]. The IPNLMS algorithm performs better than both the NLMS and the PNLMS algorithms regardless of the impulse response’s nature. The improved IPNLMS (IIPNLMS) algorithm was proposed to identify active and inactive regions of the echo path impulse response [11]. Active regions receive updates that are more in-line with NLMS, while inactive regions received gains based upon PNLMS. Meanwhile, a partitioned block improved proportionate NLMS (PB-IPNLMS) algorithm exploits the properties of an acoustic enclosure where the early path (i.e., direct path and early reflections) of the acoustic echo path is sparse and the late reverberant part of the acoustic path is dispersive [12]. The PB-IPNLMS consists of two time-domain partitioned blocks, such that different adaptive algorithms can be used for each part.

The standard PNLMS algorithm performance depends on some predefined parameters controlling proportionality through a minimum gain that is common for all of the coefficients. The individual activation factor PNLMS (IAF-PNLMS) algorithm was proposed to use a separate time varying minimum gain for each coefficient, which is computed in terms of both the past and the current values of the corresponding coefficient magnitude, and does not rely on either the proportionality or the initialization parameters [13].

The family of zero-point attracting projection (ZAP) algorithms was recently proposed to solve the sparse system identification problem [14]-[17]. When the solution is sparse, the gradient descent recursion will accelerate the convergence of the sparse system’s near-zero coefficients. A block-sparsity-induced adaptive filter, called block-sparse LMS (BS-LMS), was recently proposed to improve the identification of block-sparse systems [18]. The basis of BS-LMS is to insert a penalty of block-sparsity (a mixed

$l_{2,0}$ norm of adaptive tap-weights with equal group partition sizes) into the cost function of the traditional LMS algorithm.

A family of proportionate algorithms is proposed here for block-sparse system identification, which can achieve faster convergence in the block-sparse application. Both the classical NLMS and the PNLMS algorithms are special cases of this proposed scheme. The computational complexities of the proposed BS-PNLMS and BS-IPNLMS algorithms are also compared to NLMS, PNLMS, and IPNLMS algorithms.

2. REVIEW OF PNLMS

The input signal $\mathbf{x}(n)$ is filtered through the unknown coefficients, $\mathbf{h}(n)$, so that the observed output signal $d(n)$ can be obtained as

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n), \tag{1}$$

where

$$\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-L+1)]^T,\ \mathbf{h}(n) = [h_1(n), h_2(n), \cdots, h_L(n)]^T,$$

$v(n)$ is the measurement noise, and L is the length of the impulse response. The estimated error is defined as

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \tag{2}$$

where $\hat{\mathbf{h}}(n)$ is the adaptive filter's coefficients.

The coefficient update of the family of PNLMS algorithms is [2]:

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mu \mathbf{G}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{G}(n-1)\mathbf{x}(n) + \delta}, \tag{3}$$

where μ is the step-size, δ is the regularization parameter, and

$$\mathbf{G}(n-1) = diag\left[g_1(n-1), g_2(n-1), \cdots, g_L(n-1)\right]. \tag{4}$$

It should be noted that the step-size for the NLMS is the same for all filter coefficients: $\mathbf{G}(n-1) = \mathbf{I}_{L\times L}$, where $\mathbf{I}_{L\times L}$ is an $L\times L$ identity matrix. Meanwhile, the matrix for the family of PNLMS is defined as

$$g_l(n-1) = \frac{\gamma_l(n-1)}{\frac{1}{L}\sum_{i=1}^{L}\gamma_i(n-1)}, \tag{5}$$

where

$$\gamma_l = \max\left\{\rho \max\left\{q, \mathrm{F}\left(\left|\hat{h}_1\right|\right), \cdots, \mathrm{F}\left(\left|\hat{h}_L\right|\right)\right\}, \mathrm{F}\left(\left|\hat{h}_l\right|\right)\right\}, \tag{6}$$

$\mathrm{F}\left(\left|\hat{h}_l\right|\right)$ is specific to the algorithm, q is a small positive value that prevents the filter coefficients $\hat{h}_l(n-1)$ from stalling when $\hat{\mathbf{h}}(0)=\mathbf{0}_{L\times 1}$ at initialization, and ρ, another small positive value, prevents the coefficients from stalling when they are much smaller than the largest coefficient [1]. The classical PNLMS employs step-sizes that are proportional to the magnitude of the estimated impulse response [2],

$$\mathrm{F}\left(\left|\hat{h}_l(n-1)\right|\right)=\left|\hat{h}_l(n-1)\right|. \tag{7}$$

Instead of (5) and (6), the improved PNLMS (IPNLMS) algorithm proposed to use [10]

$$\gamma_l=(1-\alpha)\frac{\sum_{i=1}^{L}\left|\hat{h}_i\right|}{L}+(1+\alpha)\left|\hat{h}_l\right|, \tag{8}$$

and

$$g_l(n-1)=\frac{\gamma_l(n-1)}{\sum_{i=1}^{L}\gamma_i(n-1)}=\frac{(1-\alpha)}{2L}+\frac{(1+\alpha)\left|\hat{h}_l\right|}{2\sum_{i=1}^{L}\left|\hat{h}_i\right|}, \tag{9}$$

where $-1\le\alpha<1$. IPNLMS behaves like NLMS when $\alpha=-1$ and PNLMS for α close to 1. In general, IPNLMS is a sum of two terms. The first term is an average of the absolute value of the coefficients taken from the estimated filter and the second is the absolute value of the coefficient itself. For most AEC/NEC applications, a good choice is $\alpha=0,-0.5$, with which IPNLMS behaves better than either the NLMS or the PNLMS, regardless of the impulse response nature [10].

In next section, we will show that NLMS and PNLMS are all special cases of our proposed block-sparse PNLMS (BS-PNLMS). Meanwhile, we could further take advantage of the benefits of IPNLMS algorithms to improve the performance of the proposed BS-PNLMS algorithm.

3. PROPOSED BS-PNLMS

The motivation behind the proposed family of the block-sparse proportionate algorithms is discussed at the beginning of this section, and then the proposed BS-PNLMS and BS-IPNLMS algorithms are presented next.

3.1 MOTIVATION OF THE PROPOSED BS-PNLMS

A sparse impulse response is that in which a large percentage of the energy is distributed to only a few coefficients [1]. Several different types of sparse systems exist as indicated in Figure 1. The nonzero coefficients in a general sparse system (see Figure 1(a)) may be arbitrarily located. Meanwhile, there exists a special family known as either clustering-sparse systems or block-sparse systems [18]. For example, the network echo path is typically characterized by a bulk delay that is dependent on network loading, encoding, and jitter buffer delays. This results in an "active" region in the range of 8-12 *ms* duration, and the impulse response is dominated by "inactive" regions where coefficient magnitudes are close to zero [1]. The network echo response is a typical single-clustering sparse system (see Figure 1(b)). Satellite communication is an important modern application of echo cancellation. The impulse response of the echo path in satellite-linked communications consists of several long flat delay regions and disperse active regions. Such responses are representative of multi-clustering sparse systems. The waveform in a communication link that uses single-side band suppressed carrier modulation, contains both a relatively large near-end echo, characterized by a short time delay and a far-end echo that is smaller in amplitude but with a longer delay [20]. Therefore, the echo path impulse response is primarily characterized by two active regions that correspond to the near-end signal and the far-end signal echo (see Figure 1(c)). Considering the block-sparse characteristic of the sparse impulse responses, as in Figure 1(b) and Figure 1(c), the proportionate algorithm can be further improved by exploiting this special characteristic.

It can be observed that an echo path, such as Figure 1(b), consists of the direct path and a few early reflections, which are almost always sparse, and the late reverberant part, which is always dispersive. The PB-IPNLMS algorithm splits the impulse response into

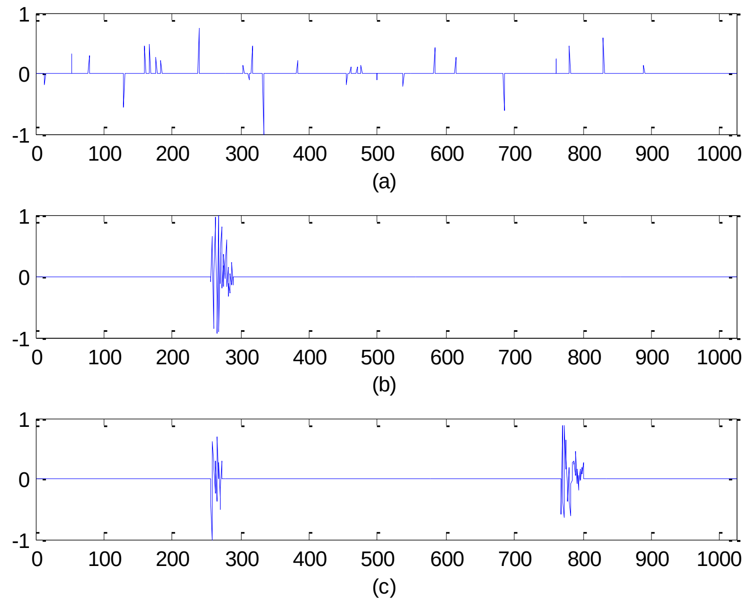


Figure 1. Three types of sparse systems, (a) a general sparse system, (b) a one-cluster block-sparse system, and (c) a two-cluster block-sparse system.

two blocks and used two IPNLMS algorithms each with a different proportionate/non-proportionate factor for the two corresponding time-domain partitioned blocks [12].

However, the PB-IPNLMS in [12] depends on the assumption of one-cluster sparse system, which does not hold for the multi-clustering case as in Figure 1(c). Additional IPNLMS algorithms could be employed to extend the PB-IPNLMS to multi-cluster sparse system. However, this must depend on the priori information of the bulk delays in the multi-cluster sparse system, which is not necessarily the case in practice.

P. Loganathan et al. in [12] noted that distributing almost equal step-sizes for the dispersive block provides better steady-state performance, which agrees with the well-known fact that for the dispersive system, NLMS is preferred over PNLMS. Meanwhile, PNLMS is only beneficial when the impulse response is close to a delta function [9]. Therefore, the block-sparse proportionate NLMS (BS-PNLMS) algorithm is proposed to accelerate the convergence by combining the above two facts together. In BS-PNLMS, considering the fact that the block-sparse system is dispersive within each block, it is preferred to use NLMS within each block. Meanwhile, the idea of PNLMS can be applied to have the NLMS step-size for each block proportionate to its relative magnitude. More details are given in the following subsection.

3.2 THE PROPOSED BS-PNLMS ALGORITHM

The proportionate NLMS algorithm can be deduced from a *basis pursuit* perspective [7]

$$\begin{aligned} &\boldsymbol{min} && \left\|\tilde{\mathbf{h}}(n)\right\|_1 \\ &\text{subject to} && d(n)=\mathbf{x}^T(n)\tilde{\mathbf{h}}(n), \end{aligned} \tag{10}$$

where $\tilde{\mathbf{h}}(n)$ is the correction component defined as [7]

$$\tilde{\mathbf{h}}(n)=\frac{\mathbf{G}(n-1)\mathbf{x}(n)d(n)}{\mathbf{x}^T(n)\mathbf{G}(n-1)\mathbf{x}(n)}.$$

Motivated by the observations in Section III.A, a family of proportionate adaptive algorithm for block-sparse system identification can be derived by replacing the l_1 norm optimization target in (10) with the following $l_{2,1}$ norm defined as

$$\left\|\tilde{\mathbf{h}}\right\|_{2,1}=\left\|\begin{bmatrix}\left\|\tilde{\mathbf{h}}_{[1]}\right\|_2\\ \left\|\tilde{\mathbf{h}}_{[2]}\right\|_2\\ \vdots\\ \left\|\tilde{\mathbf{h}}_{[N]}\right\|_2\end{bmatrix}\right\|_1=\sum_{i=1}^{N}\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2, \tag{11}$$

where $\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2=\sqrt{\tilde{\mathbf{h}}_{[i]}^T\tilde{\mathbf{h}}_{[i]}}$, $\tilde{\mathbf{h}}_{[i]}=[\tilde{h}_{(i-1)P+1},\tilde{h}_{(i-1)P+2},\cdots,\tilde{h}_{iP}]^T$, P is a predefined group partition size parameter and $N=L/P$ is the number of groups. The following convex target could be minimized with a constraint on the linear system of equations:

$$\begin{aligned} &\boldsymbol{min} && \left\|\tilde{\mathbf{h}}(n)\right\|_{2,1} \\ &\text{subject to} && d(n)=\mathbf{x}^T(n)\tilde{\mathbf{h}}(n). \end{aligned} \tag{12}$$

The Lagrange multiplier can be used to derive the proposed block-sparse proportionate NLMS algorithm [6]-[7]. The derivative of the $l_{2,1}$ norm in (11), with respect to the weight vector, is

$$\frac{\partial\left\|\tilde{\mathbf{h}}(n)\right\|_{2,1}}{\partial\tilde{\mathbf{h}}(n)}=\left[\frac{\partial\sum_{i=1}^{N}\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2}{\partial\tilde{h}_1},\frac{\partial\sum_{i=1}^{N}\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2}{\partial\tilde{h}_2},\cdots,\frac{\partial\sum_{i=1}^{N}\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2}{\partial\tilde{h}_L}\right], \tag{13}$$

in which

$$\frac{\partial\sum_{i=1}^{N}\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2}{\partial\tilde{h}_k}=\frac{\partial\left\|\tilde{\mathbf{h}}_{[j]}\right\|_2}{\partial\tilde{h}_k}=\frac{\tilde{h}_k}{\left\|\tilde{\mathbf{h}}_{[j]}\right\|_2}, \quad (j-1)P+1\le k\le jP. \tag{14}$$

The update equation for the proposed BS-PNLMS is

$$\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\frac{\mu\mathbf{G}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{G}(n-1)\mathbf{x}(n)+\delta}, \tag{15}$$

where

$$\mathbf{G}(n-1)= diag\left[\left\|\hat{\mathbf{h}}_{[1]}\right\|_2\mathbf{1}_P,\left\|\hat{\mathbf{h}}_{[2]}\right\|_2\mathbf{1}_P,\cdots,\left\|\hat{\mathbf{h}}_{[N]}\right\|_2\mathbf{1}_P\right], \tag{16}$$

and $-1\le\alpha<1$ is a P-length row vector of all ones. Equation (15) is the same as the traditional PNLMS, except that here the block-sparse definition of $\mathbf{G}(n-1)$ is used in (16). In a manner similar to (4)-(6) in PNLMS to prevent stalling issues, the proposed BS-PNLMS does so as

$$\mathbf{G}(n-1)= diag\left[g_1(n-1)\mathbf{1}_P,g_2(n-1)\mathbf{1}_P,\cdots,g_N(n-1)\mathbf{1}_P\right], \tag{17}$$

where

$$g_i(n-1)=\frac{\gamma_i}{\frac{1}{N}\sum_{l=1}^{N}\gamma_l}, \tag{18}$$

and

$$\gamma_i = \max\left\{\rho \max\left\{q, \left\|\hat{\mathbf{h}}_{[1]}\right\|_2, \cdots, \left\|\hat{\mathbf{h}}_{[N]}\right\|_2\right\}, \left\|\hat{\mathbf{h}}_{[i]}\right\|_2\right\}. \tag{19}$$

The traditional PNLMS and NLMS algorithms can each be easily verified as special cases of the proposed BS-PNLMS. If P is equal to 1, the mixed $l_{2,1}$ norm in (11) is equivalent to the l_1 norm in (10), which is the classical basis pursuit based PNLMS algorithm [7]. Meanwhile, if P is chosen as L, the mixed $l_{2,1}$ norm in (13) is the same as the l_2 norm and BS-PNLMS then becomes the traditional NLMS [7]. Therefore, the BS-PNLMS is a generalization of NLMS and PNLMS.

3.3 EXTENSION TO THE BS-IPNLMS ALGORITHM

Meanwhile, in order to further improve the robustness of the proposed BS-PNLMS algorithm to both sparse and dispersive impulse responses, an improved BS-PNLMS (BS-IPNLMS) algorithm is proposed using the similar idea of IPNLMS algorithm

$$\gamma_l = (1-\alpha)\frac{\sum_{i=1}^{N}\left\|\hat{\mathbf{h}}_{[i]}\right\|_2}{N} + (1+\alpha)\left\|\hat{\mathbf{h}}_{[l]}\right\|_2, \tag{20}$$

$$g_l(n) = \frac{\gamma_l(n)}{P\sum_{i=1}^{N}\gamma_i(n)} = \frac{(1-\alpha)}{2L} + \frac{(1+\alpha)\left\|\hat{\mathbf{h}}_{[l]}\right\|_2}{2P\sum_{i=1}^{N}\left\|\hat{\mathbf{h}}_{[i]}\right\|_2}. \tag{21}$$

This section is concluded with a brief discussion about the proposed BS-PNLMS and BS-IPNLMS algorithms. Unlike the PB-IPNLMS, the proposed BS-PNLMS and BS-IPNLMS algorithms only require prior information about the length of the active regions to determine the group size, which are usually known for both the NEC and the satellite link channels, etc., and not their actual locations. The BS-PNLMS could be interpreted as transferring the block-sparse system into a multi-delta system in the coefficient space to fully benefit from PNLMS. However, if the impulse system is dispersive, or the group size is much smaller than the actual block size in the impulse response, the BS-IPNLMS could outperform both the PNLMS and the BS-PNLMS, as well. The details of the proposed BS-PNLMS and BS-IPNLMS algorithms are summarized in Table 1. The superior performance of BS-PNLMS, and BS-IPNLMS over NLMS, PNLMS, and IPNLMS will be demonstrated in the simulations of Section 5.

Table 1. The block-sparse algorithms

Initializations:

$\hat{\mathbf{h}}(n) = \mathbf{0}_{L\times 1}$, $N = L/P$

General Computations:

$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1)$

$\mathbf{G}(n-1) = diag\left[g_1\mathbf{1}_P, g_2\mathbf{1}_P, \cdots g_N\mathbf{1}_P\right]$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mu\mathbf{G}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{G}(n-1)\mathbf{x}(n)+\delta}$$

for $i = 1, 2, \cdots, N$

$$\left\|\hat{\mathbf{h}}_{[i]}\right\|_2 = \sqrt{\sum_{k=1}^{P}\hat{h}^2_{(i-1)P+k}}$$

end for

BS-PNLMS:

for $i = 1, 2, \cdots, N$

$$\gamma_i = \max\left\{\rho\max\left\{q, \left\|\hat{\mathbf{h}}_{[1]}\right\|_2, \cdots, \left\|\hat{\mathbf{h}}_{[N]}\right\|_2\right\}, \left\|\hat{\mathbf{h}}_{[i]}\right\|_2\right\}$$

$$g_i(n) = \frac{\gamma_i}{\frac{1}{N}\sum_{l=1}^{N}\gamma_l}$$

end for

BS-IPNLMS:

for $i = 1, 2, \cdots, N$

$$g_l(n) = \frac{(1-\alpha)}{2L} + \frac{(1+\alpha)\left\|\hat{\mathbf{h}}_{[l]}\right\|_2}{2P\sum_{i=1}^{N}\left\|\hat{\mathbf{h}}_{[i]}\right\|_2}$$

end for

4. COMPUTATIONAL COMPLEXITY

The computational complexity of BS-PNLMS and BS-IPNLMS algorithms is compared with traditional NLMS, PNLMS and IPNLMS algorithms in Table 2 in terms of the total number of additions (A), multiplications (M), divisions (D), comparisons (C), square roots (Sqrt) and memory words (MW), needed per sample. The additional computational complexity for the BS-PNLMS family arises from the computation of the l_2 norm of the block responses using the square root operations. The complexity of the square root can be reduced through the use of a look up table or a Taylor series expansion [22]. Meanwhile, it should be noted that the "comparison operations" and the required memory words for the family of BS-PNLMS are decreased from that of PNLMS. Finally, the computational complexity of the proposed block-sparse family algorithms is also related to the number of groups, N, where $N = L/P$.

Table 2. Computational complexity of the algorithms' coefficient updates – Addition (A), Multiplication (M), Division (D), Comparison (C), Square Root (Sqrt) and Memory Word (MW).

Algorithm	A	M	D	C	Sqrt	MW
NLMS	$2L+3$	$2L+3$	1	0	0	$4L+7$
PNLMS	$4L+2$	$5L+4$	2	$2L$	0	$8L+11$
BS-PNLMS	$4L-1$	$6L+3$	2	$N+1$	N	$5L+3N+11$
IPNLMS	$5L+2$	6L+2	4	$L-1$	0	$8L+11$
BS-IPNLMS	$4L+N-1$	$6L+N+1$	2	0	N	$5L+3N+11$

5. SIMULATION RESULTS

Simulations were conducted to evaluate the performance of the proposed BS-PNLMS and BS-IPNLMS algorithms. The algorithms were tested using zero mean white Gaussian noise (WGN), colored noise and speech signals at sampling rate 8 KHz. The WGN was filtered through a first order system with a pole at 0.8 to generate the colored input signals. An independent WGN was added to the system's background at a signal-to-noise ratio (SNR) of 30dB. The regularization parameter for NLMS was $\delta_{NLMS} = 0.01$, and the regularization parameters for PNLMS, BS-PNLMS, IPNLMS, and BS-IPNLMS were δ_{NLMS}/L according to [19]. The values of N used for both the IPNLMS and the BS-IPNLMS algorithms were 0. For both the PNLMS and the BS-PNLMS algorithms, $\rho = 0.01$, and $q = 0.01$.

The convergence state of adaptive filter was evaluated with the normalized misalignment defined as

$$10\log_{10}(\|\mathbf{h}-\hat{\mathbf{h}}\|_2^2 / \|\mathbf{h}\|_2^2).$$

In all the simulations except for the ones in section 5.3, the length of the unknown system throughout the simulation was $L = 1024$, and the adaptive filter had the same length. A 32 taps impulse response in Figure 1 (b) with a single cluster of nonzero coefficients at [257, 288] was used. In order to compare the tracking ability for different algorithms, an echo path change was incurred at 40000 sample by switching to the two-clusters response located at [257, 272] (16 taps) and [769, 800] (32 taps) as illustrated in in Figure 1 (c). All the algorithms were simulated for five times and averaged in order to evaluate their performance.

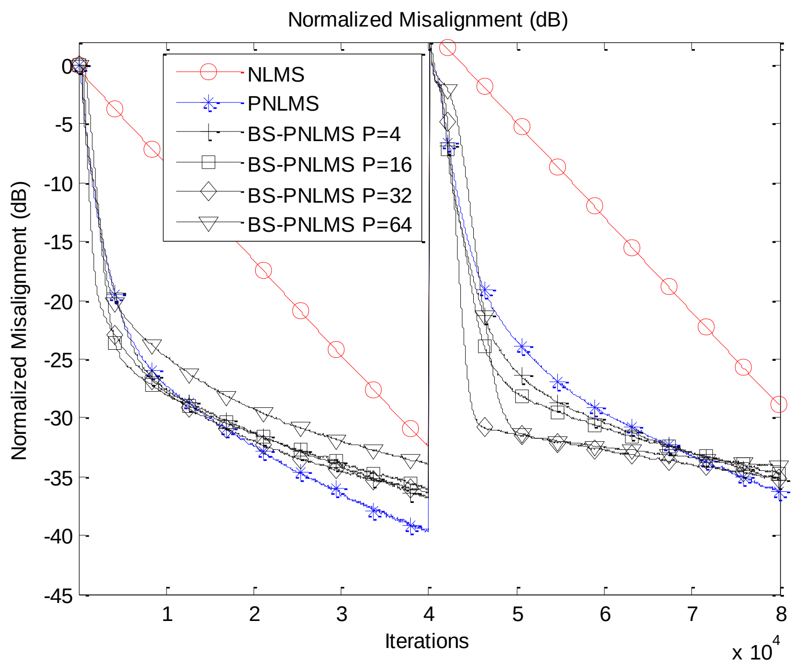
5.1 EFFECT OF *P* ON THE PERFORMANCE OF BS-PNLMS

In order to demonstrate the effect of *P*, the performance of the proposed BS-PNLMS was tested for different group sizes (4, 16, 32, and 64) separately. Meanwhile, the performance of NLMS, which is the same as BS-PNLMS with $P = 1024$, and PNLMS (the same as BS-PNLMS with $P = 1$) algorithms were also included. In the first simulation in Figure 2 (a), the input was WGN, and the step-size μ was set to 0.1. The simulation results for a colored input signal and speech input signal are illustrated in Figure 2 (b) and Figure 2 (c) separately, where the step-sizes were $\mu = 0.2$ for both the colored input and the speech input. Meanwhile, the remaining parameters for the three simulations were the same.

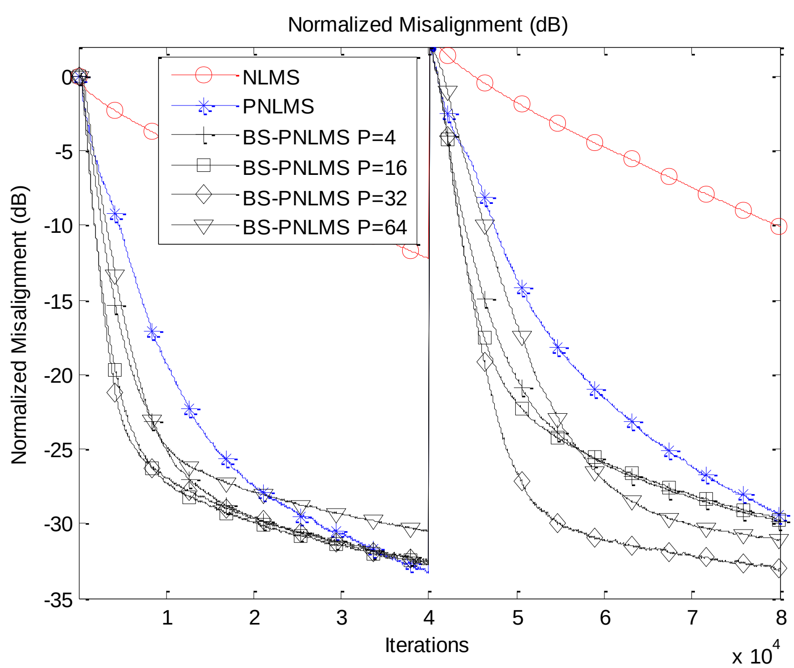
Simulation results in Figure 2 indicate that the group size P should be chosen properly in order to gain better performance than either the NLMS or the PNLMS. Due to the fact that there are a total 32 taps in the single-cluster impulse response, it is reasonable that the group size larger than 32 will likely degrade the performance before the echo path change. Meanwhile, there are two clusters with length 16 taps separately in the two-cluster impulse response, and the group size should be smaller than 16. Because the groups are evenly spaced, the actual block could have been split into multiple groups too. Therefore, the group size should be smaller than the length of cluster's actual minimum size in the impulse response. The cluster's size is typically known in real-world applications. For example, the NEC's "active" region is in the range of 8-12 *ms* duration [1]. If the group size is significantly larger than the cluster size of block-sparse system, the convergence speed will become worse than the traditional PNLMS. This fact is intuitive, considering that NLMS, which uses $P = 1024$, converges slower than PNLMS with $P = 1$ for a block-sparse system. Thus, both NLMS and PNLMS represent extreme cases. The NLMS algorithm should be chosen when the unknown system is dispersive, i.e. the cluster size is the length of the full filter, and when the unknown system is generally sparse as illustrated in Figure 1(a), PNLMS should be used because the cluster size is 1.

5.2 CONVERGENCE PERFORMANCE OF BS-PNLMS AND BS-IPNLMS FOR BLOCK-SPARSE SYSTEMS

The performances of NLMS, PNLMS, IPNLMS, proposed BS-PNLMS with $P = 16$ and the proposed BS-IPNLMS with $P = 4$ were compared for the two block-sparse systems in Figure 3.

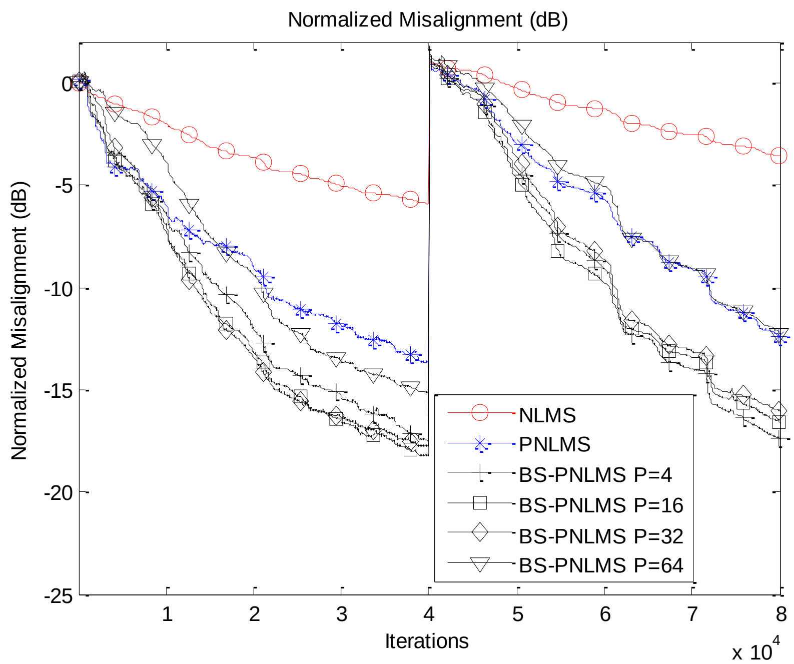


(a) WGN input with $\mu = 0.1$



(b) Colored noise input with $\mu = 0.2$

Figure 2. Comparison of the BS-PNLMS algorithms with different group sizes for block-sparse systems in Figure 1 (b) and Figure 1 (c) at SNR=30dB: (a) white, (b) colored noise and (c) speech input signals.



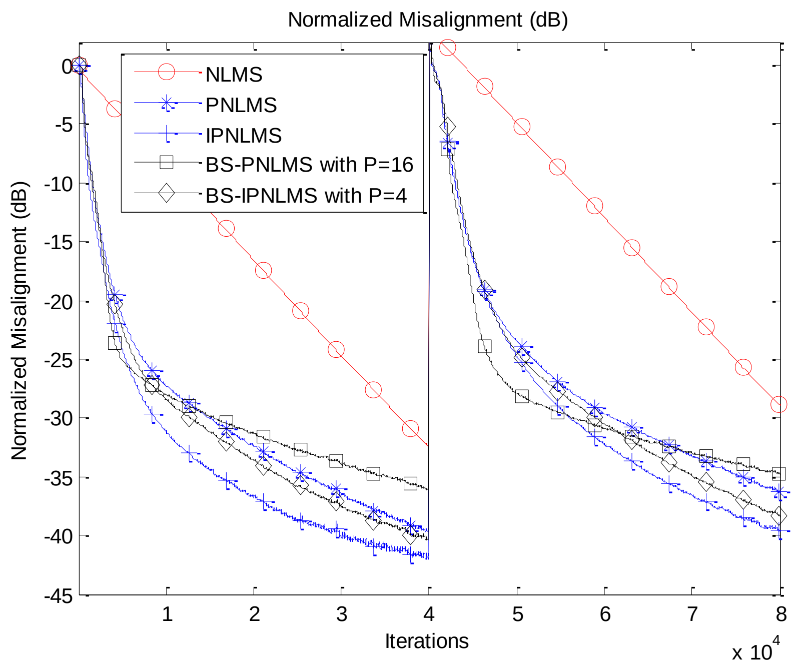
(c) Speech input with $\mu = 0.2$

Figure 2. Comparison of the BS-PNLMS algorithms with different group sizes for block-sparse systems in Figure 1 (b) and Figure 1 (c) at SNR=30dB: (a) white, (b) colored noise and (c) speech input signals (cont.).

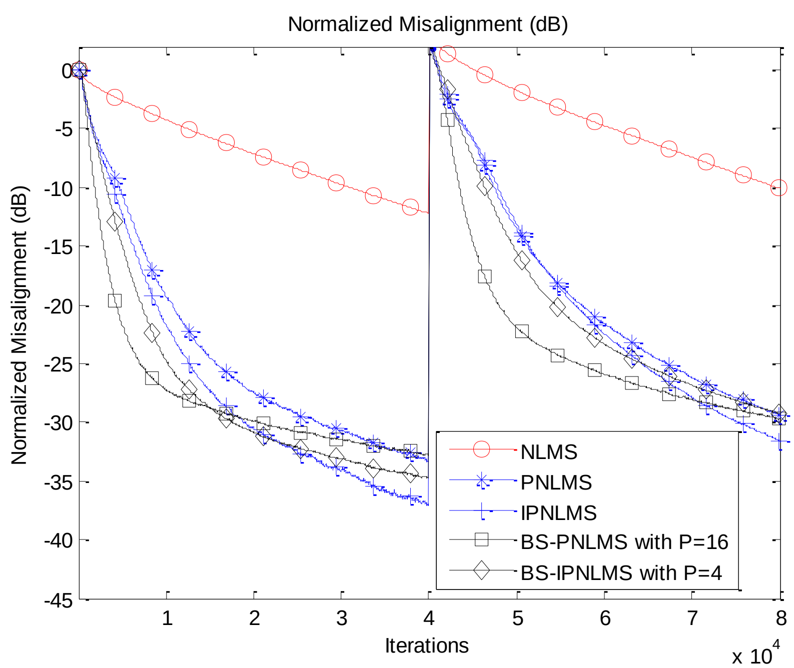
The WGN was used as the input signal in Figure 3 (a) with the step-sizes as $\mu_{NLMS} = \mu_{PNLMS} = 0.1$, and $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.1$. The simulation results for the colored and speech input are illustrated in Figure 3 (b) and Figure 3 (c), where $\mu_{NLMS} = \mu_{PNLMS} = 0.2$, and $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.2$.

The proposed BS-PNLMS algorithm provides faster convergence rate and tracking ability than either the NLMS or the traditional PNLMS algorithms for the block-sparse impulse responses. Meanwhile, the convergence rate of BS-IPNLMS outperformed both the NLMS and the IPNLMS algorithms.

It is interesting to observe that the BS-PNLMS algorithm outperformed the BS-IPNLMS algorithm. This is due to fact that the two block-sparse systems in Figure 1 (b) and Figure 1 (c) are very sparse. Meanwhile, the BS-PNLMS transformed them into highly sparse systems with only 2 or 3 non-zero elements which fully benefits from PNLMS.

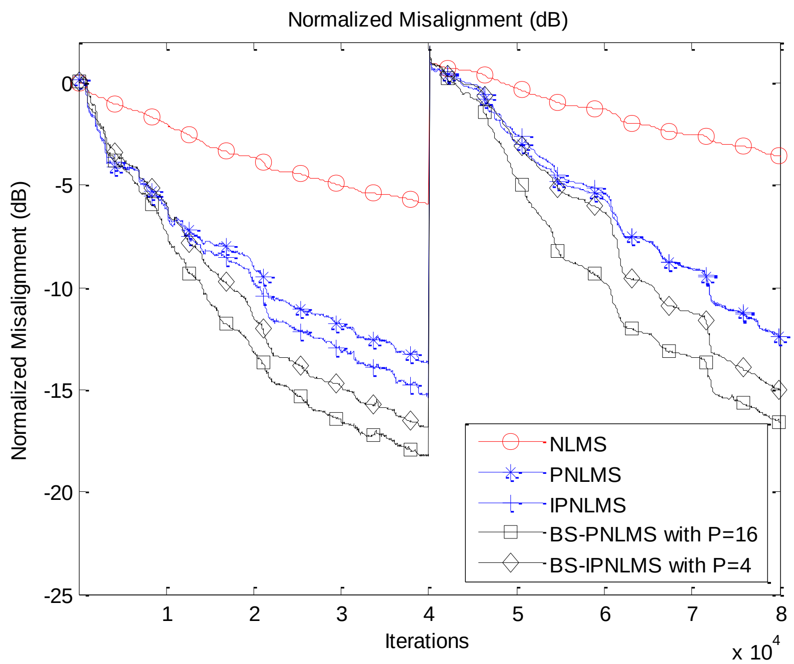


(a) $\mu_{NLMS} = \mu_{PNLMS} = 0.1$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.1$



(b) $\mu_{NLMS} = \mu_{PNLMS} = 0.2$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.2$

Figure 3. Comparison of NLMS, PNLMS, IPNLMS, BS-PNLMS and BS-IPNLMS algorithms for block-sparse systems in Figure 1 (b) and Figure 1 (c) at SNR=30dB: (a) WGN input, (b) colored noise and (c) speech input signals.



(c) $\mu_{NLMS} = \mu_{PNLMS} = 0.2$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.2$

Figure 3. Comparison of NLMS, PNLMS, IPNLMS, BS-PNLMS and BS-IPNLMS algorithms for block-sparse systems in Figure 1 (b) and Figure 1 (c) at SNR=30dB: (a) WGN input, (b) colored noise and (c) speech input signals (cont.).

Meanwhile, the benefits of BS-IPNLMS for the dispersive impulse responses will be demonstrated in the next subsection.

5.3 CONVERGENCE PERFORMANCE OF BS-PNLMS AND BS-IPNLMS FOR THE ACOUSTIC ECHO PATH AND A RANDOM DISPERSIVE SYSTEM

In order to verify the performance of the proposed BS-IPNLMS algorithm for dispersive impulse response, simulations were conducted to compare the performances of NLMS, PNLMS, IPNLMS, the proposed BS-PNLMS with $P = 16$, and the proposed BS-IPNLMS with $P = 16$. An echo path change was incurred at 40000 samples by switching from a 512 taps measured acoustic echo path in Figure 4 (a) to a random impulse response in Figure 4 (b). The simulation results for WGN, colored noise and speech input signals are illustrated in Figure 5.

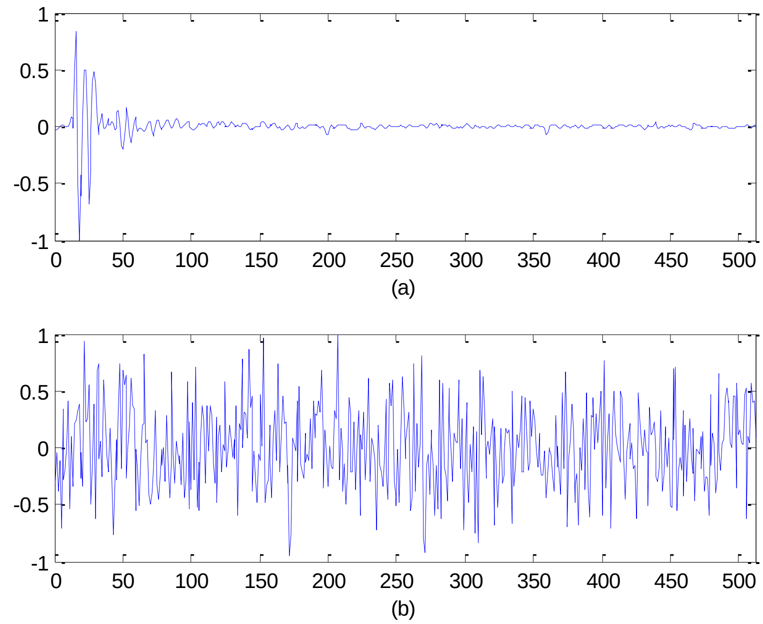
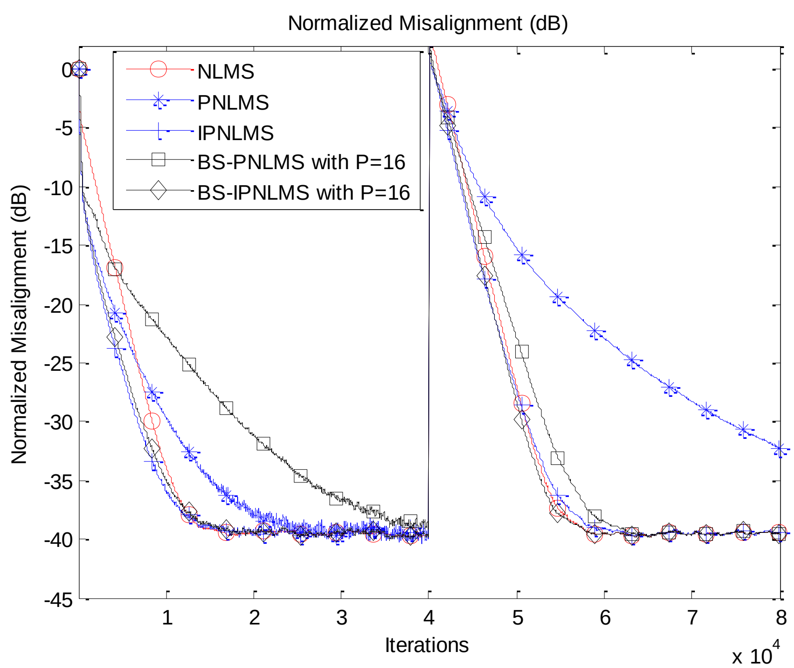
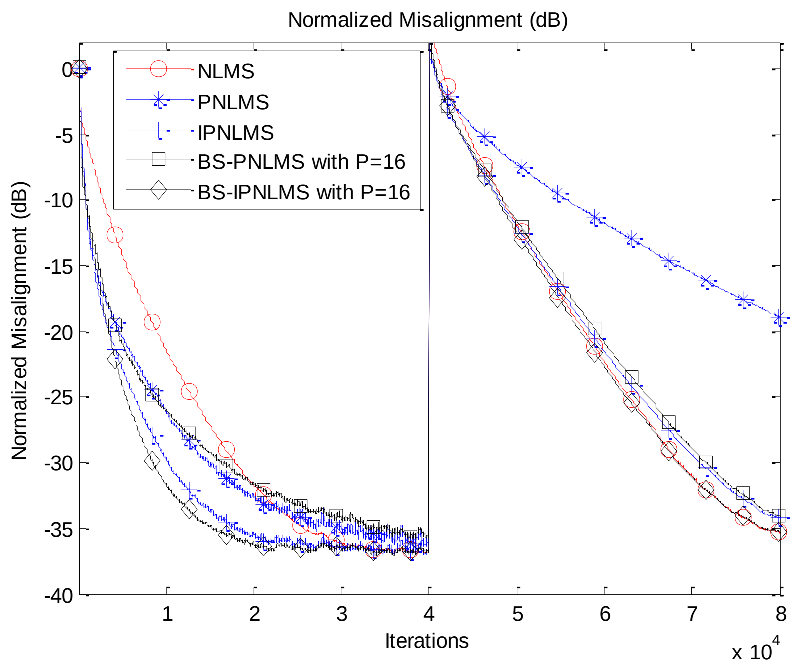


Figure 4. Two impulse responses (a) a measured quasi-sparse acoustic echo path, (b) a random dispersive impulse response.

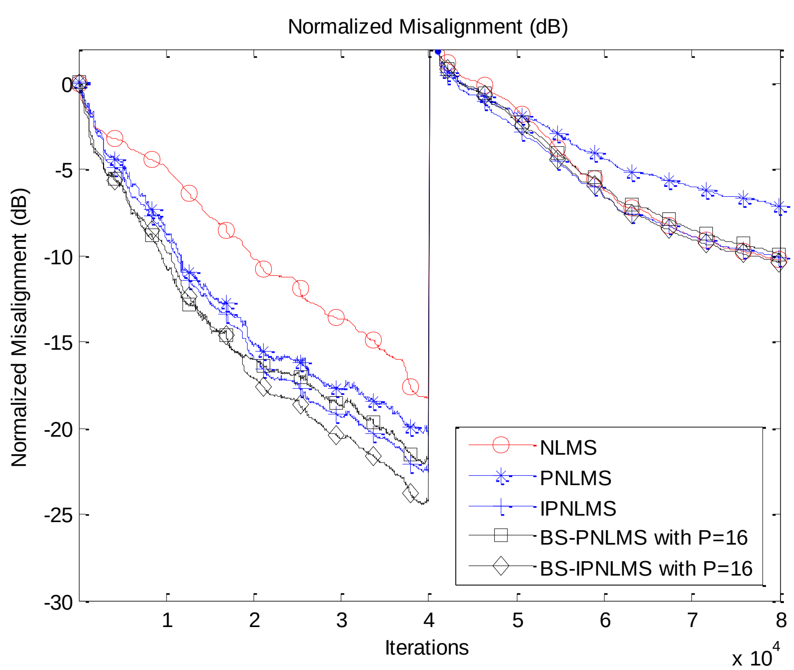


(a) $\mu_{NLMS} = \mu_{PNLMS} = 0.2$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.2$

Figure 5. Comparison of NLMS, PNLMS, IPNLMS, BS-PNLMS and BS-IPNLMS algorithms for acoustic echo path and dispersive system in Figure 4 and SNR=30dB: (a) WGN input (b) colored noise and (c) speech input.



(b) $\mu_{NLMS} = \mu_{PNLMS} = 0.4$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.4$



(c) $\mu_{NLMS} = \mu_{PNLMS} = 0.4$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.4$

Figure 5. Comparison of NLMS, PNLMS, IPNLMS, BS-PNLMS and BS-IPNLMS algorithms for acoustic echo path and dispersive system in Figure 4 and SNR=30dB: (a) WGN input (b) colored noise and (c) speech input (cont.).

The step-size parameters were $\mu_{NLMS} = \mu_{PNLMS} = 0.2$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.2$ for the WGN input, and $\mu_{NLMS} = \mu_{PNLMS} = 0.4$, $\mu_{BS-PNLMS} = \mu_{BS-IPNLMS} = 0.4$ for both the colored noise and the speech input signals.

It can be observed that the BS-IPNLMS algorithm outperformed the BS-PNLMS algorithm for both the acoustic echo path and the random dispersive impulse response. Meanwhile, both BS-PNLMS and BS-IPNLMS work better than the traditional PNLMS algorithm for the random dispersive impulse responses.

It should be noted that, neither the acoustic echo path nor the random dispersive impulse response are typical block-sparse impulse systems, therefore, the family of BS-IPNLMS should be used to obtain better performance instead of the BS-PNLMS algorithms.

6. CONCLUSION AND FUTURE WORK

A new family of proportionate algorithms for block-sparse system identification (known as BS-PNLMS and BS-IPNLMS) were proposed. These algorithms were based on the optimization of a mixed $l_{2,1}$ norm of the adaptive filter's coefficients. The computational complexities of the proposed algorithms were presented. Simulation results demonstrated that, the new BS-PNLMS algorithm outperforms the NLMS, PNLMS and IPNLMS algorithms for the block-sparse system, and the new BS-IPNLMS algorithm is more preferred for the dispersive system.

This block-sparse proportionate idea proposed in this paper could be further extended to many other proportionate algorithms, including proportionate affine projection algorithm (PAPA) [23], proportionate affine projection sign algorithm (PAPSA) [24], and their corresponding low complexity implementations [25]-[26] etc. The proof of convergence for the proposed BS-PNLMS and BS-IPNLMS algorithms can also be part of the future work. Finally, it will be interesting to explore the variable and non-uniform group split to further improve the performance of the BS-PNLMS and the BS-IPNLMS algorithms.

7. ACKNOWLEDGEMENT

The authors would like to thank the Associate Editor and the reviewers for the valuable comments and suggestions.

8. REFERENCES

[1] K. Wagner and M. Doroslovački, *Proportionate-type Normalized Least Mean Square Algorithms*: John Wiley & Sons, 2013.

[2] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *Speech and Audio Processing, IEEE Transactions on,* vol. 8, no. 5, pp. 508-518, 2000.

[3] H. Deng and M. Doroslovački, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *Signal Processing Letters, IEEE,* vol. 12, no. 3, pp. 181-184, 2005.

[4] H. Deng and M. Doroslovački, "Proportionate adaptive algorithms for network echo cancellation," *Signal Processing, IEEE Transactions on,* vol. 54, no. 5, pp. 1794-1803, 2006.

[5] K. Wagner, M. Doroslovački, and H. Deng, "Convergence of proportionate-type LMS adaptive filters and choice of gain matrix," *40th Asilomar Conference on Signals, Systems and Computers*, 2006 (ACSSC '06), Pacific Grove, CA, pp. 243–247, November 2006.

[6] C. Paleologu, J. Benesty and S. Ciochină, "An improved proportionate NLMS algorithm based on l_0 norm," *IEEE ICASSP*, pp. 309-312, 2010.

[7] J. Benesty, C. Paleologu, and S. Ciochină, "Proportionate adaptive filters from a basis pursuit perspective," *Signal Processing Letters, IEEE,* vol. 17, no. 12, pp. 985-988, 2010.

[8] J. Liu and S. L. Grant, "A generalized proportionate adaptive algorithm based on convex optimization," in *Proc. Signals and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*, pp. 748-752, 2014.

[9] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," *Conference Record of the 32nd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, vol. 1, pp. 394–398, November 1998.

[10] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," *IEEE ICASSP*, pp. 1881-1884, 2002.

[11] J. Cui, P. Naylor, and D. Brown, “An improved IPNLMS algorithm for echo cancellation in packet-switched networks,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004 (ICASSP ’04), vol. 4, Montreal, Quebec, Canada, pp. 141–144, May 2004.

[12] P. Loganathan, E. A. P. Habets and P. A. Naylor, “A partitioned block proportionate adaptive algorithm for acoustic echo cancellation,” *Proc. of the APSIPA Annual Summit and Conference* 2010, Biopolis, Singapore, Dec 2010.

[13] F. D. C. De Souza, O. J. Tobias, R. Seara *et al.*, “A PNLMS algorithm with individual activation factors,” *Signal Processing, IEEE Transactions on,* vol. 58, no. 4, pp. 2036-2047, 2010.

[14] Y. Chen, Y. Gu and A. O. Hero III, “Sparse LMS for system identification,” *IEEE ICASSP*, pp. 3125-3128, 2009.

[15] Y. Gu, J. Jin, and S. Mei, “l0 norm constraint LMS algorithm for sparse system identification,” *Signal Processing Letters, IEEE,* vol. 16, no. 9, pp. 774-777, 2009.

[16] J. Liu and S. L. Grant, “A new variable step-size zero-point attracting projection algorithm,” in *Proc. Signals, Systems and Computers, 2013 Asilomar Conference*, pp. 1524-1528, 2013.

[17] J. Liu and S. L. Grant, “An improved variable step-size zero-point attracting projection algorithm,” *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on* , pp.604-608, 19-24 April 2015

[18] S. Jiang and Y. Gu, “Block-Sparsity-Induced Adaptive Filter for Multi-Clustering System Identification,” in *Signal Processing, IEEE Transactions on* , vol.63, no.20, pp.5318-5330, Oct.15, 2015

[19] J. Benesty, C. Paleologu, and S. Ciochină, “On regularization in adaptive filtering,” *Audio, Speech, and Language Processing, IEEE Transactions on,* vol. 19, no. 6, pp. 1734-1742, 2011.

[20] P. A. Marques, F. M. Sousa, and J. Leitao, “A DSP based long distance echo canceller using short length centered adaptive filters,” In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, vol. 3, pp. 1885-1888. IEEE, 1997.

[21] P. Loganathan, A. W. Khong, and P. Naylor, “A class of sparseness-controlled algorithms for echo cancellation,” *Audio, Speech, and Language Processing, IEEE Transactions on,* vol. 17, no. 8, pp. 1591-1601, 2009.

[22] D. Liu, *Embedded DSP processor design: application specific instruction set processors*: Morgan Kaufmann, 2008.

[23] T. Gansler, J. Benesty, S. L. Gay, and M. Sondhi. “A robust proportionate affine projection algorithm for network echo cancellation,” In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, vol. 2, pp. II793-II796. IEEE, 2000.

[24] Z. Yang, Y. R. Zheng, and S. L. Grant, “Proportionate affine projection sign algorithms for network echo cancellation,” *Audio, Speech, and Language Processing, IEEE Transactions on,* vol. 19, no. 8, pp. 2273-2284, 2011.

[25] C. Paleologu, S. Ciochină, and J. Benesty, “An efficient proportionate affine projection algorithm for echo cancellation,” *Signal Processing Letters, IEEE* 17, no. 2 (2010): 165-168.

[26] F. Albu and H. K. Kwan, “Memory improved proportionate affine projection sign algorithm,” *Electronics letters* 48, no. 20 (2012): 1279-1281.

II. PROPORTIONATE AFFINE PROJECTION ALGORITHMS FOR BLOCK-SPARSE SYSTEM IDENTIFICATION

Jianming Liu and Steven L. Grant

Abstract

A new family of block-sparse proportionate affine projection algorithms (BS-PAPA) is proposed to improve the performance for block-sparse systems. This is motivated by the recent block-sparse proportionate normalized least mean square (BS-PNLMS) algorithm. It is demonstrated that the affine projection algorithm (APA), proportionate APA (PAPA), BS-PNLMS and PNLMS are all special cases of the proposed BS-PAPA algorithm. Meanwhile, an efficient implementation of the proposed BS-PAPA and block-sparse memory PAPA (BS-MPAPA) are also presented to reduce computational complexity. Simulation results demonstrate that the proposed BS-PAPA and BS-MPAPA algorithms outperform the APA, PAPA and MPAPA algorithms for block-sparse system identification in terms of both faster convergence speed and better tracking ability.

1. INTRODUCTION

The impulse responses of many applications, such as network echo cancellation (NEC), are sparse, which means a small percentage of the impulse response components have a significant magnitude while the rest are zero or small. Therefore, instead of the normalized least mean square (NLMS) [1] and the affine projection algorithm (APA) [2], the family of proportionate algorithms exploits this sparseness to improve their performance, including proportionate NLMS (PNLMS) [3], and proportionate APA (PAPA) [4]. The memory improved PAPA (MIPAPA) algorithm was proposed to not only speed up the convergence rate but also reduce the computational complexity by taking into account the memory of the proportionate coefficients [5].

It has been shown that the PNLMS algorithm and PAPA can both be deduced from a basis pursuit perspective [6]-[7]. A more general framework was further proposed to derive the PNLMS adaptive algorithms for sparse system identification, which employed convex optimization [8]. Recently, the block-sparse PNLMS (BS-PNLMS) algorithm was proposed to improve the performance of PNLMS for identifying block-sparse systems [9]. Motivated by BS-PNLMS, we propose a family of block-sparse PAPA algorithms for block-sparse system identification in this paper. The PNLMS, BS-PNLMS, APA and PAPA algorithms are all special cases of this proposed BS-PAPA algorithm. Meanwhile, in order to reduce the computational complexity, taking advantage of the block-sparse property in the proposed BS-PAPA algorithm, an efficient implementation of BS-PAPA is studied, and the block-sparse memory PAPA (BS-MPAPA) algorithm is also introduced.

2. REVIEW OF PAPA

In the typical echo cancellation problem, the input signal $\mathbf{x}(n)$ is filtered through the unknown coefficients $\mathbf{h}(n)$ to get the observed output signal $d(n)$.

$$d(n)=\mathbf{x}^T(n)\mathbf{h}(n)+v(n), \tag{1}$$

where

$$\mathbf{x}(n)=[x(n),x(n-1),\cdots,x(n-L+1)]^T,$$

$v(n)$ is the measurement noise, and L is the length of the impulse response. We define the estimated error as

$$e(n)=d(n)-\mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \tag{2}$$

where $\hat{\mathbf{h}}(n)$ is the adaptive filter's coefficients. Grouping the M most recent input vectors together gives the input signal matrix:

$$\mathbf{X}(n)=[\mathbf{x}(n),\mathbf{x}(n-1),\cdots,\mathbf{x}(n-M+1)].$$

Therefore, the estimated error vector is

$$\mathbf{e}(n)=\mathbf{d}(n)-\mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \tag{3}$$

in which

$$\mathbf{d}(n)=[d(n),d(n-1),\cdots,d(n-M+1)],$$

$$\mathbf{e}(n)=[e(n),e(n-1),\cdots,e(n-M+1)],$$

and M is the projection order. The PAPA algorithm updates the filter coefficients as follows [4]:

$$\begin{aligned}\hat{\mathbf{h}}(n)&=\hat{\mathbf{h}}(n-1)+\\&\quad\mu\mathbf{G}(n-1)\mathbf{X}(n)\left(\mathbf{X}(n)\mathbf{G}(n-1)\mathbf{X}(n)+\delta\mathbf{I}_M\right)^{-1}\mathbf{e}(n),\end{aligned} \tag{4}$$

in which μ is the step-size, δ is the regularization, $\mathbf{I}_M$ is $M\times M$ identity matrix and

$$\mathbf{G}(n-1) = diag\left[g_1(n-1), g_2(n-1), \cdots, g_L(n-1)\right], \tag{5}$$

$$g_l(n-1) = \frac{\gamma_l(n-1)}{\frac{1}{L}\sum_{i=1}^{L}\gamma_i(n-1)}, \tag{6}$$

$$\gamma_l = \max\left\{\rho \max\left\{q, \left|\hat{h}_1\right|, \cdots, \left|\hat{h}_L\right|\right\}, \left|\hat{h}_l\right|\right\}, \tag{7}$$

q prevents the filter coefficients $\hat{h}_l(n-1)$ from stalling when $\hat{\mathbf{h}}(0) = \mathbf{0}_{L\times 1}$ at initialization, and ρ prevents the coefficients from stalling when they are much smaller than the largest coefficient.

In many applications, including network echo cancellation (NEC) and satellite-linked communication echo cancellation, the impulse response is block sparse, that is, it consists of several dispersive active regions. However, PAPA does not take into account the block-sparse characteristic, and motivated by the block-sparse PNLMS (BS-PNLMS) algorithm [9], we propose a family of new block-sparse PAPA algorithms to further improve their performance for identifying the block-sparse impulse system in next section.

3. PROPOSED BS-PAPA

The block-sparse scheme for PAPA will be firstly derived based on the optimization of $l_{2,1}$ norm, then in order to reduce the computational complexity, an efficient implementation of the proposed BS-PAPA is presented by taking advantage of the block structure. Finally, block-sparse memory PAPA (BS-MPAPA) is also proposed by considering the memory of the coefficients to further reduce computational complexity.

3.1 THE PROPOSED BS-PAPA

The proportionate APA algorithm can be deduced from a *basis pursuit* perspective as below [7]

$$\begin{array}{ll} \boldsymbol{min} & \left\|\tilde{\mathbf{h}}(n)\right\|_1 \\ \text{subject to} & \mathbf{d}(n)=\mathbf{X}^T(n)\tilde{\mathbf{h}}(n), \end{array} \tag{8}$$

where $\tilde{\mathbf{h}}(n)$ is the correction component defined as [6]-[7]

$$\tilde{\mathbf{h}}(n)=\mathbf{G}(n-1)\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n)\right]^{-1}\mathbf{d}(n). \tag{9}$$

Motivated by BS-PNLMS, the proposed block-sparse scheme for PAPA is derived by replacing the l_1 norm optimization target in the basis pursuit perspective with the following $l_{2,1}$ norm defined as

$$\left\|\tilde{\mathbf{h}}\right\|_{2,1}=\left\|\begin{bmatrix}\left\|\tilde{\mathbf{h}}_{[1]}\right\|_2\\ \left\|\tilde{\mathbf{h}}_{[2]}\right\|_2\\ \vdots\\ \left\|\tilde{\mathbf{h}}_{[N]}\right\|_2\end{bmatrix}\right\|_1=\sum_{i=1}^{N}\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2, \tag{10}$$

where $\left\|\tilde{\mathbf{h}}_{[i]}\right\|_2=\sqrt{\tilde{\mathbf{h}}_{[i]}^T\tilde{\mathbf{h}}_{[i]}}$, $\tilde{\mathbf{h}}_{[i]}=[\tilde{h}_{(i-1)P+1},\tilde{h}_{(i-1)P+2},\cdots,\tilde{h}_{iP}]^T$, P is a predefined group partition size parameter and $N=L/P$ is the number of groups. Therefore,

$$\begin{aligned} &\boldsymbol{min} && \left\|\tilde{\mathbf{h}}(n)\right\|_{2,1} \\ &\text{subject to} && \mathbf{d}(n) = \mathbf{X}^T(n)\tilde{\mathbf{h}}(n). \end{aligned} \tag{11}$$

Similarly, the proposed BS-PAPA could be derived using the method of Lagrange multipliers, see [6]-[7] for more details. The update equation for the proposed BS-PAPA is then,

$$\begin{aligned} &\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) \\ &+\mu\mathbf{G}(n-1)\mathbf{X}(n)\left(\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n)+\delta\mathbf{I}_M\right)^{-1}\mathbf{e}(n), \end{aligned} \tag{12}$$

and

$$\begin{aligned} &\mathbf{G}(n-1) = \\ &\quad diag\left[\left\|\hat{\mathbf{h}}_{[1]}\right\|_2\mathbf{1}_P, \left\|\hat{\mathbf{h}}_{[2]}\right\|_2\mathbf{1}_P, \cdots, \left\|\hat{\mathbf{h}}_{[N]}\right\|_2\mathbf{1}_P\right], \end{aligned} \tag{13}$$

in which $\mathbf{1}_P$ is a P-length row vector of all ones. Equation (12) is the same as traditional PAPA in (4), except for the block-sparse definition of $\mathbf{G}(n-1)$ in (13). Similar to (5)-(7) in PAPA to prevent the stalling issues, the proposed BS-PAPA replaces (5)-(7) with

$$\begin{aligned} &\mathbf{G}(n-1) = \\ &diag\left[g_1(n-1)\mathbf{1}_P, g_2(n-1)\mathbf{1}_P, \cdots, g_N(n-1)\mathbf{1}_P\right], \end{aligned} \tag{14}$$

$$g_i(n-1) = \frac{\gamma_i}{\frac{1}{N}\sum_{l=1}^{N}\gamma_l}, \tag{15}$$

$$\gamma_i = \max\left\{\rho\max\left\{q, \left\|\hat{\mathbf{h}}_{[1]}\right\|_2, \cdots, \left\|\hat{\mathbf{h}}_{[N]}\right\|_2\right\}, \left\|\hat{\mathbf{h}}_{[i]}\right\|_2\right\}. \tag{16}$$

It should be noted that the proposed BS-PAPA includes PNLMS, BS-PNLMS, APA and PAPA. The BS-PNLMS algorithm is a special case of BS-PAPA with projection order $M=1$. In the case of P is equal to 1, the BS-PAPA algorithm degenerates to PAPA. Meanwhile, when P is chosen as L, the proposed BS-PAPA turns into APA.

3.2 EFFICIENT IMPLEMENTATION OF PROPOSED BS-PAPA

By taking advantage of the new block-sparse characteristic in the proposed BS-PAPA algorithm, we can reduce the computational complexity of the proposed BS-PAPA, especially for higher projection order. Equation (12) can be rewritten as

$$\mathbf{P}(n)=\mathbf{G}(n-1)\mathbf{X}(n), \tag{17}$$

$$\begin{aligned}\hat{\mathbf{h}}(n)&=\hat{\mathbf{h}}(n-1)\\&\quad+\mu\mathbf{P}(n)\left(\mathbf{X}^T(n)\mathbf{P}(n)+\delta\mathbf{I}_M\right)^{-1}\mathbf{e}(n).\end{aligned} \tag{18}$$

Considering the blocks of $\mathbf{G}(n-1)$ in (14), (17) can be rewritten as (19) below,

$$\mathbf{P}(n)=\begin{bmatrix} g_1(n-1)\mathbf{x}_P(n), & g_1(n-1)\mathbf{x}_P(n-1), & \cdots, & g_1(n-1)\mathbf{x}_P(n-M+1) \\ g_2(n-1)\mathbf{x}_P(n-P), & g_2(n-1)\mathbf{x}_P(n-P-1), & \cdots, & g_2(n-1)\mathbf{x}_P(n-M+1-P) \\ \vdots & \vdots & \ddots & \vdots \\ g_N(n-1)\mathbf{x}_P(n-(N-1)P), & g_N(n-1)\mathbf{x}_P(n-(N-1)P-1), & \cdots, & g_N(n-1)\mathbf{x}_P(n-M+1-(N-1)P) \end{bmatrix}, \tag{19}$$

where

$$\mathbf{x}_P(n)=[x(n)\,x(n-1)\cdots x(n-P+1)]^T. \tag{20}$$

The direct implementation of (17) will need ML multiplications, which is the case of classical PAPA. However, considering the block-sparse characteristic in (14), the computational complexity of (19) can be further reduced. The *ith* submatrix of $\mathbf{P}(n)$ is defined as $\mathbf{P}_i(n)$ in (21).

$$\mathbf{P}_i(n)=\left[g_i(n-1)\mathbf{x}_P(n-(i-1)P), g_i(n-1)\mathbf{x}_P(n-(i-1)P-1),\cdots, g_i(n-1)\mathbf{x}_P(n-M+1-(i-1)P)\right]. \tag{21}$$

Considering the shift property of $\mathbf{x}_P(n)$ in (20), we only need to calculate the vector $\mathbf{p}_i(n)$ in (22)

$$\mathbf{p}_i(n)=\left[g_i(n-1)x(n-(i-1)P), g_i(n-1)x(n-(i-1)P-1),\cdots, g_i(n-1)x(n-M+2-iP)\right]^T, \tag{22}$$

which requires $P+M-1$ multiplications then use a sliding window to construct $\mathbf{P}_i(n)$. Therefore, the number of multiplications of (19) in the proposed BS-PAPA will become $(P+M-1)N$.

It should be noted that, the proposed efficient implementation will not damage the performance of the BS-PAPA algorithm. Meanwhile, the advantage of proposed efficient implementation becomes more apparent when the projection order and block size increase.

3.3 MEMORY BS-PAPA

In order to further reduce the computational complexity of (19), we could consider the memory of proportionate coefficients as in [5], and approximate the matrix $\mathbf{P}(n)$ by $\mathbf{P'}(n)$ in (23)

$$\mathbf{P'}(n)=\begin{bmatrix} g_1(n-1)\mathbf{x}_P(n), & g_1(n-2)\mathbf{x}_P(n-1), & \cdots, g_1(n-M)\mathbf{x}_P(n-M+1) \\ g_2(n-1)\mathbf{x}_P(n-P), & g_2(n-2)\mathbf{x}_P(n-P-1), & \cdots, g_2(n-M)\mathbf{x}_P(n-M+1-P) \\ \vdots & \vdots & \ddots \quad \vdots \\ g_N(n-1)\mathbf{x}_P(n-(N-1)P), & g_N(n-2)\mathbf{x}_P(n-(N-1)P-1), & \cdots, g_N(n-M)\mathbf{x}_P(n-M+1-(N-1)P) \end{bmatrix}. \quad (23)$$

Due to time-shift property of (23), it could be implemented as

$$\mathbf{P'}(n)=\left[\mathbf{g}(n-1)\odot\mathbf{x}(n),\mathbf{P'}_{-1}(n-1)\right], \quad (24)$$

where the operation $\odot$ denotes the Hadamard product and the matrix $\mathbf{P'}_{-1}(n-1)$ contains the first $M-1$ columns of $\mathbf{P'}(n-1)$. The calculation of $\mathbf{P'}(n)$ only needs L multiplications, and the proposed BS-MPAPA updates the coefficients as below:

$$\begin{aligned}\hat{\mathbf{h}}(n)&=\hat{\mathbf{h}}(n-1)\\ &+\mu\mathbf{P'}(n)\left(\mathbf{X}^T(n)\mathbf{P'}(n)+\delta\mathbf{I}_M\right)^{-1}\mathbf{e}(n).\end{aligned} \quad (25)$$

It should be noted that the efficient implementation proposed in Section III.B could not be applied to the memory BS-PAPA, however, the computational complexity of memory BS-PAPA will be lower than BS-PAPA due to the time-shift property when considering the memory.

4. SIMULATION RESULTS

The performance of the proposed BS-PAPA and BS-MPAPA are evaluated via simulations. Throughout our simulation, the length of the unknown system is $L = 1024$, and the adaptive filter is the same length. Two block-sparse impulse systems in Figure 1 are used: the first impulse response in Figure 1(a) is with a single cluster of nonzero coefficients at [257, 288], which has 32 taps; the two clusters in the second impulse response in Figure 1(b) locate at [257, 288] (32 taps) and [769, 800] (32 taps) separately. In order to compare the tracking ability for different algorithms, an echo path change was incurred at 30000-sample by switching from the first impulse response in Figure 1(a) to the second impulse response in Figure 1(b).

The algorithms were tested using colored noise which was generated by filtering white Gaussian noise (WGN) through a first order system with a pole at 0.8. Independent WGN is added to the system background with a signal-to-noise ratio, SNR = 30dB. The projection order was $M = 8$, and the step-sizes were $\mu = 0.01$. The regularization parameters δ were set to 0.01, and we used $\rho = 0.01$, and $q = 0.01$. The convergence state of adaptive filter is evaluated with the normalized misalignment which is defined as $10\log_{10}(\|\mathbf{h} - \hat{\mathbf{h}}\|_2^2 / \|\mathbf{h}\|_2^2)$.

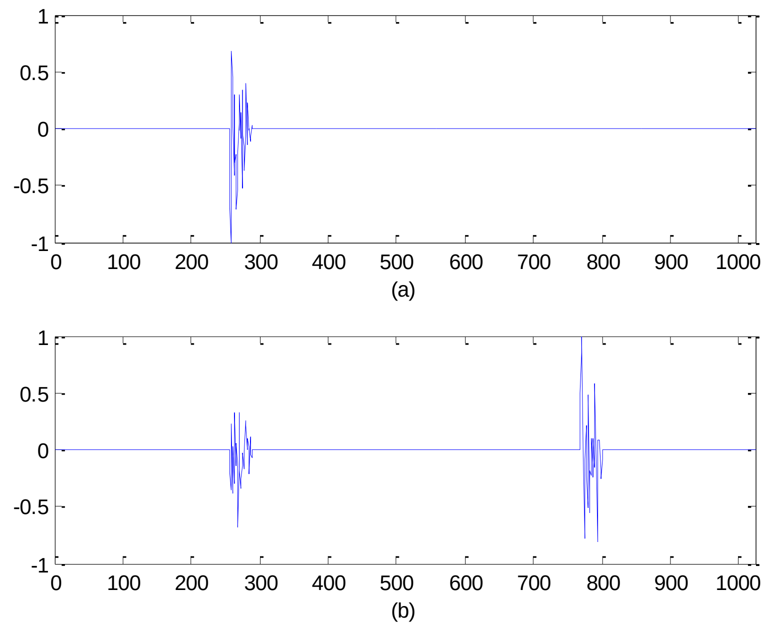


Figure 1. Block-sparse impulse systems (a) one-cluster block-sparse system, (b) two-cluster block-sparse system.

The performance of the proposed BS-PAPA was tested for different group sizes chosen as 1 (i.e. PAPA), 4, 16, 32, 64, 1024 (i.e. APA) separately in Figure 2. The impact of different group sizes on BS-MPAPA is similar. As discussed in BS-PNLMS [9], the group size should be chosen properly (around 32 here) in order to fully take advantage of the block-sparse characteristic.

In the second simulation, we compare the performance of BS-PAPA and BS-MPAPA algorithms together with APA, PAPA and MPAPA. For both the BS-PAPA and BS-MPAPA algorithms, the group size was $P = 32$. The convergence curves for colored input are shown in Figure 3. As can be seen, both proposed BS-PAPA and BS-MPAPA outperform PAPA and MPAPA in terms of convergence speed and tracking ability. Meanwhile, BS-MPAPA will be more favorable considering its lower computation complexity.

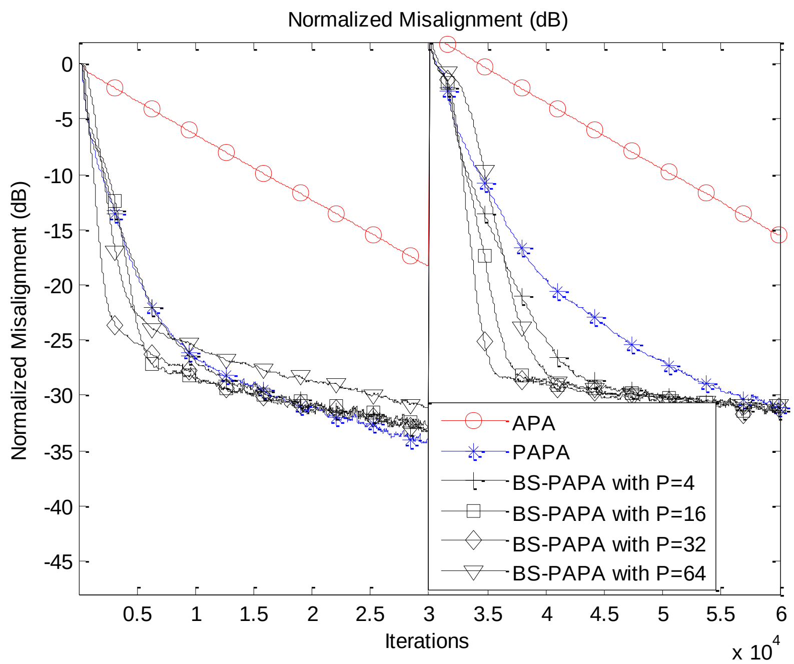


Figure 2. Comparison of BS-PAPA with different group sizes for colored input with *SNR*=30dB.

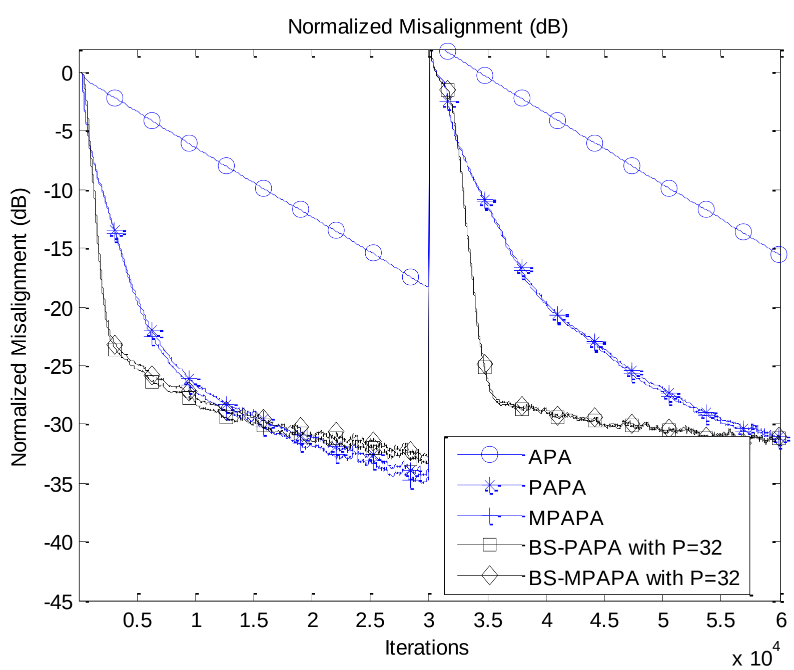


Figure 3. Comparison of APA, PAPA, MPAPA, BS-PAPA and BS-MPAPA algorithms for colored noise with *SNR*=30dB.

5. CONCLUSION

We have proposed two proportionate affine projection algorithms for block-sparse system identification, called block-sparse PAPA (BS-PAPA) and block-sparse memory PAPA (BS-MPAPA). Simulation results demonstrate that the new proportionate BS-PAPA and BS-MPAPA algorithms outperform traditional PAPA, MPAPA for block-sparse system identification.

6. REFERENCES

[1] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in network and acoustic echo cancellation*. Springer, 2001.

[2] K. Ozeki, and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electron. Commun. Jpn.*, 1984, 67-A, (5), pp. 19–27.

[3] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaption in echo cancellers," *IEEE Trans. on Speech Audio Process*., vol.8, no. 5, pp. 508-518, 2000.

[4] T. Gansler, J. Benesty, S. L. Gay, and M. Sondhi. "A robust proportionate affine projection algorithm for network echo cancellation." In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, vol. 2, pp. II793-II796. IEEE, 2000.

[5] C. Paleologu, S. Ciochina and J. Benesty, "An efficient proportionate affine projection algorithm for echo cancellation." *Signal Processing Letters, IEEE* 17, no. 2 (2010): 165-168.

[6] J. Benesty, C. Paleologu, and S. Ciochin, "Proportionate adaptive filters from a basis pursuit perspective," *Signal Processing Letters, IEEE*, vol. 17, no. 12, pp. 985-988, 2010.

[7] C. Paleologu, and J. Benesty, "Proportionate affine projection algorithms from a basis pursuit perspective." In *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pp. 2757-2760. IEEE, 2012.

[8] J. Liu and S. L. Grant, "A generalized proportionate adaptive algorithm based on convex optimization," in *Proc. Signals and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*, pp. 748-752, 2014.

[9] J. Liu and S. L. Grant, "Proportionate Adaptive Filtering for Block Sparse System Identification," submitted for possible publication at *IEEE Transactions on Audio Speech and Language Processing*.

III. BLOCK SPARSE MEMORY IMPROVED PROPORTIONATE AFFINE PROJECTION SIGN ALGORITHM

Jianming Liu and Steven L. Grant

Abstract

A block sparse memory improved proportionate affine projection sign algorithm (BS-MIP-APSA) is proposed for block sparse system identification under impulsive noise. The new BS-MIP-APSA not only inherits the performance improvement for block-sparse system identification, but also achieves robustness to impulsive noise and the efficiency of the memory improved proportionate affine projection sign algorithm (MIP-APSA). Simulations indicate that it can provide both faster convergence rate and better tracking ability under impulsive interference for block sparse system identification as compared to APSA and MIP-APSA.

1. INTRODUCTION

Adaptive filters have been widely used in various applications of system identification in which the normalized least mean square (NLMS) algorithm is well-known due to its simplicity, but suffers from slow convergence for colored input [1]. The affine projection algorithm (APA) provides better convergence for colored input compared with NLMS [2]. Meanwhile, the family of affine projection sign algorithm (APSA) has been proposed to improve the performance of APA under impulsive noise together with lower complexity [3]. In order to exploit the sparsity of some echo paths, the real-coefficient improved proportionate APSA (RIP-APSA) was proposed [4], and a memory improved proportionate APSA (MIP-APSA) was further proposed to achieve improved steady-state misalignment with similar computational complexity compared with RIP-APSA [5]. Recently, the block-sparse improved proportionate NLMS (BS-IPNLMS) algorithm was proposed to improve the performance of IPNLMS for identifying block-sparse systems [7]. In this Letter, motived by both BS-PNLMS and MIP-APSA, we will propose a block sparse memory improved proportionate APSA (BS-MIP-APSA) algorithm, which not only inherits the performance improvement for block-sparse system identification, but also achieves robustness to impulsive noise and the efficiency of MIP-APSA.

2. REVIEW OF MIP-APSA

For echo cancellation, the far-end signal $\mathbf{x}(n)$ is filtered through the echo path $\mathbf{h}(n)$ to get the desired signal $y(n)$,

$$y(n)=\mathbf{x}^T(n)\mathbf{h}(n)+v(n), \tag{1}$$

$$\mathbf{x}(n)=\left[x(n)\,x(n-1)\cdots x(n-L+1)\right]^T, \tag{2}$$

$$\mathbf{h}(n)=\left[h_0(n)\,h_1(n)\cdots h_{L-1}(n)\right]^T, \tag{3}$$

super-script T denotes transposition, L is the filter length, n is the time index, and $v(n)$ is the background noise plus near-end signals. Let $\hat{\mathbf{h}}(n)$ be the $L\times 1$ adaptive filter coefficient vector which estimates the true echo path vector $\mathbf{h}(n)$ at iteration n, and group the M most recent input vectors together:

$$\mathbf{X}(n)=\left[\mathbf{x}(n)\,\mathbf{x}(n-1)\cdots\mathbf{x}(n-M+1)\right], \tag{4}$$

$$\mathbf{e}(n)=\mathbf{y}(n)-\mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \tag{5}$$

$$\mathbf{y}(n)=\left[y(n)\,y(n-1)\cdots y(n-M+1)\right]^T, \tag{6}$$

where M is called the projection order. In [5], MIP-APSA proposed the following weight update:

$$\mathbf{g}(n)=\left[g_0(n),g_1(n),\cdots,g_{L-1}(n)\right], \tag{7}$$

$$g_l(n)=\frac{(1-\alpha)}{2L}+\frac{(1+\alpha)\left|\hat{h}_l(n)\right|}{2\sum_{i=0}^{L-1}\left|\hat{h}_i(n)\right|+\varepsilon}, \tag{8}$$

$$\mathbf{P}(n)=\left[\mathbf{g}(n)\odot\mathbf{x}(n),\mathbf{P}_{-1}(n-1)\right], \tag{9}$$

$$\mathbf{x}_{gs}(n)=\mathbf{P}(n)\operatorname{sgn}\left(\mathbf{e}(n)\right), \tag{10}$$

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \frac{\mu \mathbf{x}_{gs}(n)}{\sqrt{\delta + \mathbf{x}_{gs}^T(n)\mathbf{x}_{gs}(n)}}, \tag{11}$$

where $-1 \le \alpha < 1$, $l = 0,1,\cdots,L-1$, ε is a small positive constant that avoids division by zero, the operation $\odot$ denotes the Hadamard product, $\mathbf{P}_{-1}(n-1)$ contains the first $M-1$ columns of $\mathbf{P}(n-1)$, $y(n)$ takes the sign of each element of a vector, and δ is a small positive constant. Compared with RIP-PAPSA, MIP-PAPSA takes into account the 'proportionate history' from the last M moments of time. More details can be found in [5]-[6].

3. ALGORITHM DESIGN

In network echo cancellation, the network echo path is typically characterized by a bulk delay dependent on network loading, encoding, and jitter buffer delays and an “active” dispersive region in the range of 8-12 *ms* duration [1]. Meanwhile, it is well-known that NLMS is preferred over PNLMS for dispersive system. Therefore, considering the block-sparse characteristic of the network impulse response, the BS-PNLMS algorithm was proposed to improve the PNLMS algorithm by exploiting this special block-sparse characteristic, in which BS-PNLMS used the same step-size within each block and the step-sizes for each block were proportionate to their relative magnitude [7].

We propose to take in account the block-sparse characteristic and partition the MIP-APSA adaptive filter coefficients into N groups with group-length P, and $L = N \times P$,

$$\hat{\mathbf{h}}(n) = [\hat{\mathbf{h}}_0(n), \hat{\mathbf{h}}_1(n), \cdots, \hat{\mathbf{h}}_{N-1}(n)], \tag{12}$$

then the control matrix $\mathbf{g}(n)$ in (7)-(8) is be replaced by

$$\tilde{\mathbf{g}}(n) = \left[\tilde{g}_0(n)\mathbf{1}_P, \tilde{g}_1(n)\mathbf{1}_P, \cdots, \tilde{g}_{N-1}(n)\mathbf{1}_P\right], \tag{13}$$

$$\tilde{g}_k(n) = \frac{(1-\alpha)}{2L} + \frac{(1+\alpha)\left\|\hat{\mathbf{h}}_k(n)\right\|_2}{2P\sum_{i=0}^{N-1}\left\|\hat{\mathbf{h}}_i(n)\right\|_2 + \varepsilon}, \tag{14}$$

in which $\mathbf{1}_P$ is a P-length column vector of all ones, and $\left\|\hat{\mathbf{h}}_k(n)\right\|_2 = \sqrt{\sum_{j=1}^{P} \hat{h}_{kN+j}^2(n)}$, $k = 0,1,\cdots,N-1$. The weight update equation for BS-MIP-APSA is

$$\tilde{\mathbf{P}}(n) = \left[\tilde{\mathbf{g}}(n) \odot \mathbf{x}(n), \tilde{\mathbf{P}}_{-1}(n-1)\right], \tag{15}$$

$$\tilde{\mathbf{x}}_{gs}(n) = \tilde{\mathbf{P}}(n)\operatorname{sgn}\left(\mathbf{e}(n)\right), \tag{16}$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mu\tilde{\mathbf{x}}_{gs}(n)}{\sqrt{\delta + \tilde{\mathbf{x}}_{gs}^T(n)\tilde{\mathbf{x}}_{gs}(n)}}, \tag{17}$$

where $\tilde{\mathbf{P}}_{-1}(n-1)$ also contains the first $M-1$ columns of $\tilde{\mathbf{P}}(n-1)$.

It should be noted that the proposed BS-MIP-APSA includes both APSA and MIP-APSA. The MIP-APSA algorithm is a special case of proposed BS-MIP-APSA with group length $P=1$. Meanwhile, when P is chosen as L, the BS-MIP-APSA algorithm degenerates to APSA.

4. COMPLEXITY

Compared with traditional RIP-APSA and MIP-APSA, the extra computational complexity of the BS-MIP-APSA arises from the computation of the l_2 norm in (14), which requires L multiplications and N square roots. The complexity of the square root could be reduced through a look up table or Taylor series [7]. Meanwhile, the increase in complexity can be offset by the performance improvement as shown in the simulation results.

5. SIMULATION RESULTS

In our simulation, the echo path is a $L = 512$ finite impulse response (FIR) filter, and the adaptive filter is the same length. We generated colored input signals by filtering white Gaussian noise through a first order system with a pole at 0.8. Independent white Gaussian noise is added to the system background with a signal-to-noise ratio (SNR) of 40 dB. The impulsive noise with signal-to-interference ratio (SIR) of 0 dB is generated as a Bernoulli-Gaussian (BG) distribution. BG is a product of a Bernoulli process and a Gaussian process, and the probability for Bernoulli process is 0.1. The performance was evaluated through the normalized misalignment: $10\log_{10}(\|\mathbf{h}-\hat{\mathbf{h}}\|_2^2 / \|\mathbf{h}\|_2^2)$. In order to evaluate the tracking ability, we switch the echo path from the one-cluster block-sparse system of Figure 1(a) to the two-cluster block-sparse system of Figure 1(b).

The APSA and MIP-APSA algorithms are compared with BS-MIP-APSA. The parameters are $\mu = 0.001$, $\varepsilon = 0.01$, $\delta = 0.01$, $\alpha = 0$, $M = 2$, and $P = 4$. In the first case, we show the normalized misalignment for colored input in Figure 2. We could see that the proposed BS-MIP-APSA achieves both faster convergence rate and better tracking ability. In Figure 3, the performance of BS-MIP-APSA is compared with APSA and MIP-APSA for speech input signal, and we found that our proposed algorithm demonstrates better performance too.

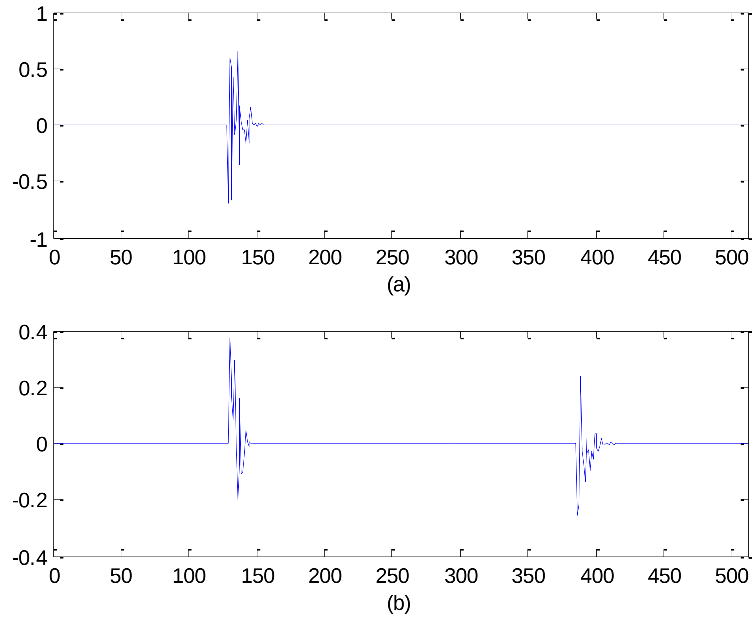


Figure 1. Two block-sparse systems used in the simulations: (a) one-cluster block-sparse system, (b) two-cluster block-sparse system.

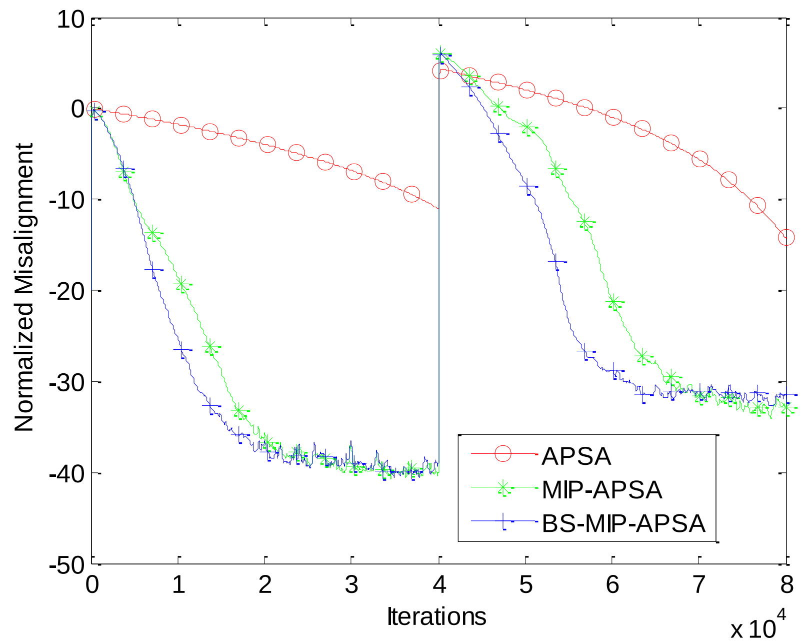


Figure 2. Normalized misalignment of APSA, MIP-APSA, and BS-MIP-APSA for colored input signal.

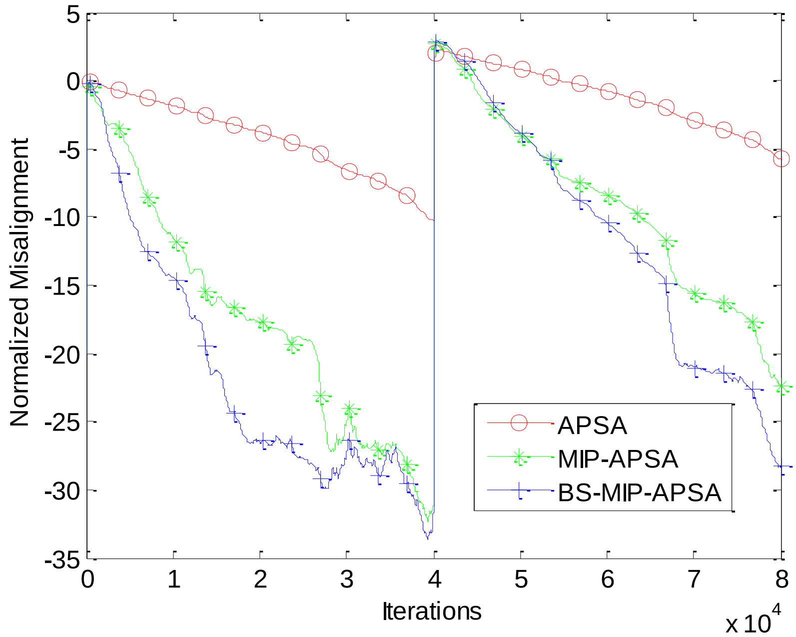


Figure 3. Normalized misalignment of APSA, MIP-APSA, and BS-MIP-APSA for speech input signal.

6. CONCLUSION

We have proposed a block-sparse memory improved affine projection sign algorithm to improve the performance of block-sparse system identification. Simulations demonstrate the proposed algorithm has both faster convergence speed and tracking ability for block-sparse system identification compared with APSA and MIP-APSA algorithms.

7. REFERENCES

[1] Benesty, J., Gänsler, T., Morgan, D.R., Sondhi, M.M, and Gay, S.L., Advances in network and acoustic echo cancellation. Springer, 2001.

[2] Ozeki, K., and Umeda, T.: "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties", *Electron. Commun. Jpn.*, 1984, 67-A, (5), pp. 19-27

[3] Shao, T., Zheng, Y. R., and Benesty, J.: "An affine projection sign algorithm robust against impulsive interferences", *IEEE Signal Process. Lett.*, 2010, 17, (4), pp. 327–330.

[4] Yang, Z., Zheng, Y.R., and Grant, S.L.: "Proportionate Affine Projection Sign Algorithms for Network Echo Cancellation", *IEEE Trans. Audio, Speech, Lang. Process.*, 19, (8), pp. 2273-2284.

[5] Albu, F., and Kwan, H.K.: "Memory improved proportionate affine projection sign algorithm", *Electron. Lett.*, 2012, **48**, (20), pp. 1279-1281.

[6] Paleologu, C., Ciochina, S., and Benesty, J.: "An efficient proportionate affine projection algorithm for echo cancellation", *IEEE Signal Process. Lett.*, 2010, 17, (2), pp. 165–168.

[7] Liu, J., and Grant, S.L.: "Proportionate adaptive filtering for block-sparse system identification", *arXiv preprint arXiv:1508.04172, 2015*.

IV. A LOW COMPLEXITY REWEIGHTED PROPORTIONATE AFFINE PROJECTION ALGORITHM WITH MEMORY AND ROW ACTION PROJECTION

Jianming Liu and Steven L. Grant

Abstract

A new reweighted proportionate affine projection algorithm (RPAPA) with memory and row action projection (MRAP) is proposed in this paper. The reweighted PAPA is derived from a family of sparseness measures, which demonstrate performance similar to mu-law and the l_0 norm PAPA but with lower computational complexity. The sparseness of the channel is taken into account to improve the performance for dispersive system identification. Meanwhile, the memory of the filter's coefficients is combined with row action projections (RAP) to significantly reduce computational complexity. Simulation results demonstrate that the proposed RPAPA MRAP algorithm outperforms both the affine projection algorithm (APA) and PAPA, and has performance similar to l_0 PAPA and mu-law PAPA, in terms of convergence speed and tracking ability. Meanwhile, the proposed RPAPA MRAP has much lower computational complexity than PAPA, mu-law PAPA, and l_0 PAPA, etc., which makes it very appealing for real time implementation.

1. INTRODUCTION

Adaptive filtering has been studied for decades and has found wide areas of application. The most common adaptive filter is the normalized least mean square (NLMS) algorithm due to its simplicity and robustness [1]. In the 1990's, the affine projection algorithm (APA), a generalization of NLMS was found to have better convergence than NLMS for colored input [2]-[3]. The optimal step size control of the adaptive algorithm has been widely studied in order to improve their performance [4]-[5]. The impulse responses in many applications, such as network echo cancellation (NEC), are sparse, that is, a small percentage of the impulse response components have a significant magnitude while the rest are zero or small. To exploit this property, the family of proportionate algorithms was proposed to improve performance in such applications [2]. These algorithms include proportionate NLMS (PNLMS) [6]-[7], and proportionate APA (PAPA) [8], etc.

The idea behind proportionate algorithms is to update each coefficient of the filter independently of the others by adjusting the adaptation step size in proportion to the magnitude of the estimated filter coefficient [6]. In comparison to NLMS and APA, PNLMS and PAPA have very fast initial convergence and tracking when the echo path is sparse. However, the big coefficients converge very quickly (in the initial period) at the cost of slowing down dramatically the convergence of the small coefficients (after the initial period). In order to combat this issue, mu-law PNLMS (MPNLMS) and mu-law PAPA algorithms were proposed [9]-[11]. Furthermore, the l_0 norm family of algorithms have recently drawn lots of attention for sparse system identification [12]. Therefore, a new PNLMS algorithm based on the l_0 norm was proposed to represent a better measure of sparseness than the l_1 norm in PNLMS [13].

On the other hand, the PNLMS and PAPA algorithms converge much slower than corresponding NLMS and APA algorithms when the impulse response is dispersive. In response, the improved PNLMS (IPNLMS) and improved PAPA (IPAPA) were proposed by introducing a controlled mixture of proportionate and non-proportionate adaptation [14]-[15]. The IPNLMS and IPAPA algorithms perform very well for both sparse and non-sparse systems. Also, recently, the block-sparse PNLMS (BS-PNLMS) algorithm was proposed to improve the performance of PNLMS for identifying block-sparse systems [16].

In order to reduce the computational complexity of PAPA, the memory improved PAPA (MIPAPA) algorithm was proposed to not only speed up the convergence rate but also reduce computational complexity by taking into account the memory of the proportionate coefficients [17]. Dichotomous coordinate descent (DCD) iterations have previous been applied to the PAPA family of algorithms to implement the MIPAPA adaptive filter [18]-[19]. Meanwhile, an iterative method based on the PAPA with row action projection (RAP) has been shown to have good convergence properties with relatively low complexity [20].

In [21] the proportionate adaptive filter was derived from a unified view of variable-metric projection algorithms. In addition, the PNLMS algorithm and PAPA can both be deduced from a basis pursuit perspective [22]-[23]. A more general framework was further proposed to derive PNLMS adaptive algorithms for sparse system identification, which employed convex optimization [24]. Here, a family of PAPA algorithms are firstly derived based on convex optimization, in which PAPA, mu-law PAPA, and l_0 PAPA are all special cases. Then, a reweighted PAPA is suggested in order to reduce the computational complexity. Finally, an efficient implementation of PAPA is proposed based on RAP and memory PAPA.

The organization of this article is as follows. The review of various PAPAs is presented in Section 2. Section 3 derives the proposed reweighted PAPA and presents an efficient memory implementation with RAP. The computational complexity is compared with PAPA, mu-law PAPA and l_0 PAPA in Section 4. In Section 5, simulation results of the proposed algorithm are presented. The last section concludes the paper with remarks.

2. REVIEW OF VARIOUS PAPAS

The input signal $\mathbf{x}(n)$ is filtered through the unknown coefficients to be identified $\mathbf{h}(n)$ to get the observed output signal $d(n)$.

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n), \tag{1}$$

where

$$\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-L+1)]^T,$$

and $v(n)$ is the measurement noise, and L is the length of impulse response. We define the estimated error as

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \tag{2}$$

where $\hat{\mathbf{h}}(n)$ is the adaptive filter's coefficients. Grouping the M most recent input vectors $\mathbf{x}(n)$ together gives the input signal matrix

$$\mathbf{X}(n) = [\mathbf{x}(n), \mathbf{x}(n-1), \cdots, \mathbf{x}(n-M+1)].$$

Therefore, the estimated error vector is

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \tag{3}$$

in which

$$\mathbf{d}(n) = [d(n), d(n-1), \cdots, d(n-M+1)],$$

$$\mathbf{e}(n) = [e(n) e(n-1) \cdots e(n-M+1)],$$

where M is the projection order. PAPA updates the filter coefficients as follows [8]:

$$\mathbf{P}(n) = \mathbf{G}(n-1)\mathbf{X}(n) \tag{4}$$

$$\begin{aligned}\hat{\mathbf{h}}(n) &= \hat{\mathbf{h}}(n-1) \\ &\quad + \mu\mathbf{P}(n)\left(\mathbf{X}^T(n)\mathbf{P}(n) + \delta\mathbf{I}_M\right)^{-1}\mathbf{e}(n).\end{aligned} \tag{5}$$

in which μ is the step-size, δ is the regularization parameter, $\mathbf{I}_M$ is the $M \times M$ identity matrix, and the proportionate step-size control matrix $\mathbf{G}(n-1)$ is defined as

$$\mathbf{G}(n-1) = diag\{\mathbf{g}(n-1)\}, \tag{6}$$

$$\mathbf{g}(n-1) = diag\left[g_1(n-1), g_2(n-1), \cdots, g_L(n-1)\right], \tag{7}$$

$$g_l(n-1) = \frac{\gamma_l(n-1)}{\frac{1}{L}\sum_{i=1}^{L} \gamma_i(n-1)}, \tag{8}$$

$$\gamma_l = \max\left\{\rho \max\left\{q, \left|\hat{h}_1\right|, \cdots, \left|\hat{h}_L\right|\right\}, \left|\hat{h}_l\right|\right\}, \tag{9}$$

where $\mathbf{I}_M$ is specific to the algorithm, q prevents the filter coefficients $\hat{h}_l(n-1)$ from stalling when $\hat{\mathbf{h}}(0) = \mathbf{0}_{L\times 1}$ at initialization and ρ prevents the coefficients from stalling when they are much smaller than the largest coefficient. The classical PAPA employs step-sizes that are proportional to the magnitude of the estimated impulse response as below [8]

$$\mathrm{F}\left(\hat{h}_l\right) = \left|\hat{h}_l\right|, \tag{10}$$

The mu-law PNLMS and the mu-law PAPA algorithm proposed in [9]-[11] use the logarithm of the coefficient magnitudes rather than magnitudes directly as below:

$$\mathrm{F}\left(\hat{h}_l\right) = \ln\left(1 + \sigma_\mu \left|\hat{h}_l\right|\right), \tag{11}$$

in which σ_μ is a positive parameter. Based on the motivation that the l_0 norm can represent an even better measure of sparseness than the l_1 norm, the improved PNLMS and PAPA algorithms based on an approximation of the l_0 norm (l_0-PNLMS) were proposed as below [13]:

$$\mathrm{F}\left(\hat{h}_l\right) = 1 - e^{-\sigma_{l0}\left|\hat{h}_l\right|}, \tag{12}$$

where σ_{l0} is a positive parameter. The main disadvantage of the mu-law or l_0 norm PAPA algorithms are their heavy computation cost because of the L logarithmic or exponential operations. Therefore, a line segment was given to approximate the mu-law function [9], where

$$\mathrm{F}\left(\hat{h}_l\right) = \begin{cases} 200\left|\hat{h}_l\right|, & \left|\hat{h}_l\right| < 0.005 \\ 1, & otherwise. \end{cases} \tag{13}$$

It should be noted that, without loss of performance, the line segment was normalized to be of unit gain for $\left|\hat{h}_l\right| \geq 0.005$, compared to the original one proposed in [9]. Meanwhile, the exponential form in (12) can be approximated by the first order Taylor series expansions of exponential functions [12]

$$e^{-\sigma_{l0}\left|\hat{h}_l\right|} = \begin{cases} 1-\sigma_{l0}\left|\hat{h}_l\right|, & \left|\hat{h}_l\right| < \dfrac{1}{\sigma_{l0}} \\ 0, & otherwise. \end{cases} \tag{14}$$

Then (12) becomes

$$\mathrm{F}\left(\hat{h}_l\right) = \begin{cases} \sigma_{l0}\left|\hat{h}_l\right|, & \left|\hat{h}_l\right| < \dfrac{1}{\sigma_{l0}} \\ 1, & otherwise. \end{cases} \tag{15}$$

It is interesting to see that the first order Taylor series approximation of l_0 PAPA in (12) is actually the same as the line segment implementation of mu-law PAPA in (11) for $\sigma_{l0} = 200$.

3. THE PROPOSED SC-RPAPA WITH MRAP

Based on the minimization of the convex target, the reweighted PAPA (RPAPA) will be firstly derived from a new sparseness measure with low computational complexity. Meanwhile, the sparseness controlled RPAPA (SC-RPAPA) is presented to improve the performance for both sparse and dispersive system identification. Finally, the SC-RPAPA with memory and RAP (MRAP) is proposed by combing the memory of the coefficients with iterative RAP to further reduce the computational complexity.

3.1 THE PROPOSED RPAPA

The proportionate APA algorithm can be deduced from a *basis pursuit* perspective [22]

$$\begin{aligned} &\boldsymbol{min} && \left\|\tilde{\mathbf{h}}(n)\right\|_1 \\ &\text{subject to} && \mathbf{d}(n)=\mathbf{X}^T(n)\tilde{\mathbf{h}}(n), \end{aligned} \tag{16}$$

where $\tilde{\boldsymbol{h}}(n)$ is the correction component defined as

$$\tilde{\mathbf{h}}(n)=\mathbf{G}(n-1)\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n)\right]^{-1}\mathbf{d}(n).$$

According to [24], the family of PAPA algorithms can be derived from the following target

$$\begin{aligned} &\boldsymbol{min} && \int\mathbf{G}^{-1}(n-1)\tilde{\mathbf{h}}(n)d\tilde{\mathbf{h}} \\ &\text{subject to} && \mathbf{d}(n)=\mathbf{X}^T(n)\tilde{\mathbf{h}}(n), \end{aligned} \tag{17}$$

where $\mathbf{G}^{-1}(n-1)$ is the inverse matrix of proportionate matrix $\mathbf{G}(n-1)$, which is also a diagonal matrix. If the optimization target in (17) is convex, the family of PAPA algorithms can be derived using Lagrange Multipliers. It should be noted that, using the approximation,

$$\int\mathbf{G}^{-1}(n-1)\tilde{\mathbf{h}}(n)d\tilde{\mathbf{h}}\approx\frac{1}{2}\tilde{\mathbf{h}}^T(n)\mathbf{G}^{-1}(n-1)\tilde{\mathbf{h}}(n), \tag{18}$$

the proposed formulation in (17) becomes the variable-metric in [21], which is an approximation of the proposed formulation. The function $\mathrm{G}(t), t \in \mathbb{R}$ should satisfy the following properties:

- $\mathrm{G}(0)=0$, $\mathrm{G}(t)$ is even and not identically zero;
- $\mathrm{G}(t)$ is non-decreasing on $[0,\infty)$;
- $\mathrm{G}(t)/t$ is non-increasing on $(0,\infty)$.

The above properties follow the requirements of the sparseness measure proposed in [25]. From the perspective of proportionate algorithms, the first two requirements are intuitive, since the family of the proportionate algorithms should be proportionate to the magnitude of the filter's coefficients. The third property will guarantee the convexity of the optimization target. PAPA, mu-law PAPA and l_0 PAPA are all special cases of the sparseness measures fulfilling all three properties. In this paper, considering the computational complexity, we propose using the following reweighted PAPA:

$$\mathrm{F}\left(\hat{h}_l\right)=\frac{\left|\hat{h}_l\right|}{\left|\hat{h}_l\right|+\sigma_r}, \tag{19}$$

where σ_r is a small positive constant.

The proposed reweighted metric is compared with PAPA, mu-law PAPA and l_0 PAPA in Figure 1. The σ parameters for each algorithm were $\sigma_\mu=1000$, $\sigma_{l0}=50$, $\sigma_r=0.01$. These parameters were recommended and widely simulated in the literature for each algorithm [9] [13]. It should be noted that, the plots in [24] set the σ parameters respectively so that they all contain the point $(0.9, 0.9)$. However, in actual application, this parameter should be tuned to maximize the performance. In order to facilitate the comparison of the different sparseness measure, they are normalized to pass through the point, $(1,1)$ here instead. Without loss of generality, it is assumed that the filter's coefficients are normalized and the maximum possible magnitude is 1. Therefore, it is convenient to compare the gain distribution of different metrics with different σ parameters.

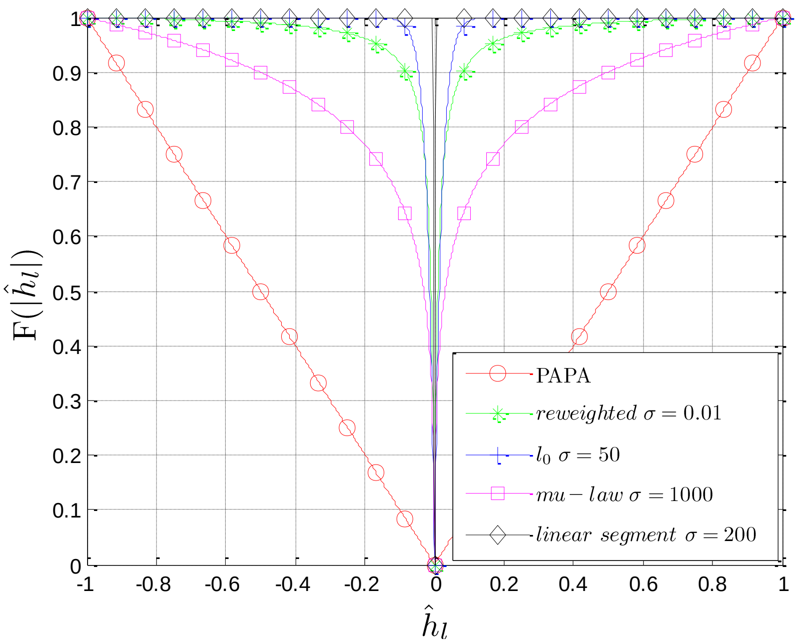


Figure 1. Comparison of the different metrics.

3.2 THE PROPOSED SC-RPAPA

It should be noted that the reweighting factor σ_r in the proposed RPAPA (19) is related to the sparseness of the impulse system. It is straightforward to verify that if $\sigma_r = 0$, reweighted PAPA simplifies to APA. If the impulse system is more sparse, σ_r should be relatively larger than $|\hat{h}_l|$, which makes it more like the PAPA. This agrees with the fact that we fully benefit from PNLMS only when the impulse response is close to a delta function [26]. Therefore, it is natural to take the sparseness of impulse response into account. The sparsity of an impulse response could be estimated as

$$\hat{\varepsilon}(n) = \frac{L}{L-\sqrt{L}}\left(1 - \frac{\|\hat{\mathbf{h}}(n)\|_1}{\sqrt{L}\|\hat{\mathbf{h}}(n)\|_2}\right), \tag{20}$$

where $L > 1$ is the length of the channel, $\|\hat{\mathbf{h}}(n)\|_1$ and $\|\hat{\mathbf{h}}(n)\|_2$ are the l_1 norm and l_2 norm of $\hat{\mathbf{h}}(n)$, respectively. The value of $\hat{\varepsilon}(n)$ is between 0 and 1. For a sparse channel, the value of the sparseness is close to 1 and for a dispersive channel, this value is close to 0. Therefore, the SC-RPAPA is

$$\mathrm{F}\left(\hat{h}_l\right)=\frac{\left|\hat{h}_l\right|}{\left|\hat{h}_l\right|+\hat{\varepsilon}\left(n\right)\sigma_{\max}}, \tag{21}$$

where $\sigma_{\max}$ is the maximum value for the sparse system identification. The plot of the reweighted metric for different σ s is presented in Figure 2. In practical implementation, we would like to apply the APA algorithm to the dispersive system under certain sparseness threshold. For example, the sparsity of the dispersive channel is about 0.4, and a heuristic implementation that works pretty well in the simulations is

$$\mathrm{F}\left(\hat{h}_l\right)=\frac{\left|\hat{h}_l\right|}{\left|\hat{h}_l\right|+\max\left\{\hat{\varepsilon}\left(n\right)-0.4,\varepsilon_{\min}\right\}\sigma_{\max}}, \tag{22}$$

where $\varepsilon_{\min}=1e^{-4}$ is a minimum sparsity in order to avoid dividing by zero for $\left|\hat{h}_l\right|=0$.

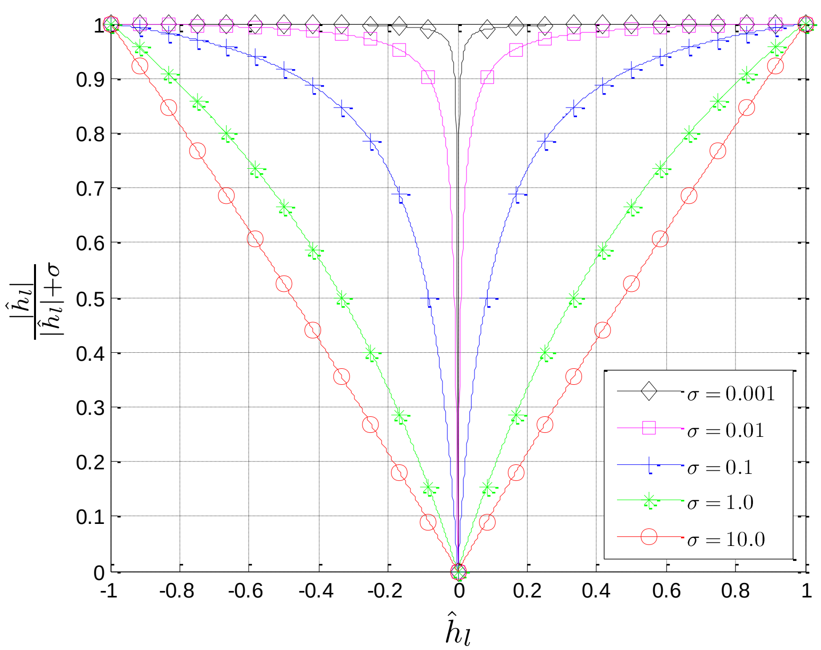


Figure 2. Reweighted metric with different σ parameters.

3.3 THE PROPOSED SC-RPAPA WITH MRAP

However, the main computational complexity of the family of PAPA algorithm is the matrix inversion in (5). Reduction in complexity is achieved by using $5M$ DCD iterations, thus requiring about $10M^2$ additions [18]. Meanwhile, a sliding-window recursive least squares (SRLS) low-cost implementation of PAPA is given based on DCD, which does not depend on M. The SRLS implementation is only efficient when the projection order is very high (e.g., such as $M = 512$) [19]. However, it is known that if the projection order increases, the convergence speed is faster, but the steady-state error also increases.

Another way to avoid the matrix inversion altogether is to use the method of RAP [27]. RAP is also known in the literature as a *data reuse* algorithm (see [28]). It has been shown in [29] that RAP is effectively the same as APA, except that the system of equations problem that is solved with a direct matrix inversion (DMI) in APA is solved iteratively in RAP [30].The iterative PAPA algorithm proposed in [31] was made efficient by implementing it using RAP in [27]. RAP is an iterative approach to solving a system of M equations. It cycles through the M equations J times performing an NLMS-like update on the coefficients for each equation. In this instance, the number of RAP iterations, J is set to one. It should be noted that, by limiting J to one, the solution of the system of equations through RAP is approximate. However, the simulation results will demonstrate that this approximation works pretty well, especially for relatively high projection order. In each sample period a new equation is added to the system of equations and the oldest equation is dropped. Thus, M RAP updates are performed on a given equation every M sample periods. The PAPA algorithm with RAP updates the coefficients

$$
\begin{array}{ll}
\textit{Initialize} & \hat{\mathbf{h}}^{[0]} = \hat{\mathbf{h}}(n-1) \\
\textit{Loop} & m = 0,1,\cdots,M-1 \\
 & \alpha(m) = \mu \big/ \left(\mathbf{x}^T(n-m)\mathbf{P}_m(n) + \delta \right) \\
 & e^{[m]} = d(n-m) - \mathbf{x}^T(n-m)\hat{\mathbf{h}}^{[m]} \\
 & \hat{\mathbf{h}}^{[m+1]} = \hat{\mathbf{h}}^{[m]} + \alpha(m)\mathbf{P}_m(n)e^{[m]} \\
 & m = m+1 \\
\textit{Update} & \hat{\mathbf{h}}(n) = \hat{\mathbf{h}}^{[M]}
\end{array}
$$

where $\mathbf{P}_m(n)$ is the m_{th} column of $\mathbf{P}(n)$ defined as

$$\mathbf{P}_m(n) = \mathbf{g}(n-1) \odot \mathbf{x}(n-m),$$

the operation $\odot$ denotes the Hadamard product and $m = 0, 1, \cdots, M-1$.

The traditional PAPA requires $M \times L$ multiplications to calculate $\mathbf{P}(n)$, and in order to further reduce the computational complexity, we propose to apply the memory of the proportionate coefficients [17] into SC-RPAPA. Therefore, the matrix $\mathbf{P}(n)$ in (4) can be approximated as $\mathbf{P}'(n)$

$$\mathbf{P}'(n) = \left[\mathbf{g}(n-1) \odot \mathbf{x}(n), \mathbf{P}'_{-1}(n-1)\right], \tag{23}$$

where $\mathbf{P}'_{-1}(n-1)$ contains the first $M-1$ columns of $\mathbf{P}'(n-1)$. Meanwhile, we define

$$\mathbf{p}(n) = [p_0(n), p_1(n), \cdots, p_{M-1}(n)],$$

in which

$$p_m(n) = \boldsymbol{x}^T(n-m)\mathbf{P}'_m(n),$$

and $\mathbf{P}'_m(n)$ is the m_{th} column of $\mathbf{P}'(n)$ defined as

$$\mathbf{P}'_m(n) = \mathbf{g}(n-m-1) \odot \mathbf{x}(n-m).$$

Considering the time-shift property, the calculation of $\mathbf{p}(n)$ could be

$$\mathbf{p}(n) = \left[\boldsymbol{x}^T(n)\mathbf{P}'_0(n), \mathbf{p}_{-1}(n-1)\right], \tag{24}$$

where $\mathbf{p}_{-1}(n-1)$ contains the first $M-1$ values of $\mathbf{p}(n-1)$. The proposed update for the PAPA with memory and RAP is

$$\begin{array}{ll}
\textit{Initialize} & \hat{\mathbf{h}}^{[0]} = \hat{\mathbf{h}}(n-1) \\
\textit{Loop} & m = 0, 1, \cdots, M-1 \\
 & \alpha(m) = \mu / \left(p_m(n) + \delta\right) \\
 & e^{[m]} = d(n-m) - \mathbf{x}^T(n-m)\hat{\mathbf{h}}^{[m]} \\
 & \hat{\mathbf{h}}^{[m+1]} = \hat{\mathbf{h}}^{[m]} + \alpha(m)\mathbf{P}'_m(n)e^{[m]} \\
 & m = m+1 \\
\textit{Update} & \hat{\mathbf{h}}(n) = \hat{\mathbf{h}}^{[M]}
\end{array}$$

As mentioned in [17], the proposed RPAPA with MRAP takes into account the "history" of the proportionate factors from the last M steps. The convergence and the tracking become faster when the projection order increases. Meanwhile, combined with the RAP, the computational complexity is also significantly lower as compared to the MPAPA through avoiding the direct matrix inversion and using the memory. The proposed SC-RPAPA with MRAP algorithm is summarized in detail in Table 1.

Table 1. The SC-RPAPA algorithm with MRAP

Initialization	$\hat{\mathbf{h}}(0)=\mathbf{0}_{L\times 1}$, $\rho=0.01$, $q=0.01$, $\delta=0.01/L$, $\sigma_{\max}=0.02$, $\varepsilon_{\min}=1e^{-4}$, $\mu=0.2$
Sparseness control	$\hat{\varepsilon}(n)=\frac{L}{L-\sqrt{L}}\left(1-\frac{\left\Vert\hat{\mathbf{h}}(n)\right\Vert_1}{\sqrt{L}\left\Vert\hat{\mathbf{h}}(n)\right\Vert_2}\right)$ $\mathrm{F}\left(\hat{h}_l\right)=\frac{\left\vert\hat{h}_l\right\vert}{\left\vert\hat{h}_l\right\vert+\max\left\{\hat{\varepsilon}(n)-0.4,\varepsilon_{\min}\right\}\sigma_{\max}}$ $\gamma_l=\max\left\{\rho\max\left\{q,\left\vert\hat{h}_1\right\vert,\cdots,\left\vert\hat{h}_L\right\vert\right\},\left\vert\hat{h}_l\right\vert\right\}$ $g_l(n-1)=\frac{\gamma_l(n-1)}{\frac{1}{L}\sum_{i=1}^{L}\gamma_i(n-1)}$ $\mathbf{g}(n-1)=diag\left[g_1(n-1),g_2(n-1),\cdots,g_L(n-1)\right]$
Memory update	$\mathbf{P'}(n)=\left[\mathbf{g}(n-1)\odot\mathbf{x}(n),\mathbf{P'}_{-1}(n-1)\right]$ $\mathbf{p}(n)=\left[\mathbf{x}^T(n)\mathbf{P}_0^{'}(n),\mathbf{p}_{-1}(n-1)\right]$
Error output	$e(n)=d(n)-\mathbf{x}^T(n)\hat{\mathbf{h}}(n-1)$
RAP iteration	$\hat{\mathbf{h}}^{[0]}=\hat{\mathbf{h}}(n-1)$ $for\ \ m=0,1,\cdots,M-1$ $\alpha(m)=\mu/\left(p_m(n)+\delta\right)$ $e^{[m]}=d(n-m)-\mathbf{x}^T(n-m)\hat{\mathbf{h}}^{[m]}$ $\hat{\mathbf{h}}^{[m+1]}=\hat{\mathbf{h}}^{[m]}+\alpha(m)\mathbf{P'}_m(n)e^{[m]}$ $m=m+1$
Filter update	$\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}^{[M]}$

4. COMPUTATIONAL COMPLEXITY

The computational complexity of the SC-RPAPA with MRAP algorithm is compared with traditional PAPA, MPAPA, RPAPA, and SC-RPAPA in Table 2, in terms of the total number of additions (A), multiplications (M), divisions (D), comparisons (C), square root (Sqrt), and direct matrix inversion (DMI) needed per algorithm iteration. All the algorithms require L $|\cdot|$ operations for calculating the magnitude of the filter's coefficients.

Compared with traditional PAPA, the MPAPA reduced the complexity of $\mathbf{GX}$, but the calculation of $\mathbf{X}^T\mathbf{P}'$ still requires M^2L multiplications. Meanwhile, due to the memory and the iterative RAP structure, only L multiplications are needed to update $\boldsymbol{p}(n)$ instead.

What's more important is that, both the PAPA and the MPAPA algorithms require a $M \times M$ direct matrix inversion, which is especially expensive for high projection orders. The combination of the memory and the iterative RAP structure, not only avoids the $M \times M$ direct matrix inversion, but also reduces the computational complexity required for the calculation of both $\mathbf{GX}$ and $\mathbf{X}^T\mathbf{GX}$.

The additional computational complexity for the SC-RPAPA with MRAP algorithm arises from the computation of the sparseness measure $\hat{\varepsilon}$. As in [32], given that $L/\left(L-\sqrt{L}\right)$ can be computed offline, the remaining *l*-norms require an additional $2L$ additions and L multiplications. Furthermore, this sparseness measure can be reused in many other sparseness controlled algorithms too, for example [32]. The calculation of the $\mathbf{F}$ in (22) requires additional L divisions, $L+1$ additions, one multiplication, and one comparison more than PAPA. The complexity of division is much lower than the L exponential or logarithmic operations required by either the mu-law or the l_0 PAPA. Meanwhile, (22) also offers the robustness to dispersive system identification.

Table 2. Computational complexity of the algorithms' coefficient updates.

Algorithm	A	M	D	C	Sqrt	DMI
PAPA	$(M^2+2M+1)\ L-M-1$	$(M^2+3M+1)\ L+2M^2+2$	L	2L	0	Yes, $M\times M$
MPAPA	$(M^2+2M+1)\ L-M-1$	$(M^2+3M+1)\ L+2M^2+2$	L	$2L$	0	Yes, $M\times M$
RPAPA	$(M^2+2M+1)\ L-M-1$	$(M^2+3M+1)\ L+2M^2+2$	$2L$	$2L$	*0*	Yes, $M\times M$
SC-RPAPA	$(M^2+2M+1)\ L-M-1$	$(M^2+3M+1)\ L+2M^2+2$	$2L+1$	$2L$ +1	1	Yes, $M\times M$
SC-RPAPA MRAP	$4L+N-1$	$(M^2+3M+1)\ L+2M^2+2$	$2L+M+1$	$2L+1$	1	Yes, $M\times M$

5. SIMULATION RESULTS

The performance of the proposed SC-RPAPA with MRAP was evaluated via simulations. Throughout our simulation, the length of the unknown system was $L=512$, and the adaptive filter was with the same length. The sampling rate was 8 kHz. The parameters for each algorithm were $\delta = 0.01/L$, $\rho = 0.01$, $q = 0.01$. The step-size for all the algorithms was set to $\mu = 0.2$.

The algorithms were tested using both the white Gaussian noise (WGN), and colored noise as inputs. The colored input signals were generated by filtering the WGN through a first order system with a pole at 0.8. Independent WGN was added to the system background with a signal-to-noise ratio (SNR) as 30dB.

Two impulse responses were used to verify the performance of the proposed SC-RPAPA MRAP algorithm, as shown in Figure 3. The first one in Figure 3.(a) is a sparse impulse response of typical network echo with sparseness 0.92. Figure 3.(b) is a dispersive channel with sparseness 0.44. In order to demonstrate the tracking ability, an echo path change was incurred through switching the impulse response from the sparse system in Figure 3.(a) to the dispersive one in Figure 3.(b). The convergence state of adaptive filter is evaluated with the normalized misalignment which is defined as

$$20\log_{10}\left(\left\|\mathbf{h}-\hat{\mathbf{h}}\right\|_2 \Big/ \left\|\mathbf{h}\right\|_2\right)$$

5.1 THE PERFORMANCE OF THE PROPOSED RPAPA

The proposed reweighted PAPA in (19) was firstly compared to PAPA, mu-law PAPA, and l_0 PAPA. The parameters for the algorithm were $\sigma_\mu = 1000$, $\sigma_{l0} = 200$, and $\sigma_r = 0.01$. The affine projection order was selected as $M = 2$.

In the first simulation shown in Figure 4, the input signal was the WGN. According to the results, the proposed RPAPA could outperform PAPA, and has similar performance with respect to mu-law and l_0 PAPA. However, the reweighted PAPA has much lower computational complexity. In the second simulation, the input signal was colored, and a similar result could be obtained according to Figure 5.

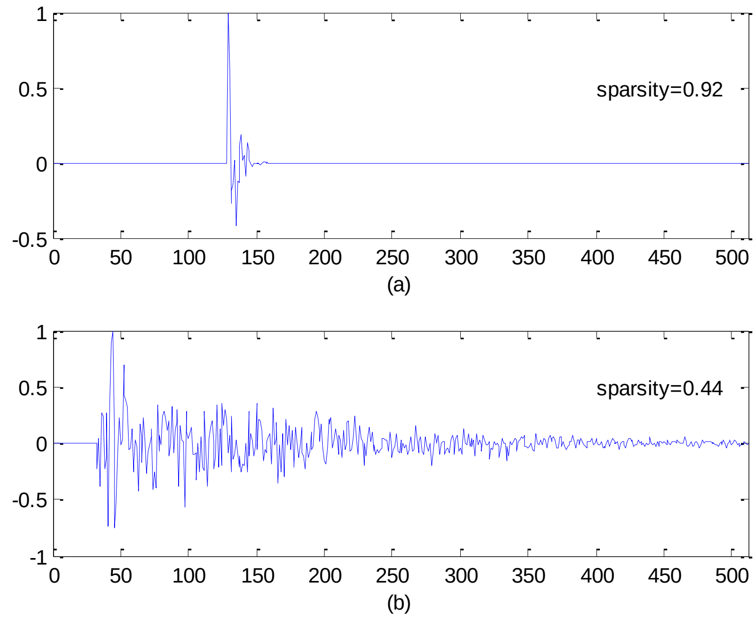


Figure 3. Two impulse responses used in the simulation (a) the sparse network echo path, and (b) the dispersive echo path.

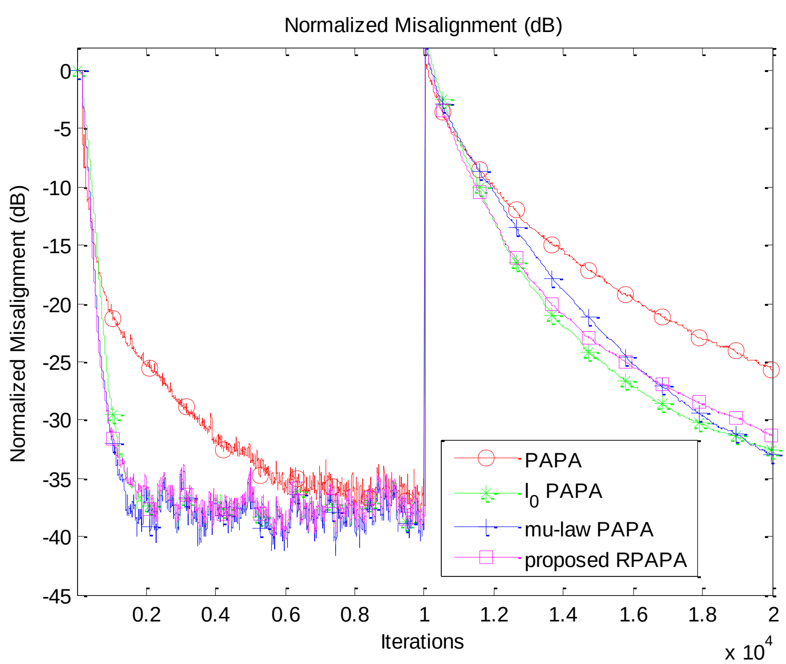


Figure 4. Comparison of RPAPA with PAPA, l_0 PAPA and mu-law PAPA for WGN input, SNR=30 dB, $M=2$, $\mu=0.2$.

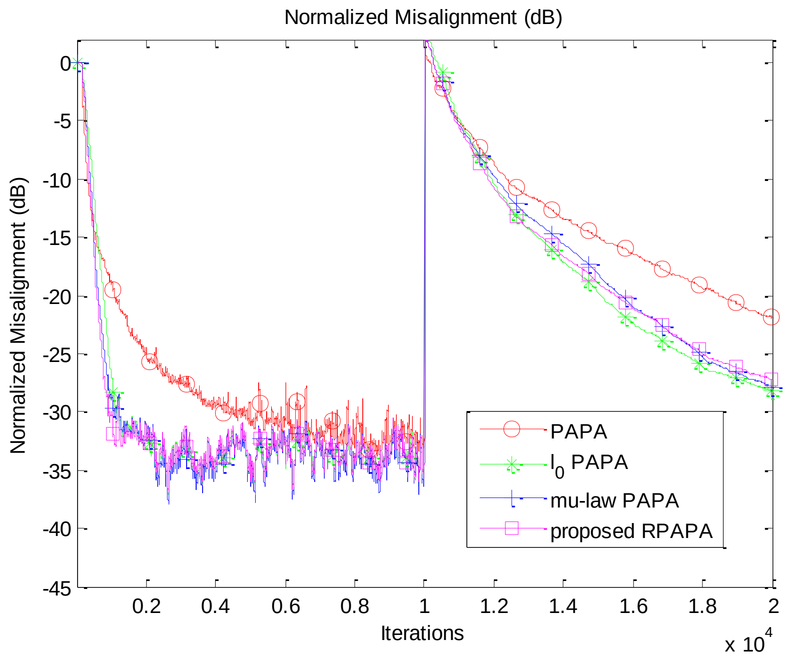


Figure 5. Comparison of RPAPA with PAPA, l_0 PAPA and mu-law PAPA for colored input, SNR=30 dB, $M = 2$, $\mu = 0.2$.

5.2 THE PERFORMANCE OF THE PROPOSED SC-RPAPA

To demonstrate the benefit of sparseness control, the proposed SC-RPAPA algorithm was simulated using an echo path change from the sparse to the dispersive impulse response in Figure 3. The SC-RPAPA algorithm was compared with APA, PAPA, and the above RPAPA algorithms. The parameters for the algorithm were $\sigma_r = 0.01$, and $\sigma_{\max} = 0.02$. The affine projection order was selected as $M = 2$. In Figure 6, the input signal was the WGN input. Both the proposed RPAPA and SC-RPAPA algorithms had similar performance for sparse system identification, which outperformed APA and PAPA. Meanwhile, due to the sparseness control, SC-RPAPA outperformed RPAPA as expected for the dispersive system. The colored input was used in Figure 7, and similar results are observed.

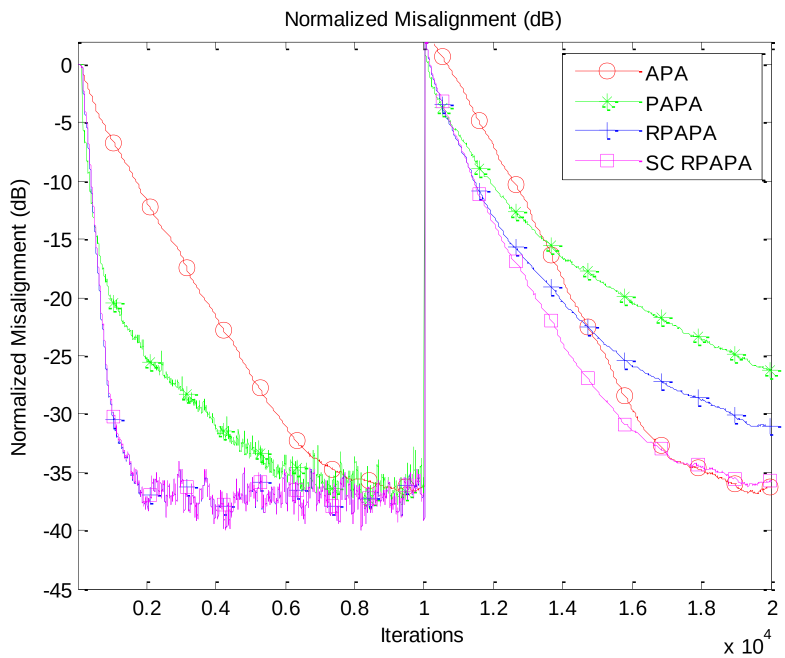


Figure 6. Comparison of SC-RPAPA with APA, PAPA, and RPAPA for WGN input, SNR=30 dB, $M = 2$, $\mu = 0.2$.

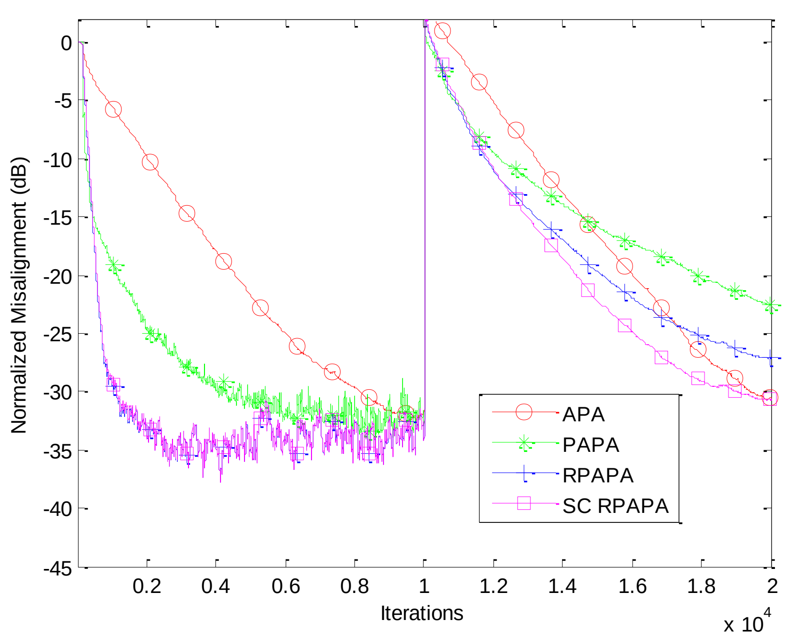


Figure 7. Comparison of SC-RPAPA with APA, PAPA, and RPAPA for colored input, SNR=30 dB, $M = 2$, $\mu = 0.2$.

5.3 THE PERFORMANCE OF THE PROPOSED SC-RPAPA WITH MRAP

An efficient implementation of the SC-RPAPA algorithm was proposed through combining the memory of the filter's coefficients with RAP. The new SC-RAPA with MRAP algorithm significantly decreases computational complexity. In this subsection, the performance of the efficient implementation was compared with APA, PAPA and SC-RPAPA through simulations.

In the first simulation, the WGN input was used. As shown in Figure 8, SC-RPAPA with MRAP worked as well as SC-RPAPA for sparse system identification. However, for dispersive system, the performance of SC-RPAPA MRAP was worse than SC-RPAPA and the APA. This fact becomes more apparent for the colored input as shown in Figure 9. This was caused by the relatively low projection order ($M = 2$), and the implementation of the MRAP was slower than the direct matrix inversion. However, this drawback could be mitigated through increasing the projection order. Furthermore, the memory of the filter's coefficients will also improve the performance as the projection order increases. We verify this point through simulations with $M = 32$ for both the WGN (see Figure 10) and the colored input (see Figure 11). It could be observed that the SC-RPAPA with MRAP works better than APA, PAPA, and SC-RPAPA for sparse system identification. Meanwhile, the performance for dispersive system with colored input has been significantly improved too.

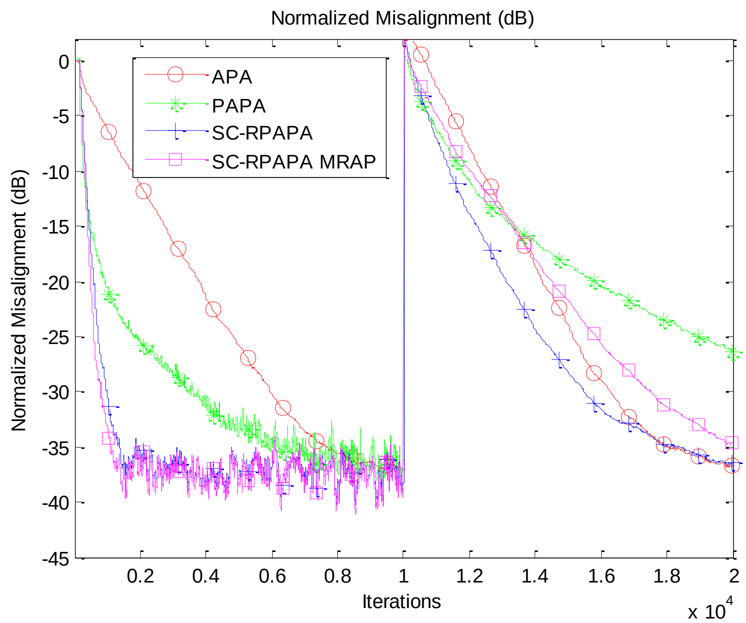


Figure 8. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for WGN input, SNR=30 dB, $M = 2$, $\mu = 0.2$.

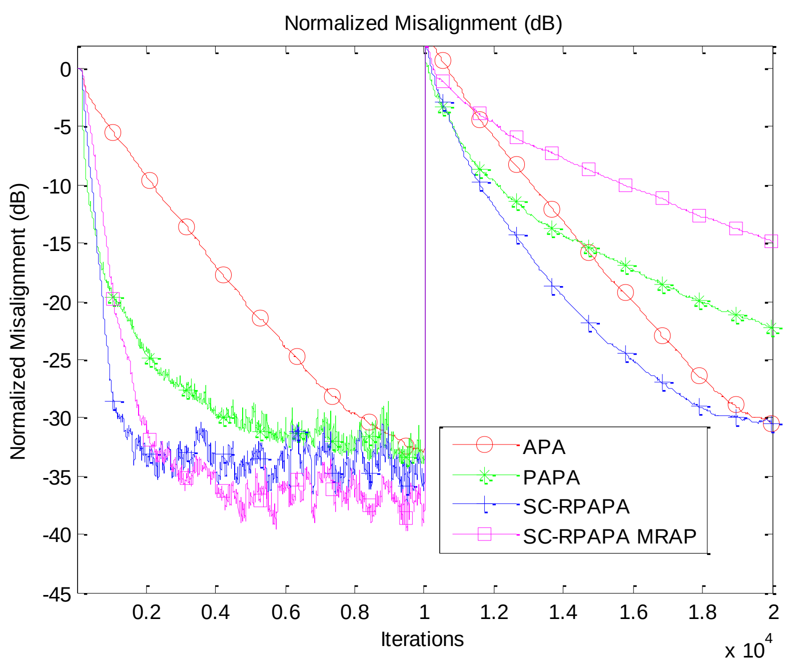


Figure 9. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for colored input, SNR=30 dB, $M = 2$, $\mu = 0.2$.

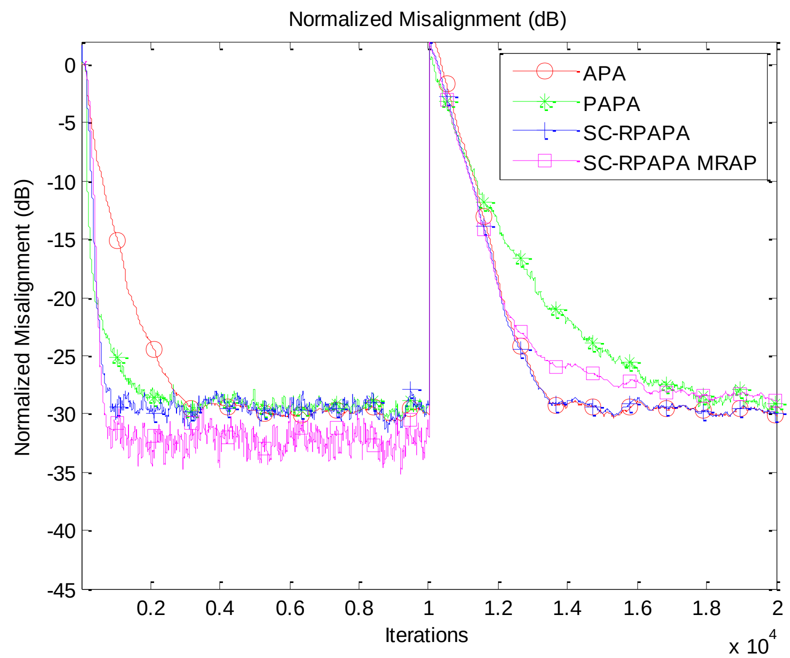


Figure 10. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for WGN input, SNR=30 dB, $M = 32$, $\mu = 0.2$.

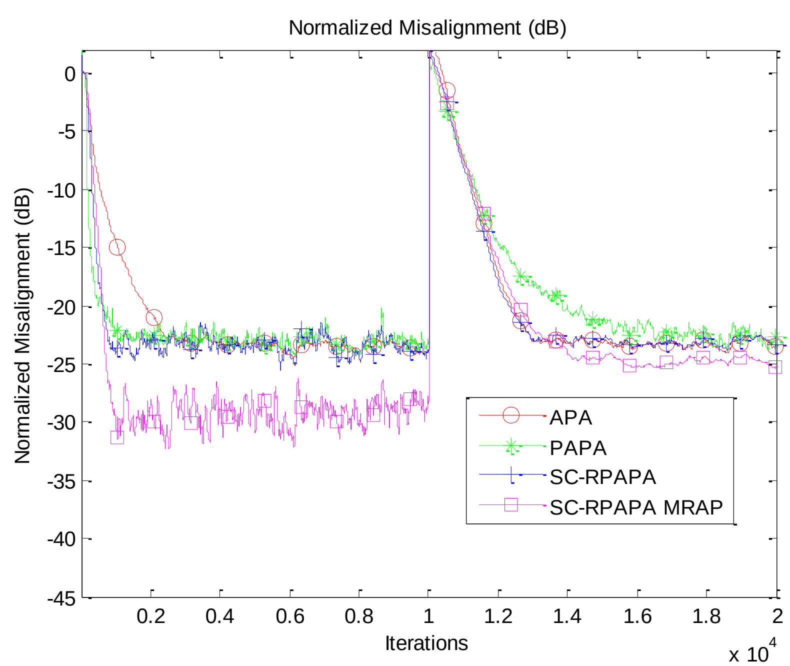


Figure 11. Comparison of SC-RPAPA MRAP with APA, PAPA and RPAPA for colored input, SNR=30 dB, $M = 32$, $\mu = 0.2$.

6. CONCLUSION

A low complexity reweighted proportionate affine projection algorithm was proposed in this paper. The sparseness of the channel was taken into account to improve the performance for dispersive systems. In order to reduce computational complexity, the direct matrix inversion of PAPA was iteratively implemented with RAP. Meanwhile, the memory of the filter's coefficients were exploited to improve the performance and further reduce the complexity for high projection orders. Simulation results demonstrate that the proposed sparseness controlled reweighted proportionate affine projection algorithm with memory and RAP outperforms traditional PAPA, with much lower computational complexity compared to mu-law and l_0 PAPA.

7. REFERENCES

[1] Hansler, E., Schmidt, G.: Acoustic Echo and Noise Control: a Practical Approach vol. 40. John Wiley & Sons, Hoboken, New Jersey (2005)

[2] Hansler, E., Schmidt, G.: Topics in Acoustic Echo and Noise Control: Selected Methods for the Cancellation of Acoustical Echoes, the Reduction of Background Noise, and Speech Processing. Springer, Berlin, Heidelberg (2006)

[3] Ozeki, K., Umeda, T.: An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. *Electronics and Communications in Japan* (Part I: Communications) 67(5), 19-27 (1984)

[4] Hansler, E., Schmidt, G.U.: Hands-free telephones - joint control of echo cancellation and postfiltering. *Signal processing* 80(11), 2295-2305 (2000)

[5] Mader, A., Puder, H., Schmidt, G.U.: Step-size control for acoustic echo cancellation filters - an overview. *Signal Processing* 80(9), 1697-1719 (2000)

[6] Duttweiler, D.L.: Proportionate normalized least-mean-squares adaptation in echo cancelers. *Speech and Audio Processing, IEEE Transactions on* 8(5), 508-518 (2000)

[7] Wagner, K., Doroslovacki, M.: Proportionate-type Normalized Least Mean Square Algorithms. John Wiley & Sons, Hoboken, New Jersey (2013)

[8] Gansler, T., Benesty, J., Gay, S.L., Sondhi, M.M.: A robust proportionate affine projection algorithm for network echo cancellation. In: *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference On*, vol. 2, pp. 793-796 (2000). IEEE

[9] Deng, H., Doroslovacki, M.: Improving convergence of the PNLMS algorithm for sparse impulse response identification. *Signal Processing Letters*, IEEE 12(3), 181-184 (2005)

[10] Deng, H., Doroslovacki, M.: Proportionate adaptive algorithms for network echo cancellation. *Signal Processing, IEEE Transactions on* 54(5), 1794-1803 (2006)

[11] Liu, L., Fukumoto, M., Saiki, S., Zhang, S.: A variable step-size proportionate affine projection algorithm for identification of sparse impulse response. *EURASIP J. Adv. Signal Process* 2009, 49-9499 (2009). doi:10.1155/2009/150914

[12] Gu, Y., Jin, J., Mei, S.: l0 norm constraint LMS algorithm for sparse system identification. *Signal Processing Letters*, IEEE 16(9), 774-777 (2009)

[13] Paleologu, C., Benesty, J., Ciochina, S.: An improved proportionate NLMS algorithm based on the l0 norm. In: *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference On*, pp. 309-312. IEEE

[14] Benesty, J., Gay, S.L.: An improved PNLMS algorithm. In: *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference On*, vol. 2, pp. 1881-1884. IEEE

[15] Hoshuyama, O., Goubran, R., Sugiyama, A.: A generalized proportionate variable step-size algorithm for fast changing acoustic environments. In: *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP'04). IEEE International Conference On*, vol. 4, p. 161 (2004). IEEE

[16] Liu, J., Grant, S.L.: Proportionate adaptive filtering for block-sparse system identification. arXiv preprint arXiv:1508.04172 (2015)

[17] Paleologu, C., Ciochina, S., Benesty, J.: An efficient proportionate affine projection algorithm for echo cancellation. *Signal Processing Letters*, IEEE 17(2), 165-168 (2010)

[18] Stanciu, C., Anghel, C., Paleologu, C., Benesty, J., Albu, F., Ciochina, S.: A proportionate affine projection algorithm using dichotomous coordinate descent iterations. In: *Signals, Circuits and Systems (ISSCS), 2011 10th International Symposium On*, pp. 1-4 (2011)

[19] Zakharov, Y., Nascimento, V.H.: Sliding-window RLS low-cost implementation of proportionate affine projection algorithms. *Audio, Speech, and Language Processing, IEEE/ACM Transactions on* 22(12), 1815-1824 (2014)

[20] Grant, S.L., Shah, P., Benesty, J.: An efficient iterative method for basis pursuit adaptive filters for sparse systems. In: *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, pp. 1-4. IEEE

[21] Yukawa, M., Yamada, I.: A unified view of adaptive variable-metric projection algorithms. *EURASIP Journal on Advances in Signal Processing* 2009, 34 (2009)

[22] Benesty, J., Paleologu, C., Ciochina, S.: Proportionate adaptive filters from a basis pursuit perspective. *Signal Processing Letters*, IEEE 17(12), 985-988 (2010)

[23] Paleologu, C., Benesty, J.: Proportionate affine projection algorithms from a basis pursuit perspective. In: *Circuits and Systems (ISCAS), 2012 IEEE International Symposium On*, pp. 2757-2760 (2012). IEEE

[24] Liu, J., Grant, S.L.: A generalized proportionate adaptive algorithm based on convex optimization. In: *Signal and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference On*, pp. 748-752. IEEE

[25] Gribonval, R., Nielsen, M.: Highly sparse representations from dictionaries are unique and independent of the sparseness measure. *Applied and Computational Harmonic Analysis* 22(3), 335-355 (2007)

[26] Gay, S.L.: An efficient, fast converging adaptive filter for network echo cancellation. In: *Signals, Systems, Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference On*, vol. 1, pp. 394-398. IEEE

[27] Kaczmarz, S.: Angenaherte auosung von systemen linearer gleichungen. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres* 35, 355-357 (1937)

[28] Benesty, J., Gansler, T.: On data-reuse adaptive algorithms. In: *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)* (2003)

[29] Gay, S.L.: Fast projection algorithms with application to voice echo cancellation. PhD thesis, Ph. D. thesis, The State University of New Jersey, 1994. 118 (1994)

[30] Grant, S.L., Shah, P., Benesty, J.: An efficient iterative method for basis pursuit adaptive filters for sparse systems. In: *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2012 Asia-Pacific, pp. 1-4 (2012). IEEE

[31] Shah, P., Grant, S.L., Benesty, J.: On an iterative method for basis pursuit with application to echo cancellation with sparse impulse responses. In: *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference On*, pp. 177-180. IEEE

[32] Loganathan, P., Khong, A.W., Naylor, P.: A class of sparseness-controlled algorithms for echo cancellation. *Audio, Speech, and Language Processing, IEEE Transactions on* 17(8), 1591-1601 (2009)

V. A NEW VARIABLE STEP-SIZE ZERO-POINT ATTRACTING PROJECTION ALGORITHM

Jianming Liu and Steven L. Grant

Abstract

This paper proposes a new variable step-size (VSS) scheme for the recently introduced zero-point attracting projection (ZAP) algorithm. The proposed variable step-size ZAPs are based on the gradient of the estimated filter coefficients' sparseness that is approximated by the difference between the sparseness measure of current filter coefficients and an averaged sparseness measure. Simulation results demonstrate that the proposed approach provides both faster convergence rate and better tracking ability than previous ones.

1. INTRODUCTION

In many practical applications, such as the network echo cancellation, the impulse response is usually sparse, which means only a small percentage of coefficients are active and most of the others are zero or close to zero [1]. Classical normalized least-mean-square (NLMS) suffers from slow convergence rate and many adaptive algorithms have been proposed to exploit the sparse nature of the system to improve performance. These include the proportionate family, in which the most popular proportionate adaptive algorithms are proportionate NLMS (PNLMS) [2], improved proportionate NLMS (IPNLMS) [3] and mu-law proportionate NLMS (MPNLMS) [4], etc.

Recently, a new LMS algorithm with l_0 norm constraint was proposed to accelerate sparse system identification [5]. It applied the constraint to the standard LMS cost function and when the solution is sparse, the gradient descent recursion will accelerate the convergence of near-zero coefficients of the sparse system. Another similar approach was proposed in [6], but it is based on l_1 norm penalty. The above scheme was referred as *zero-point attraction projection* (ZAP) in [7] and their performance analysis have been report in [8]-[10]. Analysis showed that the step-size of the ZAP term denotes the importance or the intensity of attraction. A large step-size for ZAP results in a faster convergence, but the steady-state misalignment also increases with a large step-size.

So, the step-size of ZAP is also a trade-off between convergence rate and steady-state misalignment, which is similar to the step-size trade-off of LMS. However, the variable step-size (VSS) ZAP algorithms have not been exploited too much and most of the previous algorithms are based on theoretical results which could not be calculated in practice [9]-[10]. As far as we know, the only variable step-size scheme for ZAP was proposed by You, etc. in [11], in which it was initialized to be a large value and reduced by a factor when the algorithm is convergent. However, this heuristic strategy cannot track the change in the system response due to the very small steady-state step-size.

This paper is organized as follows. Section 2 reviews the recently proposed ZAP and VSS algorithm for ZAP, and in Section 3 we present the proposed VSS ZAP algorithm. The simulation results and comparison to the previous algorithms are presented in Section 4. Finally conclusions are drawn in Section 5.

2. REVIEW OF VSS ZAP

In the scenario of echo cancellation, the far-end signal $\mathbf{x}(n)$ is filtered through the room impulse response $\mathbf{h}(n)$ to get the echo signal $y(n)$.

$$y(n) = \mathbf{x}(n) * \mathbf{h}(n) = \mathbf{x}_n^T \mathbf{h}_n, \tag{1}$$

where

$$\mathbf{x}_n = [x(n), x(n-1), \cdots, x(n-L+1)]^T, \mathbf{h}_n = [h_0, h_1, \cdots, h_{L-1}]^T,$$

and L is the length of echo path. This echo signal is added to the near-end signal $v(n)$ (including both speech and back ground noise, etc.) to get the microphone signal $d(n)$,

$$\begin{aligned} d(n) &= \mathbf{x}(n) * \mathbf{h}(n) + v(n) \\ &= \mathbf{x}_n^T \mathbf{h}_n + v(n). \end{aligned} \tag{2}$$

We define the estimation error of the adaptive filter output with respect to the desired signal as

$$e(n) = d(n) - \mathbf{x}_n^T \mathbf{w}_n. \tag{3}$$

This error, $e(n)$ is used to adapt the adaptive filter $\mathbf{w}(n)$. The LMS algorithm updates the filter coefficients as below [1]:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \mathbf{x}_n e(n), \tag{4}$$

in which μ is the step-size of adaption. The LMS algorithm with l_0 norm constraint added a zero attractor and update is as below [5]:

$$\begin{aligned} \mathbf{w}(n) = \mathbf{w}(n-1) + \mu \mathbf{x}_n e(n) \\ - \kappa\beta \operatorname{sgn}(\mathbf{w}(n-1)) \otimes e^{-\beta|\mathbf{w}(n-1)|}, \end{aligned} \tag{5}$$

where κ is the step-size of zero attractor, β is a constant, and $\otimes$ means component-wise multiplication. $\operatorname{sgn}(\cdot)$ is a component-wise sign function defined as

$$\operatorname{sgn}(x) = \begin{cases} \dfrac{x}{|x|}, & x \neq 0; \\ 0, & elsewhere. \end{cases} \tag{6}$$

The LMS algorithm with l_1 norm constraint was proposed in [6], and its update equation is

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu \mathbf{x}_n e(n) - \kappa \operatorname{sgn}(\mathbf{w}(n-1)). \tag{7}$$

The variable step-size used in [11] is rather direct: κ is initialized to be a large value, and reduced by a factor η when the algorithm is convergent. This reduction is conducted until is sufficiently small, i.e. $\kappa < \kappa_{\min}$, which means that the error reaches a low level. However, as mentioned in the introduction, this heuristic strategy will not react to a change in the system response since it will get stuck due to the very small steady-state step-size. Therefore, in order to solve this issue, we will propose a variable step-size ZAP algorithm in next section which could both converge fast and track the change efficiently.

3. PROPOSED VSS ZAP

Our proposed new variable step-size ZAP algorithm is based on the measurement of the sparseness gradient approximated by the difference between the sparseness measure of current filter coefficients and an averaged sparseness measurement. Therefore, the proposed VSS ZAP can track the change of system quickly and demonstrate a good balance between fast convergence rate and lower stable state misalignment.

For the measurement of sparsity, we could use a class of sparsity-inducing penalties. The penalty is defined as

$$J(\mathbf{w}(n)) = \sum_{i=1}^{L} G(w_i(n)), \tag{8}$$

where $G(\cdot)$ belongs to a class of sparseness measures [12]. Some commonly used sparseness measures are introduced in Table 1, where χ_P denotes the indicator function:

$$\chi_P = \begin{cases} 1 & P\, is\, true \\ 0 & P\, is\ false \end{cases} \tag{9}$$

They are mainly from [12], but they are still included in this paper for completeness. Besides to the sparseness measures listed in Table. 1, another popular measurement of channel sparsity was proposed in [13] as below. For a channel $\mathbf{h}(n)$, its sparsity can be defined as

$$\varepsilon(\mathbf{h}(n)) = \frac{L}{L-\sqrt{L}}\left(1 - \frac{\|\mathbf{h}(n)\|_1}{\sqrt{L}\,\|\mathbf{h}(n)\|_2}\right), \tag{10}$$

where $L > 1$ is the length of the channel $\mathbf{h}(n)$, and $\|\mathbf{h}(n)\|_1$ and $\|\mathbf{h}(n)\|_2$ are the l_1 norm and l_2 norm of $\mathbf{h}(n)$.

The value of $\varepsilon(\mathbf{h}(n))$ is between 0 and 1. For a sparse channel the value of sparsity is close to 1 and for a dispersive channel, this value is close to 0. Therefore, this property could be used to remove the ZAP term when the channel is dispersive, which is preferable. Instead of calculating the sparseness of the real channel, the sparsity of the current adaptive

Table 1. Sparseness measures in [12]

No.	$G(t)$	Param. Require.
1.	$\lvert t \rvert$	------
2.	$\dfrac{\lvert t \rvert}{\left(\lvert t \rvert+\sigma\right)^{1-p}}$	$0 \le p < 1$
3.	$1-e^{-\sigma \lvert t \rvert}$	$\sigma > 0$
4.	$\ln\left(1+\sigma \lvert t \rvert\right)$	$\sigma > 0$
5.	$\operatorname{atan}\left(\sigma \lvert t \rvert\right)$	$\sigma > 0$
6.	$\left(2\sigma \lvert t \rvert-\sigma^2 t^2\right)\chi_{\lvert t \rvert \le \frac{1}{\sigma}}+\chi_{\lvert t \rvert > \frac{1}{\sigma}}$	$\sigma > 0$

filter $\mathbf{w}(n)$ is estimated as [13],

$$\varepsilon\left(\mathbf{w}(n)\right)=\frac{L}{L-\sqrt{L}}\left(1-\frac{\left\|\mathbf{w}(n)\right\|_1}{\sqrt{L}\left\|\mathbf{w}(n)\right\|_2}\right). \tag{11}$$

The gradient of sparseness measure could be approximated by the difference between the sparseness measure of current filter coefficients and an averaged sparseness measurement. The averaged sparseness measure could be estimated adaptively with a forgetting factor as below:

$$\phi(n)=(1-\lambda)\phi(n-1)+\lambda J\left(\mathbf{w}(n)\right),\ 0<\lambda<1. \tag{12}$$

The difference between the sparseness measure of current filter coefficients and the averaged sparseness measurement is calculated by:

$$\delta(n)=J\left(\mathbf{w}(n)\right)-\phi(n-1). \tag{13}$$

Similar to [14], in order to obtain a good and stable estimate of the gradient, a long-term average using infinite impulse response filters is used to calculate the proposed variable step-size as below:

$$\kappa(n)=(1-\alpha)\kappa(n-1)+\alpha\gamma\left|\delta(n)\right|,\ 0<\alpha<1 \tag{14}$$

in which α is a smoothing factor and γ is a correction factor.

4. SIMULATION RESULTS

In this section, we do the results of computer simulations in the scenario of echo cancellation. In order to evaluate the performance of our proposed VSS ZAP in both sparse and dispersive impulse response, we use a sparse impulse response as in Figure 1 and a dispersive random impulse response as in Figure 2. They are both with the same length, $L = 512$, and the LMS adaptive filter is the same length. The convergence state of adaptive filter is evaluated using the normalized misalignment which is defined as

$$20\log_{10}(\|\mathbf{h}-\mathbf{w}\|_2 / \|\mathbf{h}\|_2).$$

In this simulation, we compare the proposed VSS algorithm to LMS, LMS with fixed step-size ZAP and You's VSS ZAP in [11]. For the l_1 norm constraint ZAP, we will use the No. 1 sparseness measure in Table 1 for simple, and in order to save computation efforts, for the l_0 norm constraint ZAP, we will use the same No. 3 sparseness measure as in Table 1. Meanwhile, to evaluate the performance under dispersive system, we also use the measurement of sparsity as in (11), and compare it to the above algorithms.

The input is white Gaussian noise signal and independent white Gaussian noise is added to the system background with a signal-to-noise ratio, SNR = 30 dB. The parameters of VSS ZAPs are chosen to allow all the VSS ZAPs to have similar final steady-state misalignment (about -25 dB) as standard LMS.

In order to compare the tracking, we simulate the echo path change at sample 5000 by switching to another sparse impulse response. We plot the normalized misalignment and variable step-size for l_1 norm constraint ZAP as in Figure 3 and Figure 4.

Similarly, the normalized misalignment and variable step-size for l_0 norm constraint ZAP are plotted in Figure 5, and Figure 6. It should be noted that we call the sparseness measure from Table. 1 as proposed VSS 1, and the measurement of sparsity in (11) as proposed VSS 2. We could clearly observe that the proposed VSS ZAPs are superior to standard LMS, fixed step-size ZAP LMS and previous You's VSS ZAP in the terms of convergence rate, and the tracking ability.

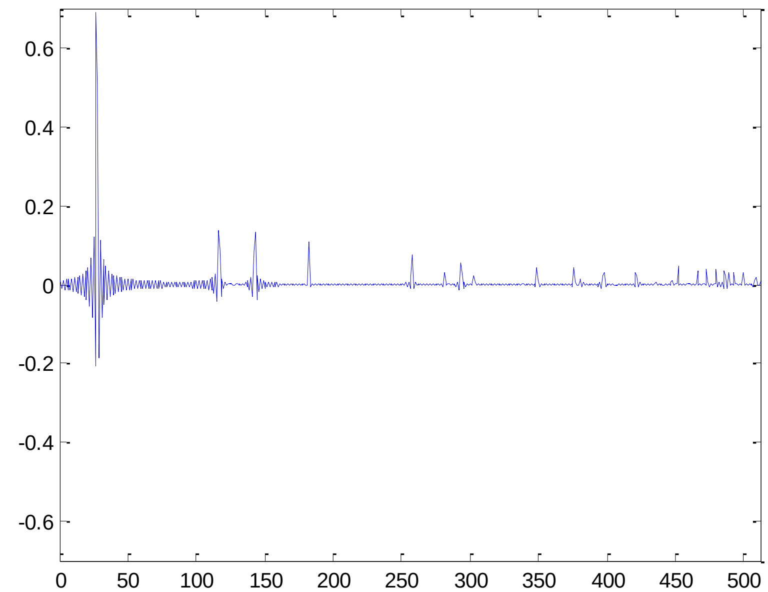


Figure 1. Sparse impulse response.

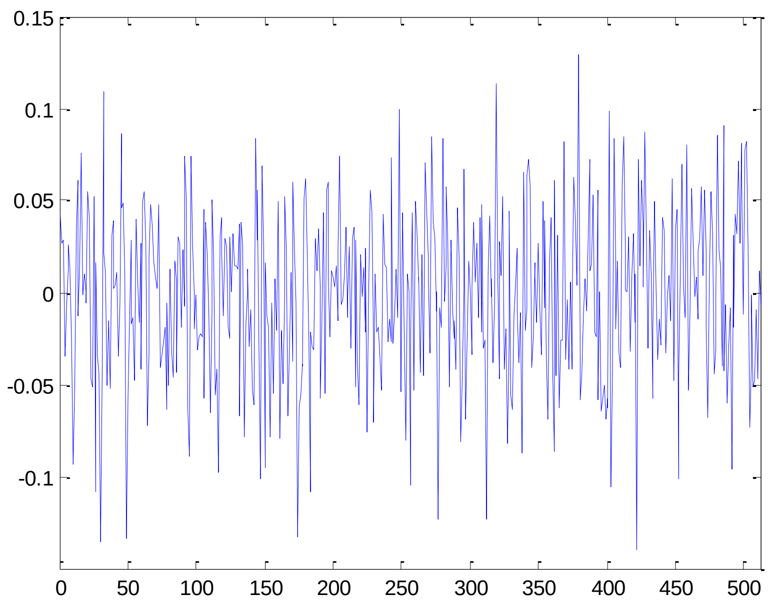


Figure 2. Dispersive random impulse response.

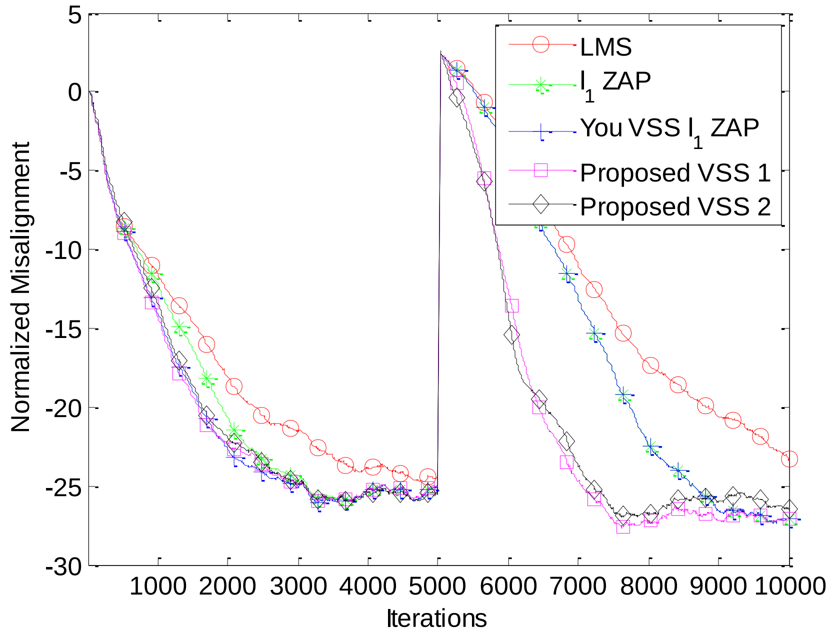


Figure 3. Comparison of normalized misalignment for l_1 norm constraint ZAP under sparse system.

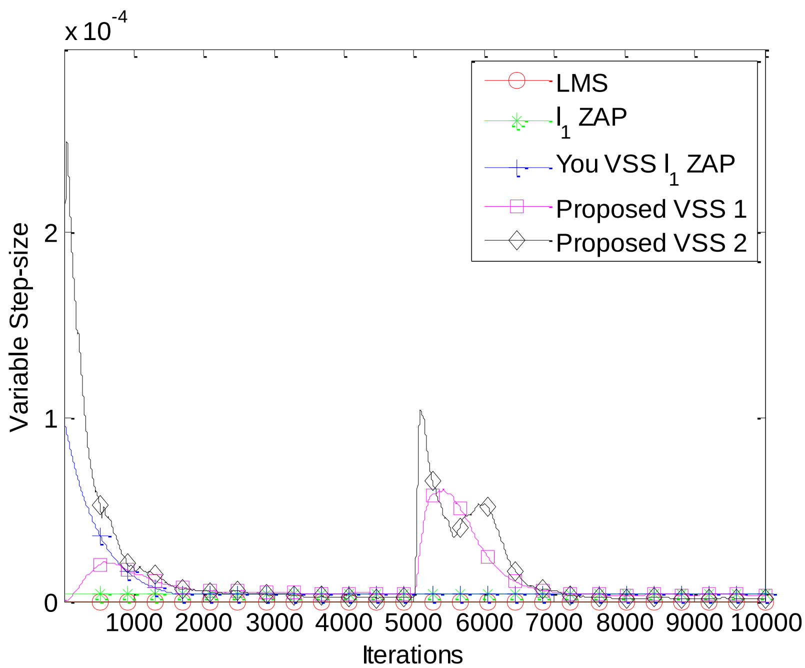


Figure 4. Comparison of variable step-size for l_1 norm constraint ZAP under sparse system.

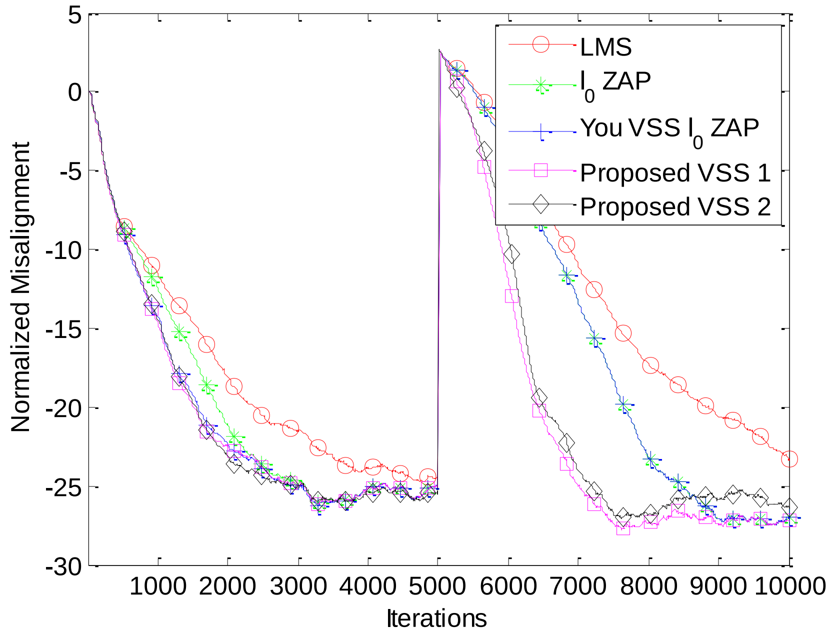


Figure 5. Comparison of normalized misalignment for l_0 norm constraint ZAP under sparse system.

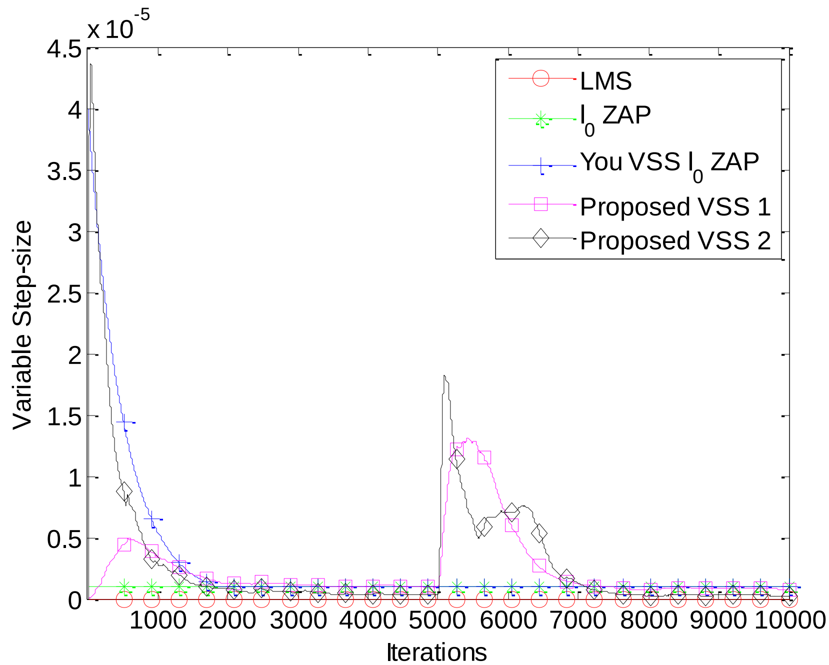


Figure 6. Comparison of variable step-size for l_0 norm constraint ZAP under sparse system.

Finally, in order to demonstrate the performance for dispersive channel, we switch the sparse echo path in Figure 1 to a dispersive random echo path as in Figure 2. The performance and VSS for l_1 norm constraint ZAP are plotted in Figure 7 and Figure 8, and l_0 norm constraint ZAP in Figure 9 and Figure 10. It is clear that the sparsity measurement in (11) could remove the impact of ZAP term under non-sparse system and performs better than the sparseness measure in Table 1. This is because the steady-state step-size of proposed VSS 1 ZAP is bigger which will cause performance degradation under non-sparse system.

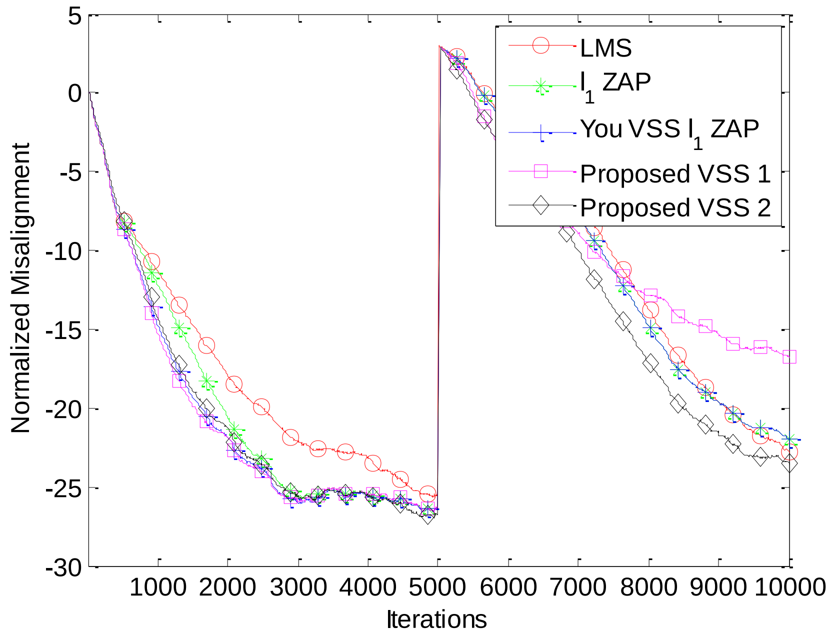


Figure 7. Comparison of normalized misalignment for l_1 norm constraint ZAP under dispersive system.

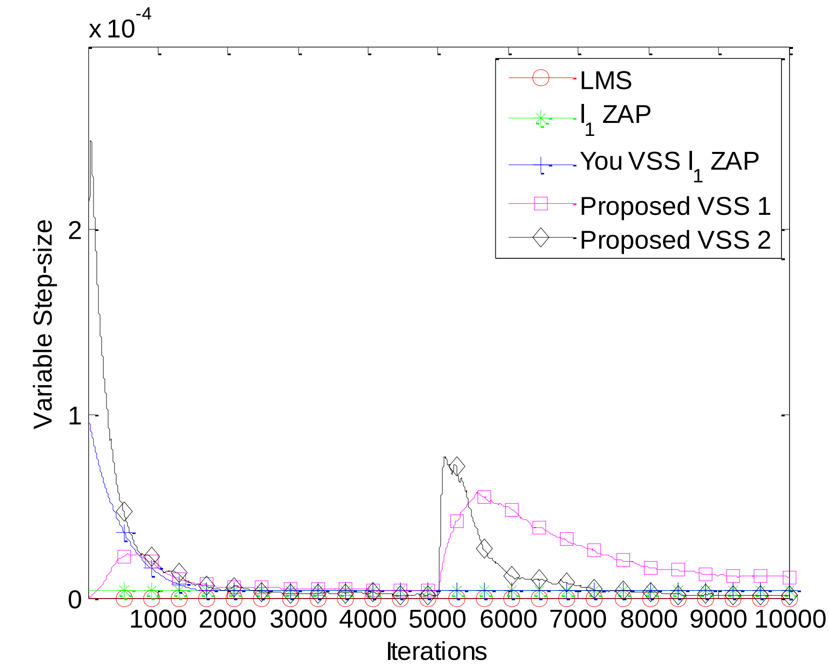


Figure 8. Comparison of variable step-size for l_1 norm constraint ZAP under dispersive system.

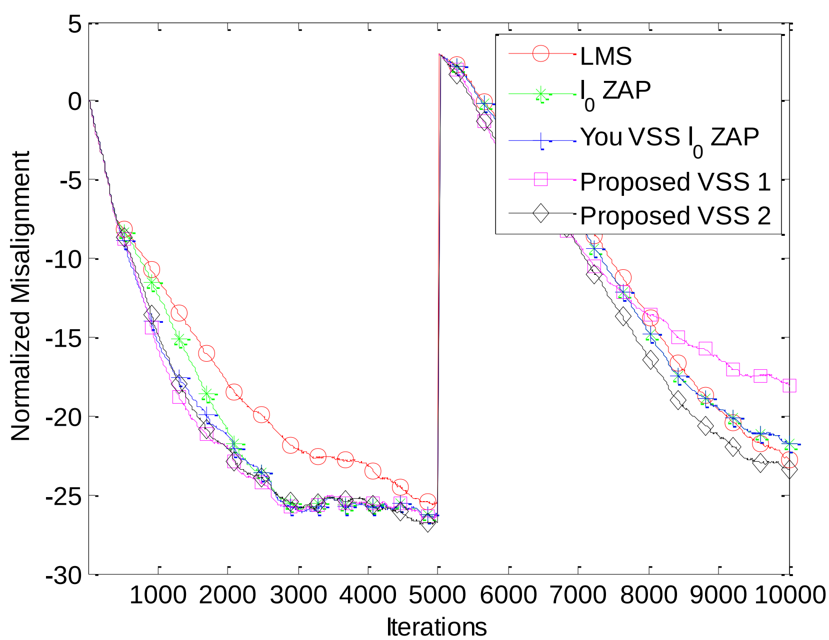


Figure 9. Comparison of normalized misalignment for l_0 norm constraint ZAP under dispersive system.

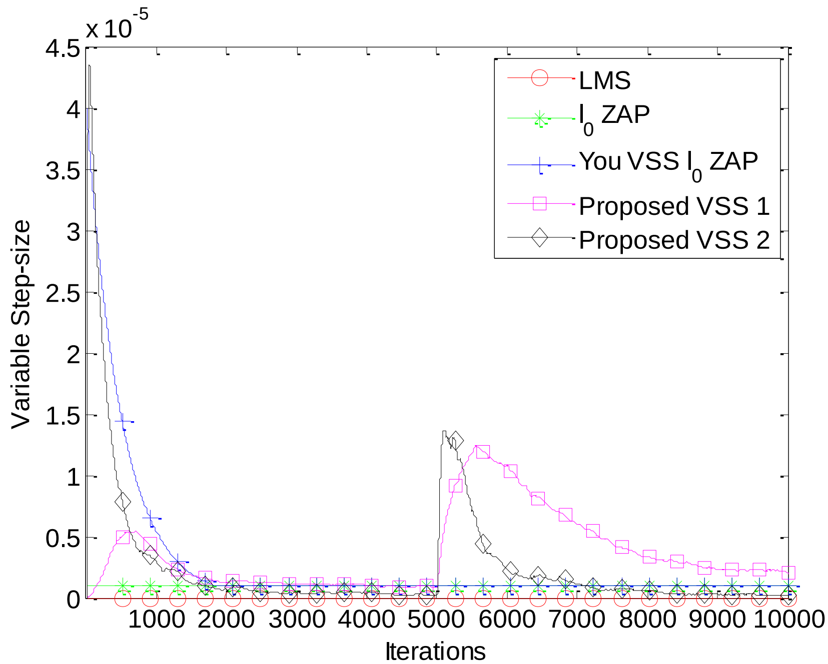


Figure 10. Comparison of variable step-size for l_0 norm constraint ZAP under dispersive system.

5. CONCLUSION

A new variable step-size scheme for the zero-point attraction projection algorithm was proposed in this paper, which is based on the estimation of sparseness gradient. Simulation results demonstrate that, for sparse system identification, the proposed VSS ZAP could provide both faster convergence rate and better tracking ability than previous VSS algorithms. Meanwhile, it could remove the impact of ZAP term for dispersive impulse response, which is preferable.

6. REFERENCES

[1] Benesty, Jacob, Tomas Gänsler, Denis R. Morgan, M. Mohan Sondhi, and Steven L. Gay. Advances in network and acoustic echo cancellation. Springer, 2001.

[2] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaption in echo cancellers," *IEEE Trans. on Speech Audio Process.*, vol.8, no. 5, pp.508-518, 2000.

[3] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," *IEEE ICASSP*, pp. 1881-1884, 2002.

[4] H. Deng and M. Doroslovacki, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *IEEE Signal Processing Lett.*, vol. 12, no. 3, pp. 181– 184, 2005.

[5] Yuantao Gu, Jian Jin, and Shunliang Mei, "l_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Proces.s Lett.*, vol 16, no. 9, pp. 774-777, Sep. 2009.

[6] Yilun Chen, Yuantao Gu and Alfred O. Hero III, "Sparse LMS for system identification," *IEEE ICASSP*, pp. 3125-3128, 2009.

[7] J. Jin, Y. Gu, and S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filter framework," *IEEE Journal of Selected Topics in Signal Process.*, vol. 4, no. 2, pp. 409-420, Apr. 2010.

[8] K. Shi and P. Shi, "Convergence analysis of sparse LMS algorithms with l1-norm penalty based on white input signal," *Signal Process.*, 2010, 90, (12), pp. 3289-3293.

[9] X, Wang, Y. Gu, and L. Chen, "Proof of convergence and performance analysis for sparse recovery via zero-point attracting projection," *IEEE Trans. Signal Process.*, vol.60, no. 8, pp.4081-4093, Aug. 2012.

[10] G, Su, J. Jin, Y. Gu, and J. Wang, "Performance analysis of l0 norm constraint least mean square algorithm," *IEEE Trans. Signal Process.*, vol.60, no. 5, pp.2223-2235, May. 2012.

[11] Y. You, L. Chen, Y. Gu, W. Feng and H. Dai, "Retrieval of sparse solutions of multiple-measurement vectors via zero-point attracting projection," *Signal Process.*, 2012, 92, (12), pp. 3075-3079.

[12] Laming Chen, and Yuantao Gu, “From least squares to sparse: A non-convex approach with guarantee,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2013. ICASSP-2013. , pp. 5875-5879.

[13] Ligang Liu, Masahiro Fukumoto, and Sachio Saiki, “An Improved Mu-law Proportionate NLMS Algorithm,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2008. ICASSP, pp. 3797-3800, 2008.

[14] Hoshuyama, O., Goubran, R. A., Sugiyama, A., “A generalized proportionate variable step-size algorithm for fast changing acoustic environments,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004. ICASSP, pp. 161-164, 2004.

VI. AN IMPROVED VARIABLE STEP-SIZE ZERO-POINT ATTRACTING PROJECTION ALGORITHM

Jianming Liu and Steven L. Grant

Abstract

This paper proposes an improved variable step-size (VSS) scheme for zero-point attracting projection (ZAP) algorithm. The proposed VSS is proportional to the sparseness difference between filter coefficients and the true impulse response. Meanwhile, it works for both sparse and non-sparse system identification, and simulation results demonstrate that the proposed algorithm could provide both faster convergence rate and better tracking ability than previous ones.

1. INTRODUCTION

In the sparse system identification problem, such as the network echo cancellation, only a small percentage of coefficients are active and most of the others are zero or close to zero. Considering that the classical least-mean-square (LMS) algorithm is slow for sparse system identification [1], the family of proportionate algorithms has been proposed to exploit the sparse nature of the system to improve performance [2]-[4]. Besides to that, a new kind of method, zero-point attracting projection (ZAP), has been recently proposed to solve sparse system identification problem. The zero-attracting LMS (ZA-LMS) algorithm uses an l_1 norm penalty in the standard LMS cost function [5] and l_0 norm LMS was proposed in [6] too. When the solution is sparse, the gradient descent recursion will accelerate the convergence of near-zero coefficients of the sparse system.

The above scheme was referred as *zero-point attraction projection* (ZAP) in [7]. The performance analysis of ZA-LMS has been report in [8]-[10], and analysis showed that the step-size of the ZAP term denotes the importance or the intensity of attraction. A large step-size for ZAP results in a faster convergence, but the steady-state misalignment also increases. So, the step-size of ZAP is also a trade-off between convergence rate and steady-state misalignment, which is similar to the step-size trade-off of LMS.

There are some theoretical results about variable step-size ZAP but they could not be calculated in practice [9]-[11]. One practical variable step-size ZAP was proposed by You, et al. in [12], and You's VSS ZAP was simply initialized to be a large value and reduced by a factor when the algorithm is convergent. However, this heuristic strategy cannot track the change in the system response due to the very small steady-state step-size.

Another better VSS-ZAP was proposed in [13], in which a variable step-size based on the gradient of estimated filter coefficients' sparseness was proposed and the gradient is approximated by the difference between the sparseness measure of current filter coefficients and an averaged sparseness measure. This variable step-size ZAP works in the way of being an indicator whether the current filter's sparseness has reached the steady-state instead of measuring the real sparseness difference between the filter and true system response. Meanwhile, in this paper, a new variable step-size ZAP is proposed by defining the sparseness distance, then the proposed VSS is determined systematically based on sparseness difference between filter coefficients and true impulse response.

This paper is organized as follows. Section 2 reviews the recently VSS algorithms for ZAP, and in Section 3 we present the proposed VSS ZA-LMS algorithm. The simulation results and comparison to the previous VSS algorithms are presented in Section 4. Finally conclusions are drawn in Section 5.

2. REVIEW OF VSS ZAP

In this section, we will review the ZAP algorithm and the variable step-size ZAP algorithms in previous literature.

2.1 INTRODUCTION TO ZAP

Consider a linear system with its input and output related by

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n), \tag{1}$$

where $\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-L+1)]^T$ is the input vector, $\mathbf{h} = [h_0, h_1, \cdots, h_{L-1}]^T$ is unknown system with length L, and $v(n)$ is the additive noise which is independent with $\mathbf{x}(n)$. The estimation error of the adaptive filter output with respect to the desired signal is defined as

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n-1). \tag{2}$$

This error, $e(n)$ is used to adapt the adaptive filter $\mathbf{w}(n)$. The ZA-LMS algorithm with l_1 norm constraint was proposed in [6], and its update equation is

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu\mathbf{x}(n)e(n) - \kappa\,\mathrm{sgn}(\mathbf{w}(n-1)), \tag{3}$$

in which μ is the step-size of adaption, κ is the step-size of zero attractor, and $\mathrm{sgn}(\cdot)$ is a component-wise sign function defined as

$$\mathrm{sgn}(x) = \begin{cases} \dfrac{x}{|x|}, & x \neq 0; \\ 0, & elsewhere. \end{cases} \tag{4}$$

2.2 REVIEW OF VARIABLE STEP-SIZE ZAP ALGORITHMS

The variable step-size for ZAP used in [12] is rather direct: κ is initialized to be a large value, and reduced by a factor η when the algorithm is convergent. This reduction is conducted until is sufficiently small, i.e. $\kappa < \kappa_{\min}$, which means that the error reaches a low level. However, as mentioned in the introduction, this heuristic strategy will not react to a change in the system response since it will get stuck due to the very small steady-state step-size.

Therefore, in order to solve this issue, a new variable step-size ZAP algorithm was proposed in [13] by us, which is based on the measurement of the sparseness gradient approximated by the difference between the sparseness measure of current filter coefficients and an averaged sparseness measurement as below.

The averaged sparseness measure could be estimated adaptively with a forgetting factor λ:

$$\phi(n) = (1-\lambda)\phi(n-1) + \lambda J\left(\mathbf{w}(n)\right),\ 0 < \lambda < 1, \tag{5}$$

where $J\left(\mathbf{w}(n)\right)$ is a sparseness measure of the filter coefficients, and we will use the following l_1 norm sparseness measure through this paper

$$J\left(\mathbf{w}(n)\right) = \left\|\mathbf{w}(n)\right\|_1 = \sum_{i=1}^{L}\left|w_i(n)\right|. \tag{6}$$

The difference between the sparseness measure of current filter coefficients and the averaged sparseness measurement is calculated by:

$$\delta(n) = J\left(\mathbf{w}(n)\right) - \phi(n-1) \tag{7}$$

In order to obtain a good and stable estimate of the gradient, a long-term average using infinite impulse response filters is used to calculate the proposed variable step-size

$$\kappa(n) = (1-\alpha)\kappa(n-1) + \alpha\gamma\delta(n),\ 0 < \alpha < 1.$$

As mentioned in the introduction, this variable step-size ZAP indicates whether the current filter's sparseness has reached the steady-state instead measuring the sparseness distance between the filter and real system. Therefore, we will propose a variable step-size algorithm for ZA-LMS which is derived based on the difference between current filter coefficients' sparseness and the real sparseness in next section.

3. PROPOSED VSS ZA-LMS

In this section, we will propose the variable step-size ZAP, and further improve its performance for non-sparse system identification.

3.1 THE PROPOSED SCHEME OF VARIABLE STEP-SIZE ZAP

Our proposed new variable step-size ZAP algorithm is based on the idea that the step-size should be proportional to the sparseness distance which is defined as the difference between the sparseness measure of current filter coefficients and real sparseness of the system. Based on l_1 norm, we define the following averaged sparseness distance

$$\delta(n) = \frac{1}{L}\left|\left\|\mathbf{w}(n)\right\|_1 - \left\|\mathbf{h}(n)\right\|_1\right| = \frac{1}{L}\left|\sum_{i=1}^{L}\left|w_i(n)\right| - \sum_{i=1}^{L}\left|h_i(n)\right|\right|. \tag{8}$$

Then we rewrite (8) as

$$\delta(n) = \frac{1}{L}\left|\mathbf{h}^T(n)\operatorname{sgn}\left(\mathbf{h}(n)\right) - \mathbf{w}^T(n)\operatorname{sgn}\left(\mathbf{w}(n)\right)\right|. \tag{9}$$

However, considering the real system is unknown, we argue that $\operatorname{sgn}\left(\mathbf{h}(n)\right)$ could be approximated by $\operatorname{sgn}\left(\mathbf{w}(n)\right)$. This assumption is acceptable because it holds for the coefficients with large magnitude, and for the small and unstable coefficients close to zero, considering that its magnitude is relatively small, it will not cause large error in the approximation. We will verify the performance of this assumption in the simulation section later, and using this assumption in (9), we have

$$\begin{aligned}\delta(n) &\approx \frac{1}{L}\left|\left(\mathbf{h}(n) - \mathbf{w}(n)\right)^T \operatorname{sgn}\left(\mathbf{w}(n)\right)\right| \\ &= \frac{1}{L}\left|\Delta\mathbf{h}^T(n)\operatorname{sgn}\left(\mathbf{w}(n)\right)\right|.\end{aligned} \tag{10}$$

The system mismatch is defined as $\Delta\mathbf{h}(n) = \mathbf{h}(n) - \mathbf{w}(n)$. Using the similar approximation in [14], we have

$$\left|\Delta\mathbf{h}^{T}(n)\operatorname{sgn}(\mathbf{w}(n))\right| \approx L\left|\frac{\Delta\mathbf{h}^{T}(n)\mathbf{x}(n)\mathbf{x}^{T}(n)\operatorname{sgn}(\mathbf{w}(n))}{\mathbf{x}^{T}(n)\mathbf{x}(n)}\right|. \tag{11}$$

It should be noted that we use the following assumptions in [14]

$$\mathbf{R}_{xx}(n) = \mathbf{x}(n)\mathbf{x}^{T}(n) \approx \sigma_x^2\mathbf{I}, \text{ and } \mathbf{x}^{T}(n)\mathbf{x}(n) \approx L\sigma_x^2. \tag{12}$$

Furthermore, the residual error is defined as

$$\varepsilon(n) = \Delta\mathbf{h}^{T}(n)\mathbf{x}(n). \tag{13}$$

Substituting (11) and (13) into (10), we could rewrite (10) as

$$\delta(n) \approx \left|\frac{\varepsilon(n)\mathbf{x}^{T}(n)\operatorname{sgn}(\mathbf{w}(n))}{\mathbf{x}^{T}(n)\mathbf{x}(n)}\right|. \tag{14}$$

However, the residual error in (14) is still unknown, but similar to [13], to avoid over-shoot, a long-term time average should be used to calculate the proposed variable step-size as below

$$\kappa(n) = (1-\alpha)\kappa(n-1) + \alpha\gamma\delta(n),\ 0 < \alpha < 1. \tag{15}$$

in which α is a smoothing factor and γ is a correction factor. Meanwhile, considering the additive noise is independent with input, the cross-correlation between the input and residual error is the same as the cross-correlation between input and error. Therefore, we could replace the residual error in (14) with the error signal, which gives us

$$\delta(n) \approx \left|\frac{e(n)\mathbf{x}^{T}(n)\operatorname{sgn}(\mathbf{w}(n))}{\mathbf{x}^{T}(n)\mathbf{x}(n)}\right|. \tag{16}$$

3.2 IMPROVED VARIABLE STEP-SIZE ZAP FOR BOTH SPARSE AND NON-SPARSE SYSTEM

Besides to the l_1 norm sparseness measures defined in (6), another popular measurement of channel sparsity was used in [13], and for a channel $\mathbf{h}(n)$, its sparsity $\xi(\mathbf{h}(n))$ can be defined as

$$\xi(\mathbf{h}(n)) = \frac{L}{L-\sqrt{L}}\left(1-\frac{\|\mathbf{h}(n)\|_1}{\sqrt{L}\|\mathbf{h}(n)\|_2}\right), \tag{17}$$

where $L>1$ is the length of the channel $\mathbf{h}(n)$, and $\|\mathbf{h}(n)\|_1$ and $\|\mathbf{h}(n)\|_2$ are the l_1 norm and l_2 norm of $\mathbf{h}(n)$. The value of $\xi(\mathbf{h}(n))$ is between 0 and 1. For a sparse channel the value of sparsity is close to 1 and for a dispersive channel, this value is close to 0. In [13], this property was used to remove the ZAP term when the channel is dispersive, which is preferable.

We could also take advantage of this property and propose the following averaged sparseness distance as variable step-size for ZA-LMS

$$\begin{aligned}\delta(n) &= \frac{1}{L}\left|\xi(\mathbf{h}(n))-\xi(\mathbf{w}(n))\right| \\ &= \frac{1}{L\left(\sqrt{L}-1\right)}\left|\frac{\|\mathbf{h}(n)\|_1}{\|\mathbf{h}(n)\|_2}-\frac{\|\mathbf{w}(n)\|_1}{\|\mathbf{w}(n)\|_2}\right|.\end{aligned} \tag{18}$$

We assume the gain of the real channel and filter coefficients are the same, i.e.

$$\|\mathbf{h}(n)\|_2 \approx \|\mathbf{w}(n)\|_2 . \tag{19}$$

However, this assumption might not be accurate, especially at the initial phase of the adaption. Therefore, a reasonable minimum threshold of $\|\mathbf{w}(n)\|_2$ should be used to avoid this issue. Then we could further simplify (19) as

$$\delta(n) \approx \frac{1}{L\left(\sqrt{L}-1\right)}\frac{1}{\|\mathbf{w}(n)\|_2}\left|\|\mathbf{h}(n)\|_1-\|\mathbf{w}(n)\|_1\right|. \tag{20}$$

Considering (16), we obtain the proposed variable step-size for ZA-LMS which could work for both dispersive and sparse channel as below

$$\delta(n) \approx \frac{1}{\sqrt{L}-1}\frac{1}{\|\mathbf{w}(n)\|_2}\left|\frac{e(n)\mathbf{x}^T(n)\operatorname{sgn}(\mathbf{w}(n))}{\mathbf{x}^T(n)\mathbf{x}(n)}\right|. \tag{21}$$

4. SIMULATION RESULTS

In this section, we do the results of computer simulations in the scenario of echo cancellation. We use both sparse impulse response and a dispersive random impulse response. They are both with the same length, $L = 512$, and the LMS adaptive filter is with the same length.

The convergence state of adaptive filter is evaluated using the normalized misalignment which is defined as

$$20\log_{10}(\|\mathbf{h}-\mathbf{w}\|_2 / \|\mathbf{h}\|_2). \tag{22}$$

The input is white Gaussian noise signal and independent white Gaussian noise is added to the system background with a signal-to-noise ratio, SNR = 30 dB.

In the first simulation, we would like to verify the performance of the approximation $\operatorname{sgn}(\mathbf{h}(n)) = \operatorname{sgn}(\mathbf{w}(n))$ in (10) as in Figure 1. In order to demonstrate the tracking ability, there is an echo path change at sample 5000 by switching from one sparse impulse response to another sparse impulse response. It is observed that, even though the approximation is not very accurate in the initial phase, it could be very good for tracking the change of the echo path. This is predictable since the filter coefficients are initialized as zeros, then there will be larger difference between $\operatorname{sgn}(\mathbf{h}(n))$ and $\operatorname{sgn}(\mathbf{w}(n))$. However, this assumption is still good enough for the application scenario of proposed variable step-size ZAP, which will be verified by the following simulations.

In the second simulation, we compare the proposed VSS algorithm to LMS, fixed step-size ZA-LMS, You's VSS in [12] and Liu's VSS in [13] for sparse system identification. It should be noted that sparseness measure (17) is used in Liu's VSS, and (21) is used as the proposed variable step-size. Meanwhile, to evaluate the performance of the tracking ability, there is also an echo path change at sample 5000, and according to the simulation result in Figure 2, the parameters of the variable step-size are intentionally set to have similar steady-state misalignment for the first adaption before echo path change. It is observed that, because You's VSS cannot react to echo path change, it could only obtain similar tracking performance with original ZAP. Meanwhile, Liu's VSS and proposed VSS could track the echo path change quickly, and the proposed VSS outperforms the previous ones.

Next, in order to demonstrate the performance for dispersive channel, we switch one dispersive impulse response to another dispersive response at sample 5000, and use the same VSS algorithms and parameters as the second simulation. As shown in Figure 3, it is clear that the proposed VSS ZAP could also obtain much better tracking performance under non-sparse system than previous ones and avoid the possible performance degradation.

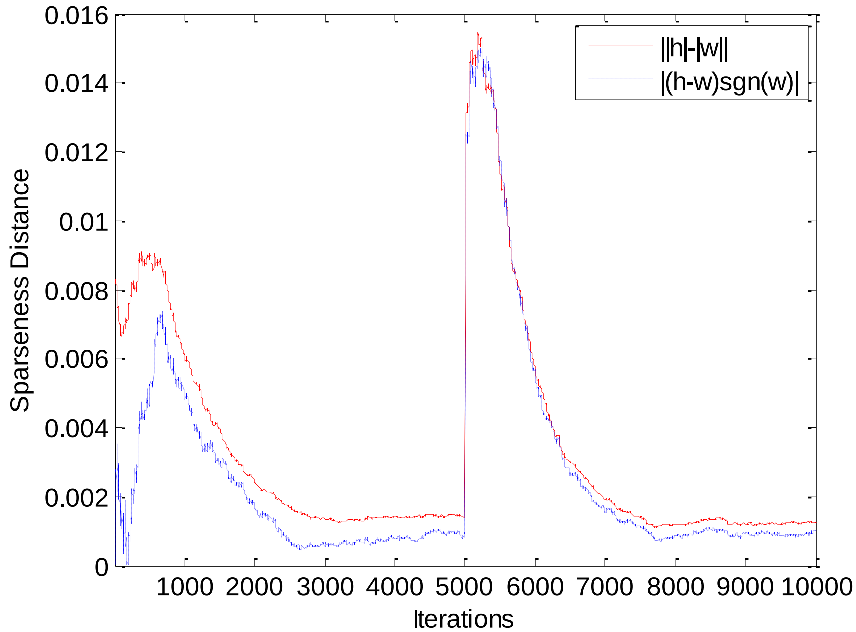


Figure 1. Performance demonstration of approximation $\mathrm{sgn}(\boldsymbol{h}(n)) = \mathrm{sgn}(\boldsymbol{w}(n))$ in (10).

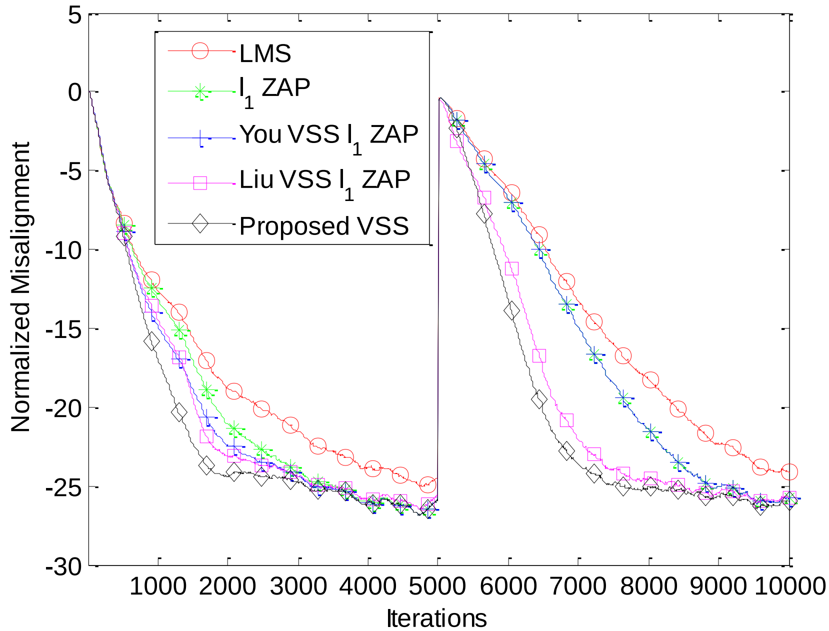


Figure 2. Comparison of normalized misalignment for sparse system identification.

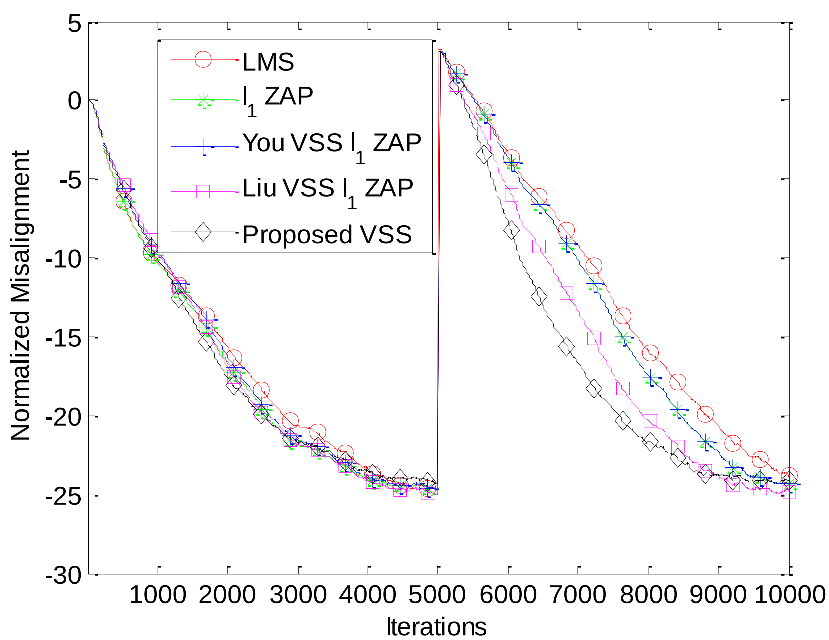


Figure 3. Comparison of normalized misalignment for dispersive system identification.

5. CONCLUSION

An improved variable step-size zero-point attraction projection algorithm was proposed based on the estimation of l_1 sparseness distance, which could work for both sparse and non-sparse system identification. Simulation results verify that the proposed VSS ZAP could provide better tracking ability than previous VSS ZAP algorithms for both sparse and non-sparse system identification.

6. REFERENCES

[1] Benesty, Jacob, Tomas Gänsler, Denis R. Morgan, M. Mohan Sondhi, and Steven L. Gay. *Advances in network and acoustic echo cancellation*. Springer, 2001.

[2] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaption in echo cancellers," *IEEE Trans. on Speech Audio Process*., vol.8, no. 5, pp.508-518, 2000.

[3] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," *IEEE ICASSP*, pp. 1881-1884, 2002.

[4] Jianming Liu, and Steven L. Grant, "A Generalized Proportionate Adaptive Algorithm based on Convex Optimization," *2014 IEEE China Summit & International Conference on*, pp. 748-752, 2014.

[5] Yuantao Gu, Jian Jin, and Shunliang Mei, "l_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett*., vol 16, no. 9, pp. 774-777, Sep. 2009.

[6] Yilun Chen, Yuantao Gu and Alfred O. Hero III, "Sparse LMS for system identification," *IEEE ICASSP*, pp. 3125-3128, 2009.

[7] J. Jin, Y. Gu, and S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filter framework," *IEEE Journal of Selected Topics in Signal Process.*, vol. 4, no. 2, pp. 409-420, Apr. 2010.

[8] K. Shi and P. Shi, "Convergence analysis of sparse LMS algorithms with l_1-norm penalty based on white input signal," *Signal Process*., 2010, 90, (12), pp. 3289-3293.

[9] X, Wang, Y. Gu, and L. Chen, "Proof of convergence and performance analysis for sparse recovery via zero-point attracting projection," *IEEE Trans. Signal Process*., vol.60, no. 8, pp.4081-4093, Aug. 2012.

[10] G, Su, J. Jin, Y. Gu, and J. Wang, "Performance analysis of l0 norm constraint least mean square algorithm," *IEEE Trans. Signal Process*., vol.60, no. 5, pp.2223-2235, May. 2012.

[11] Markus V.S. Lima, Iker Sobron, Wallace A. Martins, Paulo S.R. Diniz, "Stability and MSE Analysis of Affine Projection Algorithms for Sparse System Identification," *IEEE ICASSP*, pp. 6449-6453, 2014.

[12] Y. You, L. Chen, Y. Gu, W. Feng and H. Dai, "Retrieval of sparse solutions of multiple-measurement vectors via zero-point attracting projection," *Signal Process*., 2012, 92, (12), pp. 3075-3079.

[13] J. Liu and S. L. Grant, "A New Variable Step-size Zero-point Attracting Projection Algorithm," in Proc. Signals, Systems and Computers, *2013 Asilomar Conference*, pp. 1524-1528, 2013.

[14] Mohammad Asif Iqbal and Steven L. Grant, "Novel variable step size NLMS algorithms for echo cancellation," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2008. pp. 241-244, 2008.

SECTION

2. CONCLUSIONS

This dissertation studied the adaptive filters for sparse system identification, and proposed a new family of proportionate adaptive filters for bock-sparse system identification, a low-complexity reweighted proportionate affine projection algorithm and two variable step-size zero-point attracting projection algorithms.

For block-sparse system identification, the block-sparse proportionate normalized least mean square (BS-PNLMS) and block-sparse improved PNLMS (BS-IPNLMS) algorithms have been firstly proposed for block-sparse system identification. With a modest increase in computational complexity, the block-sparse algorithms could achieve faster convergence speed and better tracking ability in block-sparse system identification. Meanwhile, the block-sparse proportionate idea has been applied to both the proportionate affine projection algorithm (PAPA) and proportionate affine projection sign algorithm (PAPSA), yielding block-sparse PAPA (BS-PAPA) and block-sparse memory improved PAPSA (BS-MIPAPSA). The BS-PAPA is an extension of BS-PNLMS and works better at the cost of higher computational complexity especially for colored input. Meanwhile, the BS-MIPAPSA is robust to impulsive noise.

In order to further improve the performance of PAPA algorithm and reduce the computational complexity, a novel sparseness controlled reweighted PAPA (RPAPA) algorithm with memory and row action projection (SC-RPAPA with MRAP) has been proposed in this dissertation. Compared to the previous mu-law PAPA, etc., the computational complexity of the proposed algorithm is significantly reduced due to the combination of coefficients' memory and RAP. Meanwhile, SC-RPAPA works for both sparse and dispersive system due to sparseness control.

The zero-point attracting projection (ZAP) was recently proposed for sparse system identification, and the step size of the attractor is also a trade-off between the convergence rate and steady misalignment level. Therefore, two variable step size ZAP algorithms were proposed to improve the performance of ZAP algorithms.

3. PUBLICATIONS

[1] J. Liu and S. L. Grant, "An improved proportionate multi-delay filter for block-sparse system identification," to be submitted to *EURASIP Journal on Audio, Speech, and Music Processing*.

[2] J. Liu and S. L. Grant, "Proportionate affine projection algorithms for block-sparse system identification," accepted by *IEEE ICASSP* 2016.

[3] J. Liu and S. L. Grant, "Block sparse memory improved proportionate affine projection sign algorithm," published in *IET Electronics Letters*.

[4] J. Liu, S. L. Grant and J. Benesty, "A low complexity reweighted proportionate affine projection algorithm with memory and row action projection," published in *EURASIP Journal on Advances in Signal Processing*.

[5] J. Liu and S. L. Grant, "Proportionate adaptive filtering for block-sparse system identification," accepted by *IEEE/ACM Trans. on Audio, Speech, Language Processing*.

[6] H. Geng, J. Liu, P. Luo, L. Cheng, S. L. Grant, and Y. Shi. "Selective Body Biasing for Post-Silicon Tuning of Sub-Threshold Designs: An Adaptive Filtering Approach." *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on* 34, no. 5 (2015): 713-725.

[7] J. Liu and S. L. Grant, "An improved variable step-size zero-point attracting projection algorithm," *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on* , pp.604,608, 19-24 April 2015

[8] J. Liu and S. L. Grant, "A generalized proportionate adaptive algorithm based on convex optimization," in *Proc. Signals and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*, pp. 748-752, 2014.

[9] C. Zhang, P. Deng, H. Geng, J. Liu, Q. Zhu, J. Xiong and Y. Shi, "MSim: A General Cycle Accurate Simulation Platform for Memcomputing Studies," in *Proc. of the conference on Design, Automation & Test in Europe (DATE-2014),* pp 265-269, 2014. (Invited Paper).

[10] J. Liu and S. L. Grant, "A new variable step-size zero-point attracting projection algorithm," in *Proc. Signals, Systems and Computers, 2013 Asilomar Conference*, pp. 1524-1528, 2013.

[11] J. Liu and S. L. Grant, “An Improved Variable Step-size Affine Projection Sign Algorithm for Echo Cancellation,” *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pp. 1-5, 2013.

[12] H. Geng, J. Wu, J. Liu, M. Choi and Y. Shi, “Utilizing random noise in cryptography: Where is the Tofu?,” *Computer-Aided Design (ICCAD), 2012 IEEE/ACM International Conference on* , pp.163-167, 5-8 Nov. 2012.

[13] J. Liu and S. L. Grant, “A New Doubletalk and Channel Change Detection Algorithm based on Hypothesis Testing,” *European Signal Processing Conference. EUSIPCO-2012.*, pp. 2457-2461, 2012.

BIBLIOGRAPHY

[1] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in network and acoustic echo cancellation*. Springer, 2001.

[2] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaption in echo cancellers," *IEEE Trans. on Speech Audio Process*., vol.8, no. 5, pp. 508-518, 2000.

[3] T. Gansler, J. Benesty, S. L. Gay, and M. Sondhi. "A robust proportionate affine projection algorithm for network echo cancellation." In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, vol. 2, pp. II793-II796. IEEE, 2000.

[4] J. Benesty, C. Paleologu, and S. Ciochin, "Proportionate adaptive filters from a basis pursuit perspective," *Signal Processing Letters, IEEE,* vol. 17, no. 12, pp. 985-988, 2010.

[5] C. Paleologu, and J. Benesty, "Proportionate affine projection algorithms from a basis pursuit perspective." In *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pp. 2757-2760. IEEE, 2012.

[6] Y. Chen, Y. Gu and A. O. Hero III, "Sparse LMS for system identification," *IEEE ICASSP*, pp. 3125-3128, 2009.

[7] Y. Gu, J. Jin, and S. Mei, "l_0 norm constraint LMS algorithm for sparse system identification," *Signal Processing Letters, IEEE,* vol. 16, no. 9, pp. 774-777, 2009.

VITA

Jianming Liu was born in Jining, Shandong, China. He received his B.S. degree in Electronic Science and Technology from Shandong University, Jinan, China, in 2005, and M.S. degree in Electrical and Communication Engineering from Tsinghua University, Beijing, China in 2009. He joined Nokia (China) Co. Ltd, Beijing in August 2010 as an Audio Design Engineer. He started to pursue Ph.D. degree in Electrical and Computer Engineering at Missouri University of Science and Technology, Rolla, Missouri in August 2011. His research interests focus on acoustic echo cancellation, adaptive filtering, and noise reduction. In May 2016, he received his PhD in Electrical Engineering from Missouri University of Science and Technology.