Spring 2014

Sparse nonlinear optimization for signal processing and communications

Zengli Yang

Department: Electrical and Computer Engineering

SPARSE NONLINEAR OPTIMIZATION FOR SIGNAL PROCESSING AND COMMUNICATIONS

by

ZENGLI YANG

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

2014

Approved by

Yahong Rosa Zheng, Advisor
Chengshan Xiao
Reza Zoughi
Steve Grant
Serhat Hosder

PUBLICATION DISSERTATION OPTION

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

Paper I, Z. Yang, Y. R. Zheng, and S. L. Grant, "Proportionate affine projection sign algorithm for network echo cancellation," has been published in IEEE Transactions on Audio, Speech, and Language Processing, vol. 19, no. 8, pp. 2273–2284, Nov. 2011.

Paper II, Z. Yang, and Y. R. Zheng, "A comparative study of compressed sensing for 3-D synthetic aperture radar image reconstruction," has been accepted in Elsevier Digital Signal Process., pp. 1–28, Mar. 2014.

Paper III, Z. Yang, and Y. R. Zheng, "Robust adaptive channel estimation in MIMO underwater acoustic communications," has been accepted in MTS/IEEE OCEANS'14, Taipei, Taiwan, Apr. 7–10, 2014, pp. 1–6.

ABSTRACT

This dissertation proposes three classes of new sparse nonlinear optimization algorithms for network echo cancellation (NEC), 3-D synthetic aperture radar (SAR) image reconstruction, and adaptive turbo equalization in multiple-input multiple-output (MIMO) underwater acoustic (UWA) communications, respectively.

For NEC, the proposed two proportionate affine projection sign algorithms (APSAs) utilize the sparse nature of the network impulse response (NIR). Benefiting from the characteristics of l_1-norm optimization, affine projection, and proportionate matrix, the new algorithms are more robust to impulsive interferences and colored input than the conventional adaptive algorithms.

For 3-D SAR image reconstruction, the proposed two compressed sensing (CS) approaches exploit the sparse nature of the SAR holographic image. Combining CS with the range migration algorithms (RMAs), these approaches can decrease the load of data acquisition while recovering satisfactory 3-D SAR image through l_1-norm optimization.

For MIMO UWA communications, a robust iterative channel estimation based minimum mean-square-error (MMSE) turbo equalizer is proposed for large MIMO detection. The MIMO channel estimation is performed jointly with the MMSE equalizer and the maximum *a posteriori* probability (MAP) decoder. The proposed MIMO detection scheme has been tested by experimental data and proved to be robust against tough MIMO channels.

ACKNOWLEDGMENTS

First and foremost, I would like to gratefully and sincerely thank my advisor Dr. Yahong Rosa Zheng. I truly appreciate her support and guidance for my Ph.D. study and research work during the past four years. She has provided numerous valuable suggestions and discussions to make this dissertation possible, and has cultivated my ability of independent thinking. She has also given me the opportunities to work on practical wireless communication projects. These valuable experiences not only enhance my theoretical understanding but also lay a solid foundation for my future career. Her enthusiasm and attitudes towards research were contagious and motivational for me.

I would also like to thank the members of my advisory committee, Drs. Chengshan Xiao, Reza Zoughi, Steve L. Grant and Serhat Hosder, for their precious time in examining this dissertation and their constructive suggestions to my research work.

In addition, it is my pleasure to acknowledge my particular appreciation to some of my co-authors and labmates, especially Mr. Bing Han, Mr. Weimin Duan, Ms. Tiange Shao, Mr. Kajbaf Hamed, and Mr. Joseph T. Case, with whom I worked on the theoretical problems or the practical projects. I enjoyed the productive discussions and collaborations. I also cherish the time that I spent with all the other group members in the research lab.

Last, but certainly not the least, I wish to express my heartful thanks to my family for their unselfish love, everlasting support and sacrifice. Particularly, this dissertation is dedicated to my brilliant and outrageously loving and supportive wife, Hui He, our brand new, exuberant, smart and lovely baby boy, Lewis Yanche Yang, and to my always encouraging, ever faithful parents, Chuandong Yang and Lingyun Zeng.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

PAPER II

PAPER III

LIST OF TABLES

1 INTRODUCTION

1.1 BACKGROUND AND PROBLEM STATEMENT

Sparse nonlinear optimization [1] [2] is to solve the large-scale problem with sparse nature efficiently and effectively. In the last decade, sparse nonlinear optimization has made significant achievements and has been widely researched in many signal processing applications: system identification, radar imaging, magnetic resonance imaging (MRI), wireless communication, etc. This dissertation proposed three classes of sparse nonlinear optimization algorithms for network echo cancellation, 3-D synthetic aperture radar (SAR) image reconstruction, and adaptive turbo equalization in multiple-input multiple-output (MIMO) underwater acoustic (UWA) communications, respectively.

A network echo canceller (NEC) is an adaptive filter essential to modern voice communication networks. An NEC first estimates the echo path of the network, generates a replica of the network impulse response (NIR), and then subtracts the replica of the far-end echo from the near-end signal to obtain clean signals. Although NEC is a classic system identification application, modern VoIP (Voice over Internet Protocol) and telephone networks impose several challenges on conventional adaptive filters. The first is that today's large scale network often experiences long delays and causes long network impulse responses (NIRs) that require adaptive filters with 1000 or more taps; while the traditional network only requires adaptive filters with less than 100 taps. However, the response of the long-distance/local-loop interface is still typically only 6 to 12 milliseconds long, which with a typical sample rate of 8 kHz yields a flat delay of zero to nearly 900 samples followed by a damped ringing of about 50 samples. This means that the NIR is typically rather sparse, in that most of the

coefficients are close to zero. The second challenge is that the excitation signals in NECs are speech signals which are often highly colored non-Gaussian processes rather than white Gaussian noises. Many adaptive algorithms suffer reduced convergence rate in this scenario. The third challenge of NEC is due to double-talk where both near-end speech and far-end speech are active simultaneously. This often causes extremely slow convergence or even divergence in adaptive filters.

The NEC scheme is shown in Fig. 1.1, where $x(k)$ is the far-end signal, $z(k)$ and $v(k)$ are the near-end speech and background noise signal, respectively. The NIR of the true echo path is denoted by a length L coefficient vector $\mathbf{h}$.

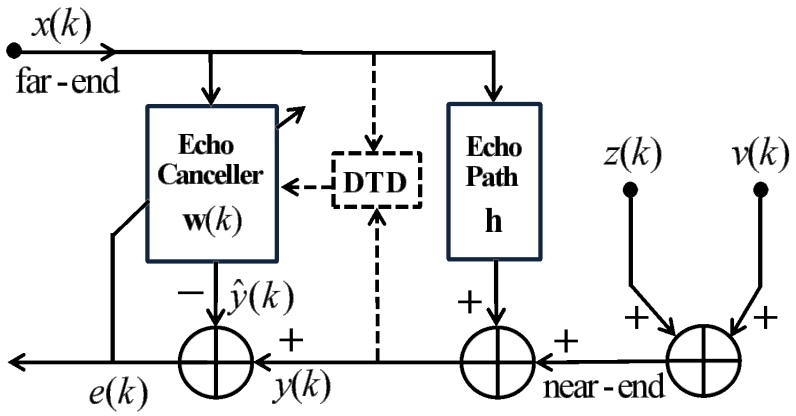


Figure 1.1. Structure of a network echo canceller (NEC).

The estimated impulse response of the NEC is denoted by $\mathbf{w}(k) = [w_0(k), w_1(k), \ldots, w_{L-1}(k)]^T$, where k is the time index and the superscript T denotes transpose. The signal $y(k)$ contains the echo, the near-end speech, and background noise. That is, $y(k) = \mathbf{x}^T(k)\mathbf{h} + z(k) + v(k)$, where $\mathbf{x}(k) = [x(k), x(k-1), \ldots, x(k-L+1)]^T$ is the far-end signal vector. Generating the replica echo $\hat{y}(k) = \mathbf{x}^T(k)\mathbf{w}(k)$, the NEC tries to minimize the difference between $y(k)$ and $\hat{y}(k)$ with an adaptive $\mathbf{w}(k)$.

Different from the l_2-norm algorithms [3], the formulation of sparse nonlinear optimization for NEC is obtained by minimizing the l_1-norm of the *a posteriori* error vector with a constraint on the filter coefficients,

$$\min_{\mathbf{w}(k+1)} \quad \|\mathbf{y}(k) - \mathbf{X}^T(k)\mathbf{w}(k+1)\|_1 \tag{1.1}$$

$$\text{subject to} \quad \|\mathbf{w}(k+1) - \mathbf{w}(k)\|_2^2 \leq \mu^2 \tag{1.2}$$

where μ^2 is a parameter to ensure the weight coefficient vector does not change too much in one iteration, $\mathbf{X}(k) = [\mathbf{x}(k), \mathbf{x}(k-1), \ldots, \mathbf{x}(k-M+1)]$, $\mathbf{y}(k) = [y(k), y(k-1), \ldots, y(k-M+1)]^T$, and M is the projection order.

Wideband 3-D SAR imaging has important applications in the area of nondestructive testing and evaluation (NDT&E). This is due to its feasibility to acquire high-resolution holographic images of specimen under test. Microwave and millimeter wave as interrogating signals can not only penetrate dielectric materials but also interact with their inner structure to render a comprehensive image for inspection. Microwave and millimeter wave NDT&E techniques have been applied to diverse applications (i.e., the detection and evaluation of corrosion under paint and composite laminates, the detection and sizing of fatigue cracks in metal surfaces, and the characterization of dielectric material). However, the speed of data acquisition for these wideband 3-D SAR imaging systems hinders their practical applications due to the slowness of mechanical scanning. For example, uniform raster scanning requires approximately one hour to scan a 120×180 mm^2 area at 2 mm spacing. In contrast, random undersampling can reduce the acquisition time when the antenna probe is placed at only a fractional number of positions on the uniform grid. Advanced sparse methods are required to reconstruct images when undersampled measurements are used.

For 3-D SAR image reconstruction, the sparse nonlinear optimization problem can be interpreted as [4]

$$\min_{\hat{\mathbf{g}}} J(\hat{\mathbf{g}}) \quad \text{subject to} \quad \|\mathbf{\Phi}\hat{\mathbf{g}} - \mathbf{r}\|_2^2 < \sigma^2, \tag{1.3}$$

where $\|\cdot\|_2$ is the l_2 norm, $\hat{\mathbf{g}} \in \mathbb{C}^{N\times 1}$ is the vectorized estimated 3-D SAR image, σ^2 is the noise variance, and $\mathbf{\Phi} \in \mathbb{C}^{M\times N}$ $(M < N)$ is the measurement matrix that reflects the acquisition of the vectorized raw measurements $\mathbf{r} \in \mathbb{C}^{M\times 1}$. For the Stolt-CS and NUFFT-CS, $\mathbf{\Phi}$ is the reverse Stolt-RMA and NUFFT-RMA, respectively. The measurement operator $\mathbf{\Phi}$ can be written as

$$\mathbf{\Phi}_{\text{Stolt}} = \mathbf{U}\mathcal{F}_{\text{2D}}^{-1}\left\{\mathbf{\Theta}^{\dagger}[\mathcal{F}_{\text{3D}}(\cdot)]\right\}, \tag{1.4}$$

$$\mathbf{\Phi}_{\text{NUFFT}} = \mathbf{U}\mathcal{F}_{\text{2D}}^{-1}\left\{\mathcal{F}_{\text{NUFFT}}[\mathcal{F}_{\text{2D}}(\cdot)]\right\}, \tag{1.5}$$

Note that the phase compensation term is omitted for brevity. Here, $\mathbf{\Theta}^{\dagger}$ represents the pseudoinverse 1-D Stolt transform with the nearest neighbor interpolation, $\mathcal{F}_{\text{3D}}$ denotes the 3-D FFT, and $\mathbf{U}$ denotes the binary matrix that is used to select the random (x, y) positions for random undersampling.

The cost function $J(\hat{\mathbf{g}})$ represents some l_1 regularization term with respect to $\hat{\mathbf{g}}$. In this paper, it is selected as

$$J(\hat{\mathbf{g}}) = \frac{\gamma_2}{2}\|\mathbf{\Psi}\hat{\mathbf{g}}\|_1 + \frac{\lambda}{2}\|\mathbf{D}\hat{\mathbf{g}}\|_1, \tag{1.6}$$

where $\|\cdot\|_1$ denotes the l_1 norm, $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ is the linear operator that transforms the image from voxel representation into a sparse representation, $\mathbf{D}$ is the discrete 3-D isotropic TV operator, and γ_2 and λ are the weights for the consistency of the l_1 norm and the TV norm, respectively.

MIMO UWA communications are challenging for three main reasons: first, the available channel bandwidth is very limited; second, the multipath delay is excessively long; third, the Doppler spread effect is significantly large. Accurate channel estimation is crucial in achieving satisfactory performance for MIMO UWA communications. The traditional block-based least squares (LS) or minimum mean-square-error (MMSE) channel estimation techniques often need the periodic training sequence and the inversion of large matrix. In comparison, the iterative channel estimation based on the adaptive algorithms, such as least mean squares (LMS) or recursive least square (RLS) algorithms [3], can better track the time-varying frequency-selective channel. As the required adaptive filter lengths grows, the conventional normalized LMS (NLMS) [3] algorithm suffers a slow convergence rate, thus requiring long training sequence. However, it is desirable to use short training sequence to reduce the overhead and increase the data transmission efficiency. With short training sequence, this slow convergence rate degrades the accuracy of the channel estimation.

This dissertation develops several robust iterative or adaptive signal processing algorithms to solve these sparse nonlinear optimization problems. For NEC, two adaptive solutions are proposed and evaluated under the environment of impulsive interference and colored input. For 3-D SAR image, two iterative reconstruction algorithms are proposed, which utilize the principle of compressed sensing (CS) and radar imaging. For MIMO UWA communications, the proposed iterative improved proportionate NLMS (IPNLMS) MIMO channel estimator is performed jointly with the MMSE turbo equalizer [5] at the receiver, which ensure efficient and effective soft information exchange between the equalizer and the decoder.

1.2 SUMMARY OF CONTRIBUTIONS

This dissertation consists of two journal publications and one conference paper as listed in the publication list. My contributions that are published or accepted are:

1. Proportionate affine projection sign algorithms for sparse system identification in impulsive interference. Two new proportionate affine projection sign algorithms (APSAs) are proposed for NEC applications where the impulse response is often real-valued with sparse coefficients and long filter length. The proposed proportionate-type algorithms can achieve fast convergence and low steady-state misalignment. Benefiting from the characteristics of l_1-norm optimization, affine projection, and proportionate matrix, the new algorithms are more robust to impulsive interferences and colored input than the proportionate least mean squares (PNLMS) algorithm [6] and the robust proportionate affine projection algorithm (Robust PAPA) [7]. The computational complexity of the new algorithms is lower than the affine projection algorithm (APA) family due to the elimination of the matrix inversion.

2. Compressed sensing approaches and comparative study for 3-D synthetic aperture radar image reconstruction. By utilizing the sparse nature of 3-D synthetic aperture radar (SAR) images, the proposed two compressed sensing (CS) approaches can reconstruct satisfactory images with undersampled measurements. Combining CS with the range migration algorithm (RMA), using either Stolt transform or nonuniform fast Fourier transform (NUFFT), yields two different approaches: Stolt-CS and NUFFT-CS. These approaches can decrease the load of data acquisition through l_1-norm optimization. Both the simulation and the experimental reconstruction results demonstrate that the NUFFT-CS achieves a good tradeoff between the reconstruction quality and the computational costs.

3. Robust iterative channel estimation based adaptive turbo equalizer in multiple-input multiple-output underwater acoustic communications. For UWA communications, a robust iterative channel estimation based MMSE adaptive turbo equalizer is proposed and studied for MIMO detection. Rather than the classical MMSE or NLMS estimation algorithms, the IPNLMS [8] is adopted for the iterative MIMO

channel estimator to utilize the sparse nature of UWA channel. The MIMO channel estimation is performed jointly with the MMSE equalizer and the maximum *a posteriori* probability (MAP) decoder. With inter-block interference removed, the MIMO MMSE equalization is performed with overlapped information subblocks without guard intervals, thus a high transmission efficiency is guaranteed and performance degradation is prevented. The proposed MIMO detection scheme has been tested by experimental data and proved to be robust against tough MIMO channels.

PAPER

I. PROPORTIONATE AFFINE PROJECTION SIGN ALGORITHMS FOR NETWORK ECHO CANCELLATION

Zengli Yang, Yahong Rosa Zheng, and Steven L. Grant

ABSTRACT—Two proportionate affine projection sign algorithms (APSAs) are proposed for network echo cancellation applications where the impulse response is often real-valued with sparse coefficients and long filter length. The proposed proportionate-type algorithms can achieve fast convergence and low steady-state misalignment by adopting a proportionate regularization matrix to the APSA. Benefiting from the characteristics of l_1-norm optimization, affine projection, and proportionate matrix, the new algorithms are more robust to impulsive interferences and colored input than the proportionate least mean squares (PNLMS) algorithm and the robust proportionate affine projection algorithm (Robust PAPA). The new algorithms also achieve much faster convergence rate in sparse impulse responses than the original APSA and the normalized sign algorithm (NSA). The new algorithms are robust to all types of NEC impulse response with different sparseness without the need to change parameters or estimate the sparseness of the impulse response. The computational complexity of the new algorithms is lower than the affine projection algorithm (APA) family due to the elimination of the matrix inversion.

1 INTRODUCTION

A network echo canceller (NEC) is an adaptive filter essential to modern voice communication networks. An NEC first estimates the echo path of the network, generates a replica of the network impulse response (NIR), and then subtracts the replica of the far-end echo from the near-end signal to obtain clean signals. Although NEC is a classic system identification application, modern VoIP (Voice over Internet Protocol) and telephone networks impose several challenges on conventional adaptive filters. The first is that today's large scale network often experiences long delays and causes long network impulse responses (NIRs) that require adaptive filters with 1000 or more taps; while the traditional network only requires adaptive filters with less than 100 taps. However, the response of the long-distance/local-loop interface is still typically only 6 to 12 milliseconds long, which with a typical sample rate of 8 kHz yields a flat delay of zero to nearly 900 samples followed by a damped ringing of about 50 samples. This means that the NIR is typically rather sparse, in that most of the coefficients are close to zero. The second challenge is that the excitation signals in NECs are speech signals which are often highly colored non-Gaussian processes rather than white Gaussian noises. Many adaptive algorithms suffer reduced convergence rate in this scenario. The third challenge of NEC is due to double-talk where both near-end speech and far-end speech are active simultaneously. This often causes extremely slow convergence or even divergence in adaptive filters.

As the required adaptive filter lengths grow, the conventional normalized least mean squares (NLMS) [1] algorithm exhibits a slower convergence rate. This slow convergence rate becomes noticeable in that echo is often heard, especially in the first few seconds of a connection. The proportionate NLMS (PNLMS) [2] has been designed to ameliorate this situation by exploiting the sparse nature of the NIR. By

selecting a proportionate matrix at each iteration, PNLMS updates each coefficient in the weight vector proportionate to its magnitude. This results in very fast initial convergence for sparse NIRs relatively independent of their length. However, the drawback of PNLMS is that, though it has fast initial convergence for sparse NIRs, it has slower convergence than NLMS for non-sparse NIRs. This problem has been addressed by several modifications to PNLMS. The first is PNLMS++ [3] which has two versions, one where the adaptation algorithm alternates between both PNLMS and NLMS in successive sample periods and another where both updates are combined in each sample period. The resulting convergence is generally the better of the two algorithms. That is, PNLMS++'s convergence is like PNLMS's for sparse NIRs and like NLMS's for dispersive NIRs. Another modification to PNLMS is the improved PNLMS (IPNLMS) [4] which has the feature of being optimal for a given NIR sparseness. This feature has later been exploited in a class of sparseness-controlled (SC) [5] algorithms which measure the sparseness of the developing coefficients on-the-fly. Another approach is to use adaptive combination of proportionate filters [6] which adaptively mix the outputs of two independent adaptive filters together based on IPNLMS. In addition, the μ-law PNLMS (MPNLMS) [7] is an optimal step-size algorithm modified from PNLMS.

A number of proportionate algorithms are also developed for the affine projection algorithm (APA) [1] which is well known for its better convergence than NLMS for colored input. These proportionate algorithms include proportionate APA (PAPA), improved proportionate APA (IPAPA) [8] and "memory"-IPAPA (MIPAPA) [9]. The IPAPA extends the proportionate matrix of IPNLMS [4] directly to APA; while the MIPAPA designs an efficient matrix to reduce computational complexity. Both algorithms improve convergence rate over that of PNLMS in practical NEC applications where the inputs are speech and the NIRs are sparse. However, they have higher

complexity than PNLMS algorithms because they generally require a matrix inversion where the size of the matrix is the order of the projection. In practice, the projection orders are typically around ten and direct matrix inversion of this size may be too expensive.

In addition to convergence rate and complexity, another challenging problem in NEC is the double-talk problem. Typically, a double-talk detector (DTD) is used to detect this situation, and the adaptation of the coefficients is inhibited by setting the step-size to zero for a period T_{hold} during double-talk. Completely reliable DTDs are notoriously difficult to design because the NEC has stringent requirements on its detection. Even the first few samples of double-talk can lead to divergence of the adaptive filter. One approach to mitigate the effect of double-talk is to design the adaptive filter to be especially robust to DTD errors. Several algorithms have been developed in this robustness approach, such as the robust PNLMS and robust PAPA [10]. Another approach treats the double-talk as variable background noise and adapts the step-size of the adaptation algorithm accordingly. This is the philosophy behind variable step-size NLMS (VSS-NLMS) [11] and variable step-size APA (VSS-APA) [12]. Yet a third strategy is to use adaptive algorithms based on l_1 rather than l_2 error norms because l_1 algorithms are especially robust to impulsive noise like speech. A conventional l_1 algorithm is the normalized sign algorithm (NSA). Unfortunately, NSA's robustness comes at the price of slower than NLMS convergence. Recently, a new affine projection sign algorithm (APSA) [13] addresses this problem and provides good robustness and fast convergence. Indeed it has been shown [13] that APSA achieves faster convergence and lower steady-state normalized misalignment than NLMS, APA, and NSA under impulsive interference. This is achieved without the need for a matrix inversion as in APA.

In this paper we combine the proportionate approach with APSA to obtain even faster convergence when the echo path is sparse. The resulting algorithm is called

proportionate APSA. Two types of proportionate matrix, one based on PNLMS [2] and another based on IPNLMS [4], are applied to the APSA for real-coefficient systems and the resulting real-coefficient proportionate APSA (RP-APSA) and real-coefficient improved proportionate APSA (RIP-APSA) achieve fast convergence in sparse NEC applications with robustness to colored input and double-talk. The computational complexity of the two proportionate APSAs is slightly higher than the original APSA but is lower than the APA family. The RIP-APSA also exhibits especially good robustness in all types of NIRs without the need to change parameters or estimate the sparseness of the NIRs.

2 PROPORTIONATE AFFINE PROJECTION SIGN ALGORITHMS

Consider the NEC scheme shown in Fig. 1, where $x(k)$ is the far-end signal, $z(k)$ and $v(k)$ are the near-end speech and background noise signal, respectively. The NIR of the true echo path is denoted by a length L coefficient vector $\mathbf{h}$. The level of the sparseness in the NIR may vary according to the changing network environment, which is measured by [14],

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

where $0 \le \xi \le 1$, and the p-norm is defined by

$$\|\mathbf{h}\|_p := \left(\sum_{l=0}^{L-1}|h_l|^p\right)^{1/p}. \tag{2}$$

In the extreme case, if $\mathbf{h}$ is a pure impulse, then $\xi = 1$. On the other hand, if all the elements of $\mathbf{h}$ have equal value, then $\xi = 0$. In other words, a larger ξ corresponds to a sparser impulse response, while a smaller ξ corresponds to a more dispersive impulse response. .

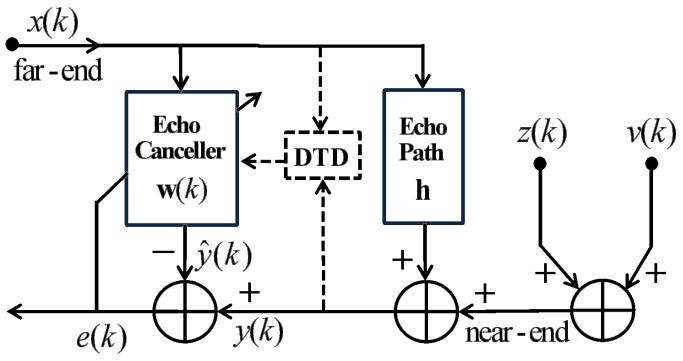


Figure 1. Structure of a network echo canceller (NEC).

The estimated impulse response of the NEC is denoted by $\mathbf{w}(k) = [w_0(k), w_1(k), \ldots, w_{L-1}(k)]^T$, where k is the time index and the superscript T denotes transpose. The signal $y(k)$ contains the echo, the near-end speech, and background noise. That is, $y(k) = \mathbf{x}^T(k)\mathbf{h} + z(k) + v(k)$, where $\mathbf{x}(k) = [x(k), x(k-1), \ldots, x(k-L+1)]^T$ is the far-end signal vector. Generating the replica echo $\hat{y}(k) = \mathbf{x}^T(k)\mathbf{w}(k)$, the NEC tries to minimize the difference between $y(k)$ and $\hat{y}(k)$ with an adaptive $\mathbf{w}(k)$.

Different from the l_2-norm algorithms [1,2,4], the original APSA algorithm [13] is obtained by minimizing the l_1-norm of the *a posteriori* error vector with a constraint on the filter coefficients,

$$\min_{\mathbf{w}(k+1)} \quad \|\mathbf{y}(k) - \mathbf{X}^T(k)\mathbf{w}(k+1)\|_1 \tag{3}$$

$$\text{subject to} \quad \|\mathbf{w}(k+1) - \mathbf{w}(k)\|_2^2 \leq \mu^2 \tag{4}$$

where μ^2 is a parameter to ensure the weight coefficient vector does not change too much in one iteration, $\mathbf{X}(k) = [\mathbf{x}(k), \mathbf{x}(k-1), \ldots, \mathbf{x}(k-M+1)]$, $\mathbf{y}(k) = [y(k), y(k-1), \ldots, y(k-M+1)]^T$, and M is the projection order. Using the method of Lagrange multipliers, we get

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{1}{2\lambda}\mathbf{X}(k)\text{sgn}[\mathbf{e}(k)] \tag{5}$$

where λ is a Lagrange multiplier, the error vector $\mathbf{e}(k) = \mathbf{y}(k) - \mathbf{X}^T(k)\mathbf{w}(k)$, and sgn$[\cdot]$ is the signum function. For sparse $\mathbf{h}$, we would like to adapt the coefficients of $\mathbf{w}(k)$ proportionately by pre-multiplying the update vector with a proportionate matrix. Then, (5) can be rewritten as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{1}{2\lambda}\mathbf{G}(k)\mathbf{X}(k)\text{sgn}[\mathbf{e}(k)] \tag{6}$$

where $\mathbf{G}(k)$ =diag$\{g_0(k), \ldots, g_{L-1}(k)\}$ is a diagonal proportionate matrix whose elements may be selected according to [2, 4]. Using (4) and (6), we obtain

$$\frac{1}{2\lambda} = \frac{\mu}{\sqrt{\mathbf{x}_{gs}^T(k)\mathbf{x}_{gs}(k)}} \tag{7}$$

where $\mathbf{x}_{gs}(k) = \mathbf{G}(k)\mathbf{X}(k)\text{sgn}[\mathbf{e}(k)]$. Substituting (7) into (6) and adding a small positive parameter δ to avoid possible division by zero, the weight updating equation for the proportionate APSAs is

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\mu\mathbf{x}_{gs}(k)}{\sqrt{\delta + \mathbf{x}_{gs}^T(k)\mathbf{x}_{gs}(k)}} \tag{8}$$

where μ is regarded as the step size satisfying $\mu > 0$.

In this paper, we choose the proportionate matrix $\mathbf{G}(k)$ according to [2] and [4], and for real-valued systems, we call the resulting proportionate-type algorithms real-coefficient proportionate APSA (RP-APSA) and real-coefficient improved proportionate (RIP-APSA), respectively. The proposed two new algorithms are summarized in Table 1. For RP-APSA, the parameter ρ prevents $w_l(k)$ from stalling when it is much smaller than the largest coefficient and q regularizes the updating when all coefficients are zero at initialization. For $\alpha = -1$, the RIP-APSA and APSA are identical. For α close to 1, the RIP-APSA behaves like the RP-APSA.

Table 1. Proportionate Affine Projection Sign Algorithms

Algorithms	RP-APSA	RIP-APSA
Initialization	$\mathbf{w}(0)=\mathbf{0}_{L\times 1}$	
Parameters	ρ=0.01 or 0.1, q=0.01	α=0 or -0.5, ε=0.01
Proportionate matrix	$\gamma_{\min}=\rho\max(q,\|w_0(k)\|,\ldots,\|w_{L-1}(k)\|)$ $\gamma_l(k)=\max(\gamma_{\min},\|w_l(k)\|)$ $g_l(k)=\frac{\gamma_l(k)}{\|\gamma_l(k)\|_1/L}$	$g_l(k)=\frac{1-\alpha}{2L}+\frac{(1+\alpha)\|w_l(k)\|}{2\|w_l(k)\|_1+\varepsilon}$
Adaptation	$\mathbf{G}(k)=\text{diag}\{g_0(k),\ldots,g_{L-1}(k)\}$ $\mathbf{e}(k)=\mathbf{y}(k)-\mathbf{X}^T(k)\mathbf{w}(k),\ \mathbf{x}_{gs}(k)=\mathbf{G}(k)\mathbf{X}(k)\text{sgn}[\mathbf{e}(k)]$ $\mathbf{w}(k+1)=\mathbf{w}(k)+\frac{\mu\mathbf{x}_{gs}(k)}{\sqrt{\delta+\mathbf{x}_{gs}^T(k)\mathbf{x}_{gs}(k)}}$	

3 COMPUTATIONAL COMPLEXITY

The computational complexity of the proportionate APSAs are compared with that of conventional algorithms in terms of the total number of additions, multiplications, divisions, comparisons, square-roots, and direct matrix inversions (DMIs). With filter length L and projection order M, the complexities are shown in Table 2. Although APA behaves better than NLMS, APA has higher complexity because the number of multiplications and the size of the DMI increase proportionately to M^2. In contrast, the APSA does not require matrix inversion thus the projection order M does not affect the number of multiplications, and the number of additions is only linearly dependent on M. Note that APSA does not require matrix inversion, thus the APSAs are more efficient than APA. With a modest increase in the proportionate matrix computation, the proportionate APSAs behave much better than the original APSA, especially for sparse NIRs.

Table 2. Computational Complexity of Algorithms' Coefficients Update
(DMI = Direct matrix inversion)

Algorithm	ADD	MUL	DIV	CMP	SQRT	DMI (size)
NLMS	L	$2L+1$	1	0	0	No
PNLMS	$2L-1$	$4L+1$	2	$2L$	0	No
IPNLMS	$3L$	$4L+1$	2	0	0	No
APA	$(M^2+M-1)L$	$(M^2+M+1)L+M^2$	0	0	0	$M\times M$
IPAPA	$(M^2+M+1)L$	$(M^2+M+2)L+M^2$	0	0	0	$M\times M$
APSA	ML	$2L$	1	0	1	No
RP-APSA	$(M+1)L-1$	$5L$	2	$2L$	1	No
RIP-APSA	$(M+2)L$	$5L$	2	0	1	No

4 SIMULATION RESULTS

The proposed algorithms were evaluated via computer simulations. The echo path had 512 coefficients with some significant (active) coefficients and many near zero (inactive) coefficients. Fig. 2 shows three typical impulse responses of the echo path with sparseness measures $\xi = 0.895$, $\xi = 0.758$ and $\xi = 0.556$, respectively. Throughout our simulations except the double-talk scenario (Section 4.2), the input signal $x(k)$ was a first order autoregressive signal (AR(1)) with a pole at 0.8; while the white Gaussian noise (WGN) $v(k)$ was added to the near-end to give a signal-to-noise ratio (SNR) of 30 dB. The near-end signal $z(k)$ was a strong impulsive interference with a signal-to-interference ratio (SIR) of -10 dB and it was modeled by a Bernoulli-Gaussian (BG) signal [13, 15]. The BG distribution was generated as the product of a Bernoulli process and a Gaussian process, *i.e.*, $z(k) = \omega(k)n(k)$, where $n(k)$ was WGN with zero mean and variance σ_n^2, and $\omega(k)$ was a Bernoulli process with the probability mass function given as $P(\omega) = 1 - Pr$ for $\omega = 0$, and $P(\omega) = Pr$ for $\omega = 1$. The average power of the BG process was $Pr \cdot \sigma_n^2$. Keeping the average power constant, the BG process was spikier when Pr was smaller. It reduced to a Gaussian process when $Pr = 1$. We chose $Pr = 0.001$ for our simulations.

For the double-talk scenario in Section 4.2, both near-end and far-end signals were speech segments and the background noise was WGN. The SIR and SNR were 6 dB and 30 dB, respectively. The classical Geigel DTD [10, 16] was used to inhibit adaptation by setting $\mu = 0$ for T_{hold} period when both near-end speech and far-end speech were detected as being active simultaneously. In Geigel DTD, double-talk is declared if

$$|y(k)| \geq \vartheta \max |x(k)|, |x(k-1)|, ..., |x(k-L+1)| \tag{9}$$

where ϑ is the detection threshold.

The performance of the algorithms was mainly measured by the normalized misalignment and the excess mean-square error (MSE) (in dB). The normalized misalignment is defined by

$$\eta(k) = 10\log_{10}\frac{\|\mathbf{h}-\mathbf{w}(k)\|_2^2}{\|\mathbf{h}\|_2^2}. \tag{10}$$

Excess MSE reflects the MSE that is in excess of the minimum MSE, which is based on ensemble averaging of the squared estimation excess error $|e_{ex}(k)|^2$. Here,

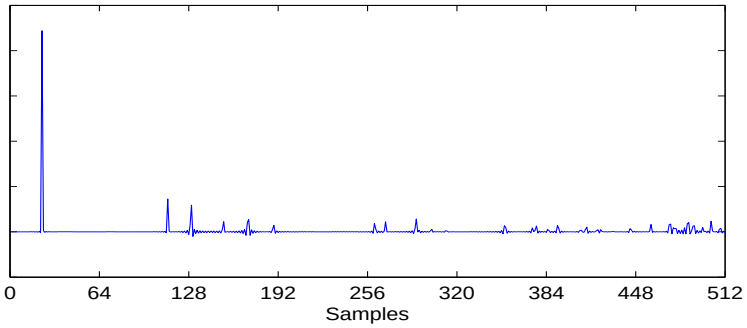
$$e_{ex}(k) = \mathbf{x}^T(k)\mathbf{h} - \mathbf{x}^T(k)\mathbf{w}(k). \tag{11}$$

In our simulations, both misalignment and excess MSE behaved similarly in all cases. Therefore, only misalignment curves are shown here for brevity. In addition, the regularization parameter $\delta = 0.01$ for all the following algorithms except the improved proportionate-type algorithms [4], where $\delta_{IP} = \frac{1-\alpha}{2L}\delta$. The simulation results shown were obtained by ensemble averaging ten independent trials.

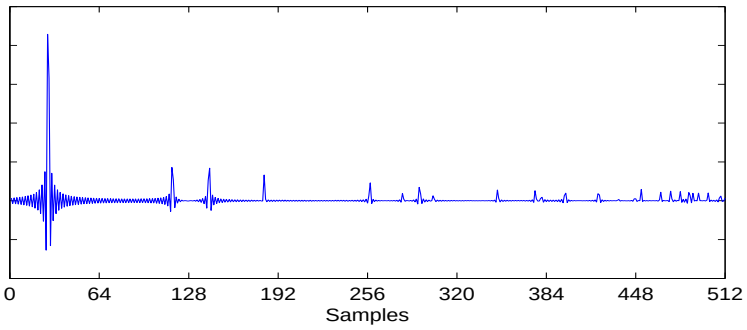
4.1 PERFORMANCE COMPARISON BETWEEN THE PROPORTIONATE APSAS AND OTHER ALGORITHMS

The normalized misalignment of the l_1-norm algorithms were compared with that of the l_2-norm ones. The NIR of the echo path was sparse with $\xi = 0.758$, as shown in Fig. 2b. The proportionate parameters ρ and α were selected as in reference papers [2,4], as shown in Fig. 3, Fig. 4 and Fig. 5. Without interference, the step-size μ was adjusted to make the algorithms achieve the steady-state normalized misalignment of -14 dB, as shown in Fig. 3. Considering only the background noise $v(k)$, RP-APSA converged slower than PAPA but faster than conventional APA, while APSA had slowest convergence.

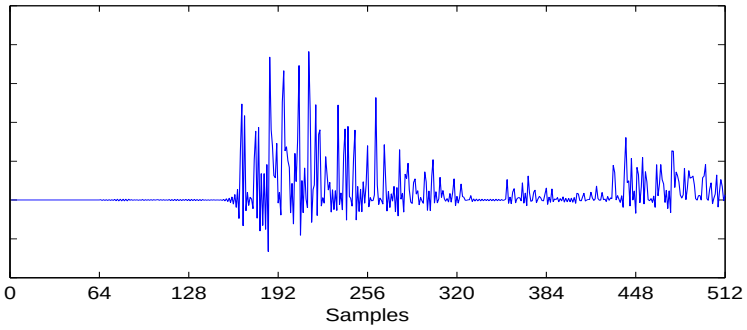
With strong interference, the step-size μ for the l_2-norm algorithms was 0.1 and was adjusted to achieve the same steady-state normalized misalignment for the



(a)



(b)



(c)

Figure 2. Typical impulse responses of the echo path for NEC applications. (a) Sparse impulse response with $\xi = 0.895$. (b) Sparse impulse response with $\xi = 0.758$. (c) Dispersive impulse response with $\xi = 0.556$.

l_1-norm algorithms, as shown in Fig. 4 and Fig. 5. Benefiting from the robustness of l_1-norm minimization, the APSA family and the NSA converged at different speeds, while the NLMS family and the APA family diverged under strong impulsive interference.

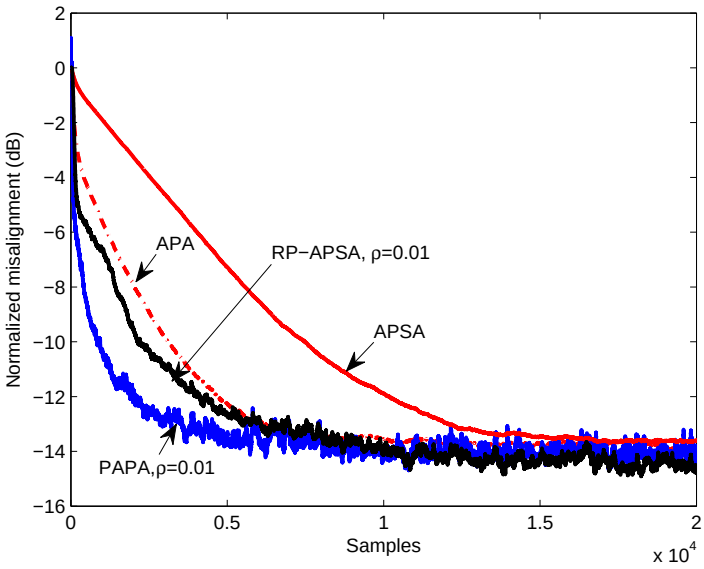


Figure 3. Normalized misalignment comparison of the APA, PAPA, APSA and RP-APSA without interference. $M = 2$, $\mu_{APA} = \mu_{PAPA} = 0.15$, $\mu_{APSA} = 0.0005$, $\mu_{RP-APSA} = 0.001$, AR(1) input with SNR=30 dB.

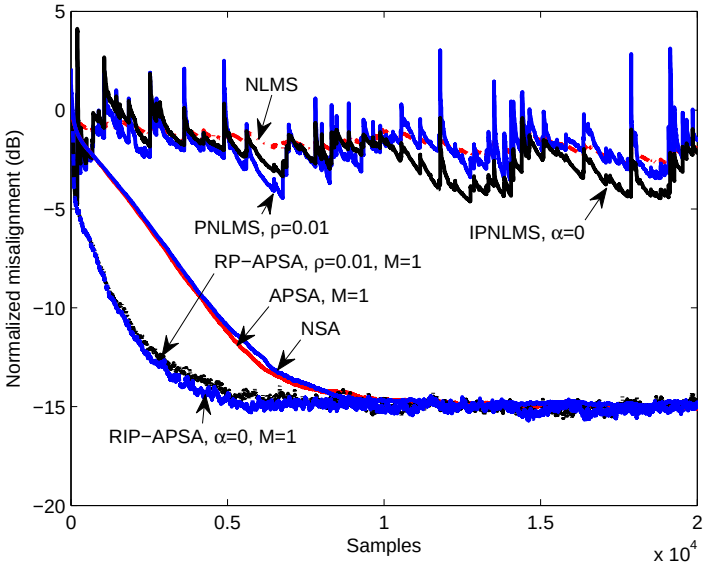


Figure 4. Normalized misalignment comparison of the NLMS, PNLMS, IPNLMS, NSA, APSA, RP-APSA and RIP-APSA with interference. $M = 1$, $\mu_{NLMS} = \mu_{PNLMS} = \mu_{IPNLMS} = 0.1$, $\mu_{NSA} = 0.3$, $\mu_{APSA} = 0.013$, $\mu_{RP-APSA} = 0.024$, $\mu_{RIP-APSA} = 0.022$, AR(1) input with SNR=30 dB and SIR=−10 dB. Strong BG interference with $Pr = 0.001$.

Moreover, the steady-state normalized misalignment of the APSA family with the projection order $M = 1$ and $M = 2$ were adjusted to be -15 dB and -20 dB, respectively. With $M = 1$ and using only sign arithmetic, APSA performed almost the same as NSA against impulsive interference. Adding the affine projection arithmetic ($M = 2$), APSA outperformed NSA, as shown in Fig. 5. Combining the decorrelation property of affine projection and exploiting the sparse nature of the NIR, proportionate APSAs had faster convergence than APSA and NSA with both values of M. Also, the RIP-APSA converged a little faster than RP-APSA under such conditions. The advantage of proportionate APSAs for a sparse NIR is clearly demonstrated.

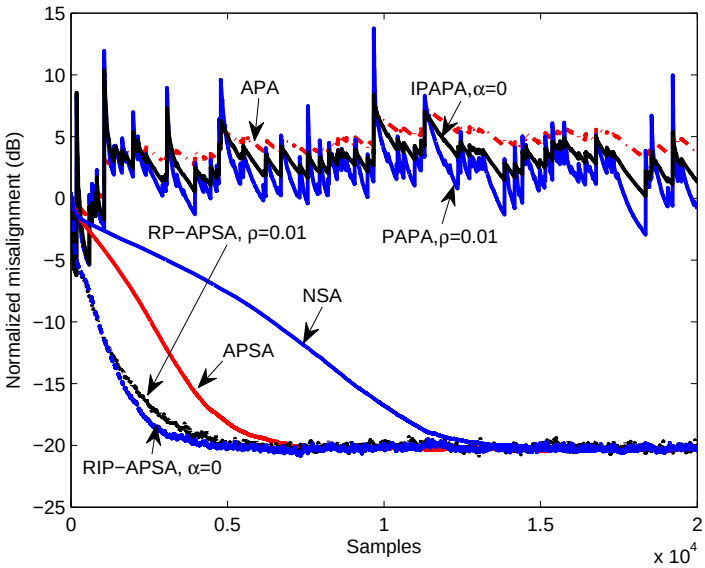


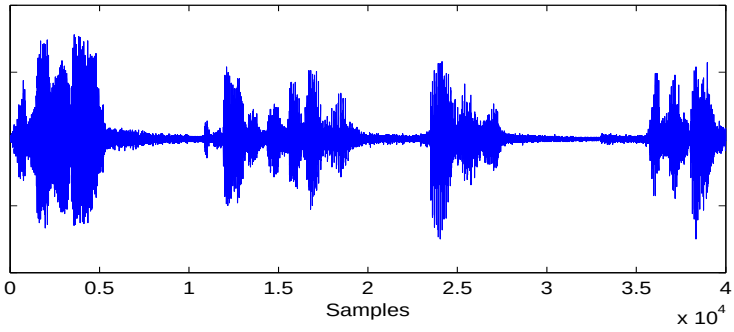
Figure 5. Normalized misalignment comparison of the APA, PAPA, IPAPA, NSA, APSA, RP-APSA and RIP-APSA with interference. $M = 2$, $\mu_{APA} = \mu_{PAPA} = \mu_{IPAPA} = 0.1$, $\mu_{NSA} = 0.15$, $\mu_{APSA} = 0.01$, $\mu_{RP-APSA} = 0.017$, $\mu_{RIP-APSA} = 0.014$, other parameters were the same as those in Fig. 4.

4.2 DOUBLE-TALK SCENARIO

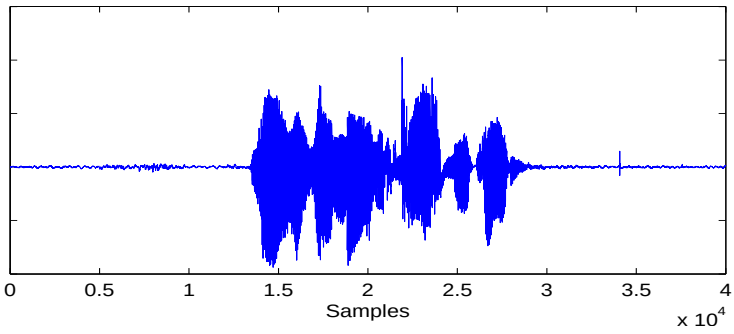
The performance of the APSA family for speech signals during double-talk were compared with the non-robust and robust PAPAs [10] with projection order $M = 5$. Parameters for the robust algorithms and double-talk were almost the same as those in [10]. In detail, $(\hat{\lambda}, k_0, \beta) = (0.997, 1.1, 0.60665)$ for the robust PAPA, the average far-end to double-talk ratio was 6 dB (SIR = 6 dB), the Geigel detector assumed 6 dB attenuation ($\vartheta = 0.5$), and the hang-over time, $T_{hold} = 240$ samples. We chose the echo path shown in Fig. 2b to obtain 20 dB hybrid attenuation. The far-end speech and near-end speech used in the double-talk simulation are shown in Fig. 6. Double-talk happened in the period with sample index of $[1.4, 2.8] \times 10^4$. We chose the step-size $\mu = 0.0002$ for the APSA family and $\mu = 0.005$ for the non-robust and robust PAPAs to achieve their best performance. A larger μ will make the robust PAPA diverge faster during double-talk under such parameter settings. As shown in Fig. 7, both PAPAs behaved almost the same most of the time, although robust PAPA outperformed non-robust PAPA against some double-talk disturbance. In comparison, the APSA family were more robust against double-talk than robust PAPA, and the proportionate APSAs achieved lower normalized misalignment and faster convergence rate than all other algorithms. Without the need to change the parameter settings to obtain a robustness feature, the proportionate APSAs exhibited better robustness to double-talk than robust PAPA.

4.3 CHARACTERISTICS OF THE PROPORTIONATE APSAS

The performance of the RP-APSA and RIP-APSA with different projection order M were studied using AR(1) input and BG interference. The step size $\mu = 0.01$, echo path and other parameters were the same as those in Fig. 4. It has been shown that, for the APA family, a larger projection order M leads to faster convergence with higher steady-state misalignment in stationary environments [17]. In contrast,



(a)

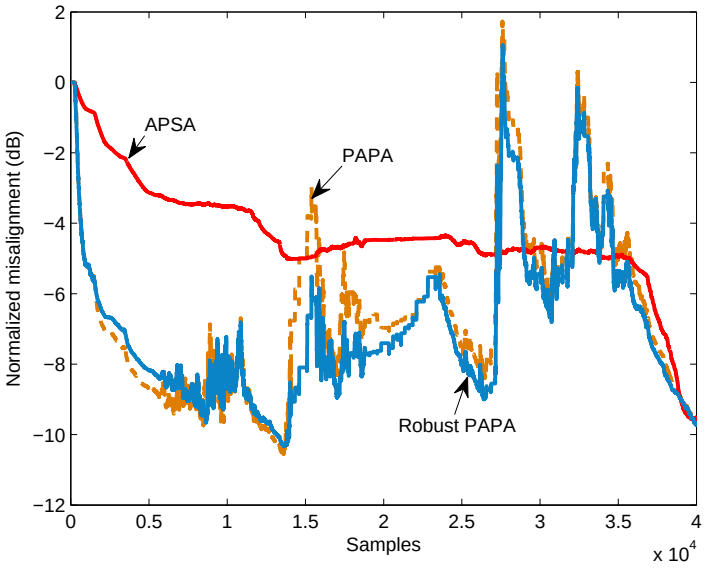


(b)

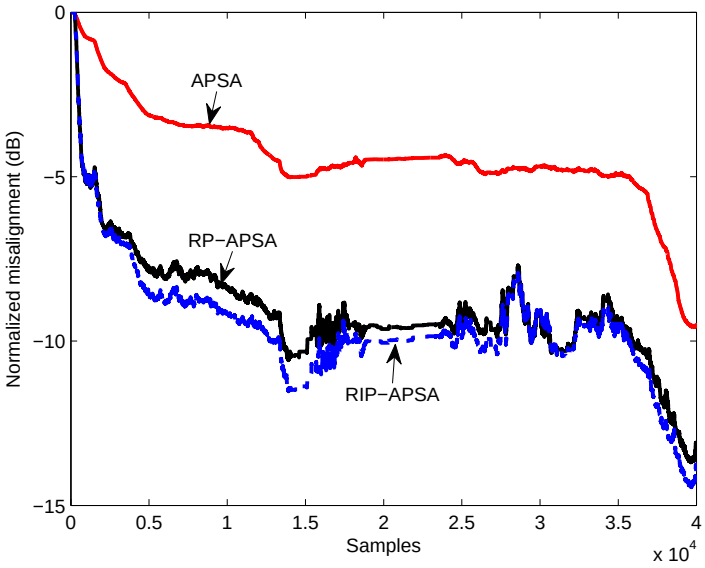
Figure 6. Speech signals used in the double-talk scenario. Average Far-end speech to double-talk ratio is 6 dB (1.4×10^4th - 2.8×10^4th samples). (a) Far-end speech. (b) Near-end speech.

for the RP-APSA and RIP-APSA, a larger M achieved both faster convergence and lower steady-state misalignment in sparse NIR, as shown in Fig. 8. Benefiting from the convergence behavior (see the Appendix), the new algorithms can exploit the advantage of affine projection even in the steady-state.

The effects of proportionate matrix were investigated for both sparse and dispersive NIRs using the AR(1) input and BG interference. Impulse responses shown in Fig. 2a and Fig. 2c were used in the simulation with step size $\mu = 0.005$. The results are shown in Fig. 9 for RP-APSA and in Fig. 10 for RIP-APSA. Similar to the effect of ρ on PNLMS as shown in [2], the RP-APSA with a smaller ρ ($\rho = 0.01$) lead to a higher convergence rate and lower steady-state misalignment simultaneously

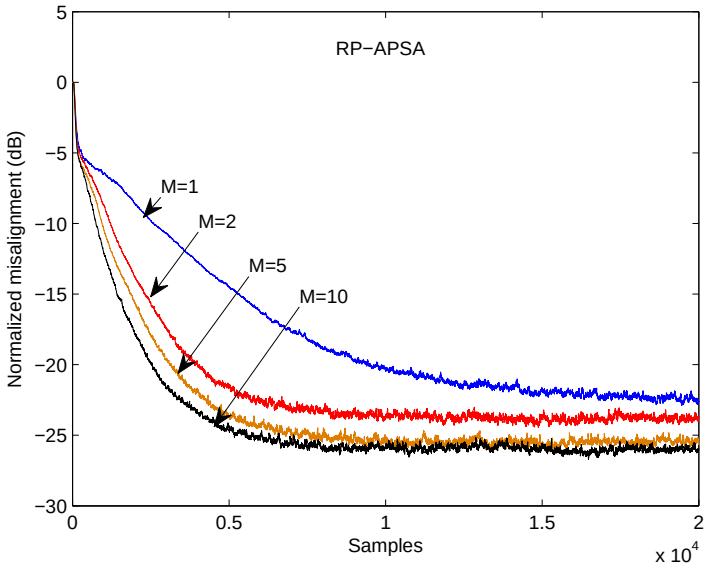


(a)

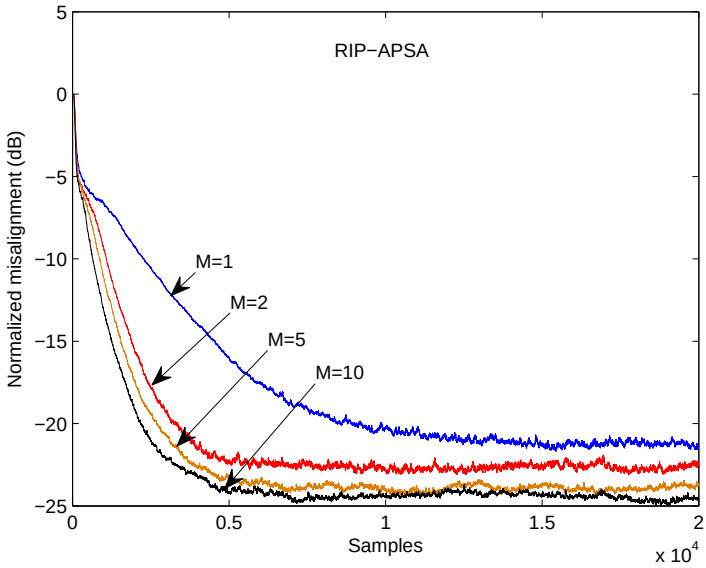


(b)

Figure 7. Performance of the APSA family, nonrobust and robust PAPA for speech signals during double-talk. $M = 5$, $\rho = 0.01$, $\alpha = 0$. (a) $\mu_{APSA} = 0.0002$, $\mu_{PAPA} = \mu_{Robust\ PAPA} = 0.005$, $(\hat{\lambda}, k_0, \beta) = 0.997, 1.1, 0.60665)$. (b) $\mu_{APSA} = \mu_{RP-APSA} = \mu_{RIP-APSA} = 0.0002$.



(a)



(b)

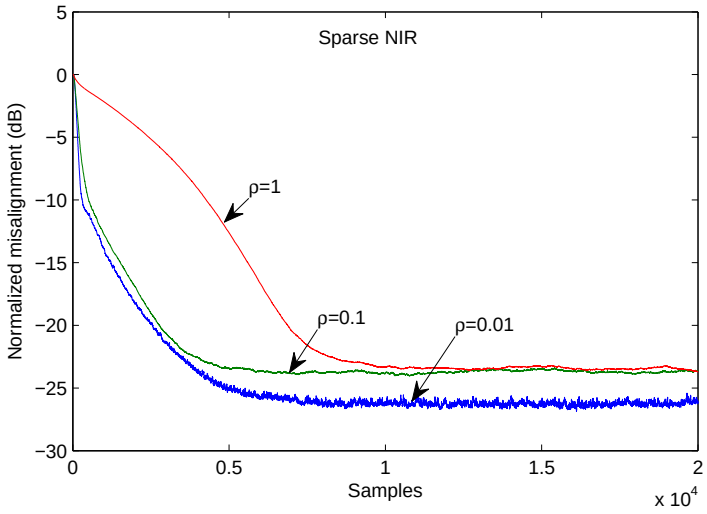
Figure 8. Normalized misalignment of the proportionate APSA with varying projection orders $M = 1, 2, 5, 10$. The input, interference and echo path were the same as those in Fig. 4. The step size $\mu = 0.01$. (a) RP-APSA, $\rho = 0.01$. (b) RIP-APSA, $\alpha = 0$.

for sparse NIR because more proportionality of the NIR was exploited. However, for the dispersive NIR, little proportionality can be utilized and most of the filter coefficients were updated at a similar rate. Therefore, if ρ is too small, it deteriorated the performance of RP-APSA with slower convergence but maintained lower steady-state misalignment than that of a large ρ.

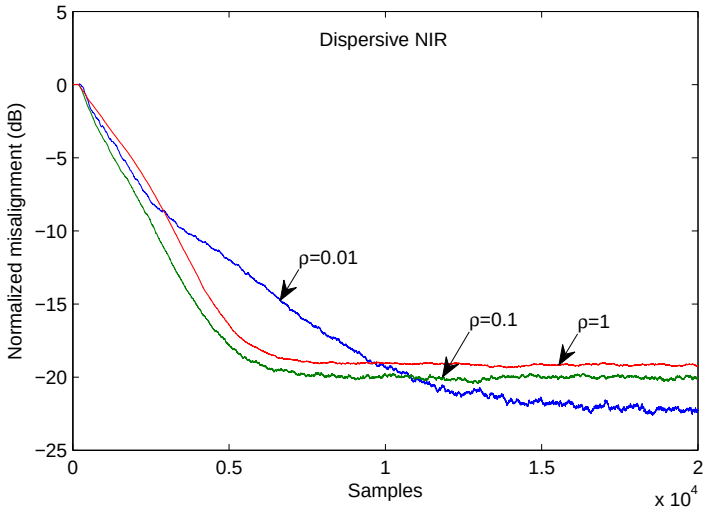
The effect of α on the convergence of the RIP-APSA are shown in Fig. 10 for both sparse and dispersive NIRs. According to [4], $\alpha = 0$ and $\alpha = -0.5$ were good choices for the improved proportionate algorithms to achieve robustness to different sparseness of NIRs in practice. The RIP-APSA with $\alpha = 0$ achieves a little lower steady-state misalignment in both NIRs. In contrast to the effect of ρ on the RP-APSA, the RIP-APSA is less sensitive to the value of α.

4.4 EFFECT OF SPARSENESS OF THE IMPULSE RESPONSES

Ten different NIRs with sparseness measure $0.556 \leq \xi \leq 0.938$ were used to study the effect of sparseness on the convergence of the proportionate APSAs. All step-sizes were adjusted so that the algorithms achieved almost the same steady-state normalized misalignment between -27 dB to -25 dB. The number of samples taken to achieve -25 dB normalized misalignment were used as a measure for the convergence rate versus the sparseness measure, as shown in Fig. 11. The projection order $M = 2$ was employed for the APSA family. The RIP-APSA with both $\alpha = -0.5$ and $\alpha = 0$ outperformed the other two algorithms over the entire sparseness region, and the number of samples taken to converge decreased approximately linearly with the increase of spareness. In comparison, APSA performed worst except in dispersive NIRs with $\xi < 0.64$, where RP-APSA with $\rho = 0.01$ had particularly slow convergence. With $\rho = 0.1$, RP-APSA behaved similarly as RIP-APSA in the entire ξ range. For sparse NIRs with $\xi > 0.7$, both RP-APSA and RIP-APSA had



(a)



(b)

Figure 9. Convergence of the RP-APSA for different ρ using $\mu = 0.005$. The input and interference were the same as those in Fig. 4. Impulse responses in Fig. 2(a) and (c) were used as the sparse and dispersive impulse responses, respectively. (a) In the sparse NIR with $\xi = 0.895$. (b) In the dispersive NIR with $\xi = 0.556$.

similar high convergence rates regardless of the selection of the proportionate matrix parameters ρ and α.

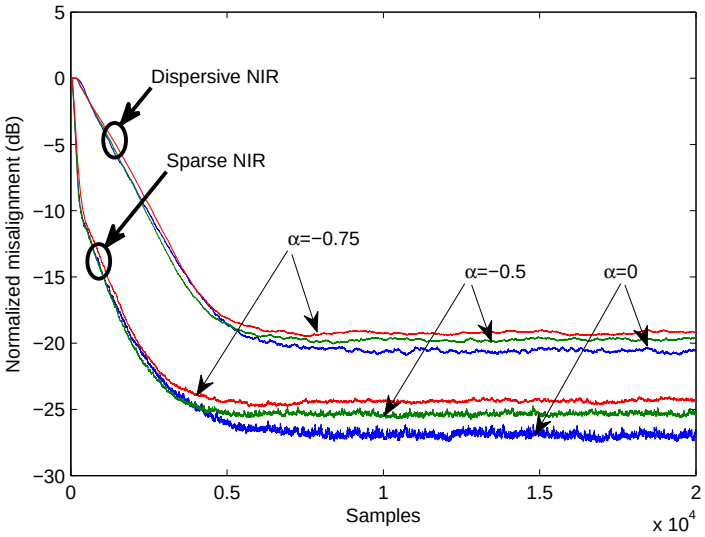


Figure 10. Convergence of the RIP-APSA for different values of α using $\mu = 0.005$. The input and the interference were the same as those in Fig. 4.

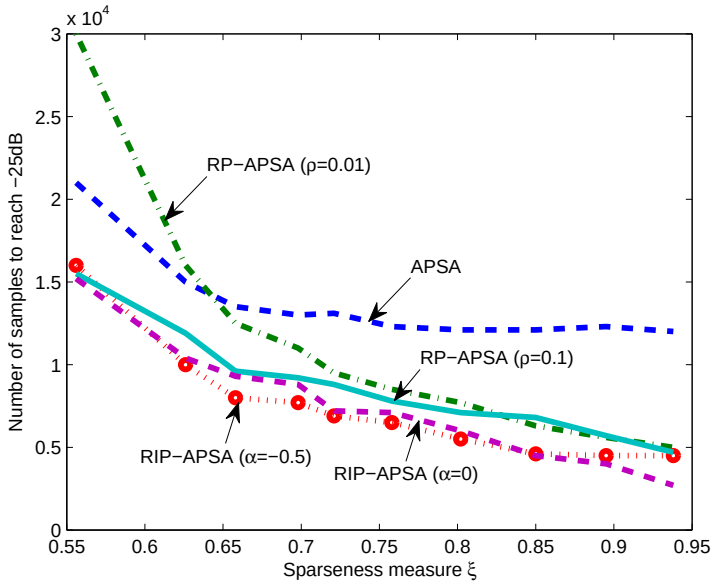


Figure 11. Number of samples to reach the -25 dB normalized misalignment against different sparseness measure of ten systems for the APSA, RP-APSA and RIP-APSA. The input and the interference were the same as those in Fig. 4. $M = 2$, $\rho = 0.01$ or 0.1 for the RP-APSA and $\alpha = -0.5$ or $\alpha = 0$ for the RIP-APSA.

4.5 TRACKING PERFORMANCE OF THE PROPORTIONATE APSAS

The tracking performance of the APSA, RP-APSA and RIP-APSA was also studied in the BG interference environment. As in the previously discussed simulations, we chose the step-sizes $\mu_{APSA} = 0.003$, $\mu_{RP-APSA} = 0.006$, $\mu_{RIP-APSA} = 0.005$ to make sure they achieved the same steady-state normalized misalignment under sparse NIR. The echo path changed from sparse (Fig. 2a) to dispersive (Fig. 2c) NIR at the 2×10^4th sample. It can be seen in Fig. 12 that the RIP-APSA tracked the change faster than the APSA and RP-APSA, but with slightly higher steady-state misalignment in the dispersive NIR. The APSA had similar slow convergence in both sparse and dispersive NIRs. Its steady-state misalignment was slightly lower than those of the RP-APSA and RIP-APSA for dispersive NIR. The tracking performance of RP-APSA was in between APSA and RIP-APSA.

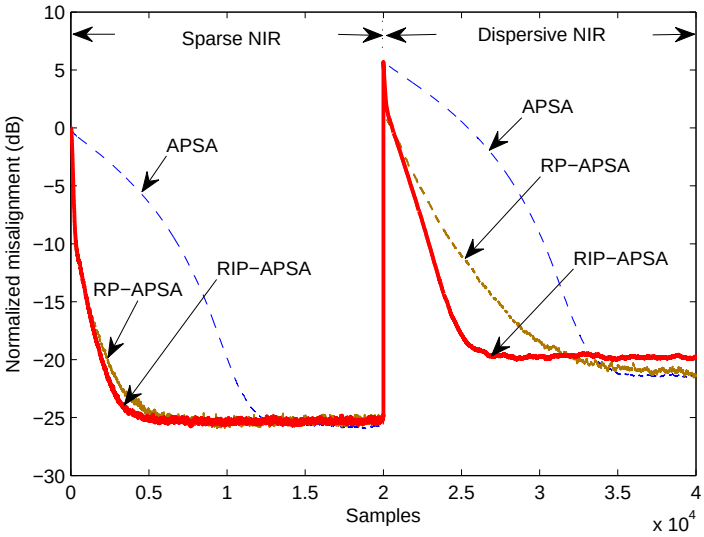


Figure 12. Tracking performance of the APSA, RP-APSA and RIP-APSA. The input and the interference were the same as those in Fig. 4. The echo path changed from sparse to dispersive at the $2{\times}10^4$th samples. $\rho = 0.01$, $\alpha = -0.5$, $\mu_{APSA} = 0.003$, $\mu_{RP-APSA} = 0.006$, $\mu_{RIP-APSA} = 0.005$.

4.6 SELECTION OF STEP SIZE

The relationship between step size and convergence behavior of the l_2 norm algorithms has been studied intensively in the literature [18, 19, 1, 20, 17, 21]. It has been shown that the step size of the NLMS and APA has to be bounded to $0 < \mu < 2$ to ensure stability. However, the stochastic model for the l_1 norm algorithms is more difficult than that of the l_2 norm family. Limited studies are devoted on the convergence analysis of the sign algorithm and its variations [22, 23, 24, 25, 26, 27, 28]. Attempts to finding a stability bound for the sign algorithm family have been reported in [22, 23, 24] using a second order stochastic model similar to that in the analysis of l_2 algorithms. However, this approach is proved to be incorrect for l_1 algorithms [25]. Instead, an interesting result for l_1 algorithms is that the sign algorithm converges asymptotically for any step size $\mu > 0$ [25]. This result is also proven to be true [27, 28] for the sign algorithm family without the Gaussianity or independence assumptions. This property does not present in the l_2 norm algorithms and it proves to be a significant advantage of the sign algorithm family. Upper bounds for the time-averaged mean absolute deviation (weight misalignment) and time-averaged MSE at steady-state are found to be functions of the step size μ [25,27], which gives guidelines for choosing step size in practical applications.

As a member of the sign algorithm family, the APSA and proportionate APSAs exhibits similar convergence behavior to that of the pure sign algorithm in that any step size $\mu > 0$ guarantees the asymptotic convergence and the bound is a function of μ. However, rigorous proof of this property for the APSA family is very involved mathematically because the weight adaptation formula (8) has nonlinear functions of the error in both the numerator and denominator. Attempt of theoretical analysis deserves a separate treatment. Here, numerical results for the excess MSE versus the step size μ are provided in Fig. 13 and Fig. 14 for the RP-APSA and RIP-APSA. The simulation results were obtained by time-averaging 500 instantaneous excess square

errors in the steady-state and then averaging 10 independent trials. The step size μ varied from 0.0005 to 10. For most simulations, we used the NIR with $\xi = 0.631$ for large filter length $L = 512$, while for small filter lengths, we used random dispersive NIRs with $\xi = 0.220$. Similar to the previous simulation, the input was AR(1) and the interference was a BG with $Pr = 0.001$. The normalized misalignment versus step size performed similarly as the excess MSE and is omitted for brevity.

The steady-state excess MSEs of the RP-APSA, as shown in Fig. 13, increased with the step size μ in a similar manner for different NIR lengths. With the same step size, a larger length L yielded higher excess MSEs, which was consistent with the results in [27]. The results for the RIP-APSA is similar to Fig. 13, thus omitted here for brevity. Other parameters, such as the projection order and proportionate regularization parameters, exhibited small impact on the steady-state excess MSE of the two proportionate APSAs, as shown in Fig. 14. For a given step size μ, both proportionate APSAs with a larger M achieved a slightly lower steady-state excess MSE when $\mu > \mu_t$ and a higher excess MSE when $\mu < \mu_t$, where $\mu_t = 0.002$ for the RP-APSA (Fig. 14a) and $\mu_t = 0.001$ for the RIP-APSA (Fig. 14b). The value μ_t might change slightly depending on the proportionate regularization parameters. Interestingly, the RP-APSA exhibited almost the same excess MSE for different projection orders at $\mu = 0.1$ and this μ might also shift slightly depending on the proportionate regularization parameters. For a given projection order M, the impact of the proportionate regularization parameters on the excess MSE were small, as shown in Fig. 14c and Fig. 14d, except that the excess MSE of the RP-APSA with $\rho = 0.01$ was a little lower than that of the other parameters. This is because the NIR of the echo path was dispersive with $\xi = 0.631$ and the recommended $\rho = 5/L$ according to [2].

Although no step sizes can cause asymptotic divergence, the steady-state excess MSE is related with the step size. This convergence behavior of the APSA family

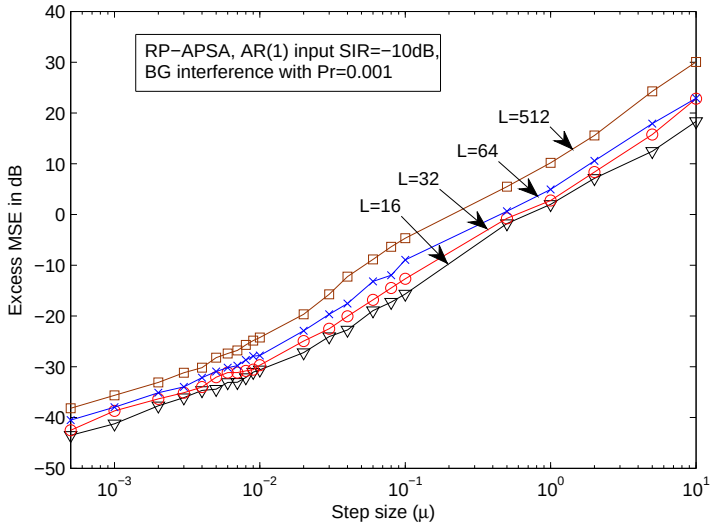
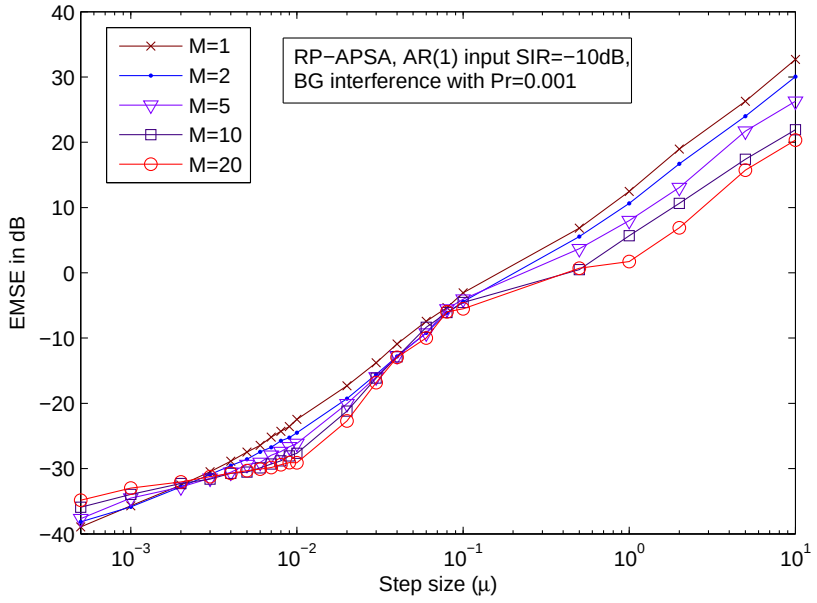
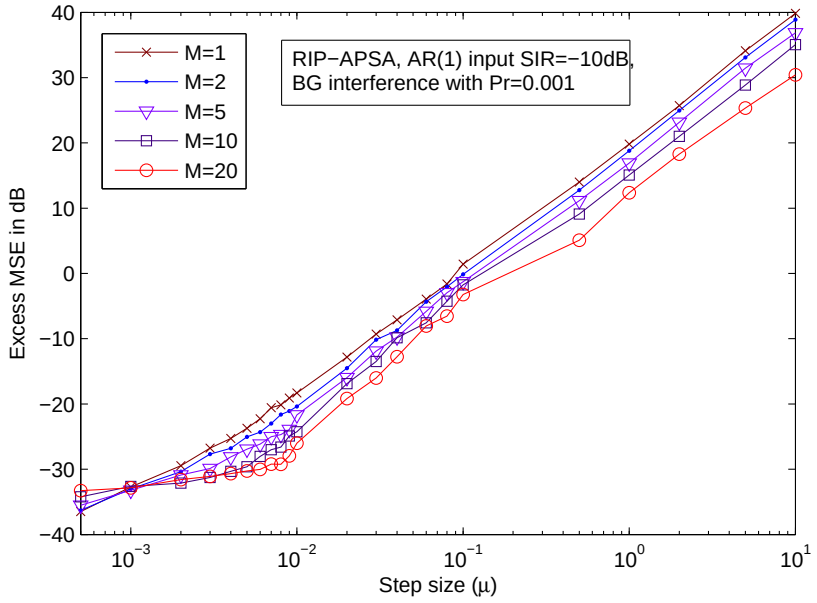


Figure 13. Steady-state excess MSE of the RP-APSA as a function of the step size μ. The input and the interference were the same as those in Fig. 4. The projection order $M = 2$ and proportionate regularization parameter $\rho = 0.01$. The impact of L for the RIP-APSA was similar to that of the RP-APSA, thus omitted for brevity.

will be illustrated by geometrical interpretation in Appendix. However, in the practical range of excess MSE less than 0 dB, the step size has to be very small for both the RP-APSA and RIP-APSA. The range of $0 < \mu < 0.1$ is recommended for practical NEC applications.

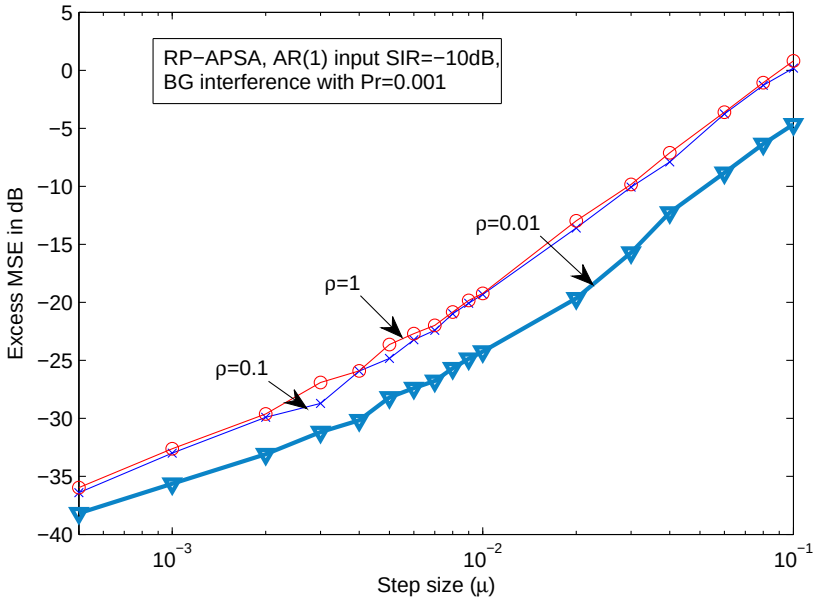


(a)

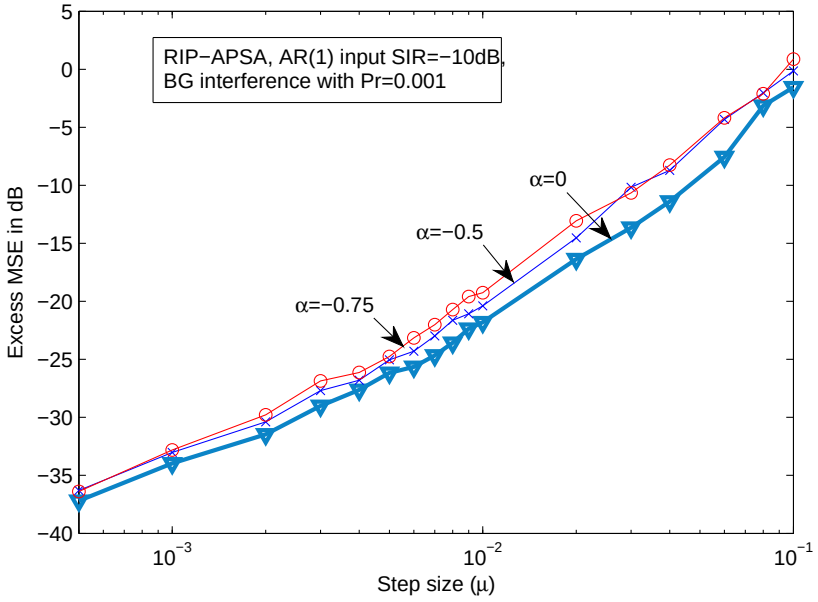


(b)

Figure 14. Steady-state excess MSE vs. the step size μ for the RP-APSA and RIP-APSA with varying projection order and proportionate parameters. The input and the interference were the same as those in Fig. 4. The filter length $L = 512$. (a) RP-APSA with varying M and $\rho = 0.01$. (b) RIP-APSA with varying M and $\alpha = -0.5$.



(c)



(d)

Figure 14. (c) RP-APSA with $M = 2$ and varying ρ. (d) RIP-APSA with $M = 2$ and varying α. (cont.)

5 CONCLUSION

Two proportionate affine projection sign algorithm (APSA) have been proposed for the identification of real-coefficient, sparse systems. With a modest increase in computational complexity over that of the original APSA, the proportionate APSAs can achieve faster convergence rate and lower in a steady-state misalignment in a sparse network echo path, colored input, and impulsive interference environment. The computational complexity of the two proportionate APSAs is lower than the APA family due to elimination of the matrix inversion. Especially, the RIP-APSA also exhibits good robustness in all types of NIRs without the need to change parameters or estimate the sparseness of the NIRs. Numerical results also demonstrate that the proportionate APSAs exhibit asymptotic convergence for all step size $\mu > 0$, but practically, a small step size of $\mu < 0.1$ is required to achieve excess MSE less than 0 dB.

6 APPENDIX: GEOMETRICAL ILLUSTRATION OF CONVERGENCE

We now demonstrate the convergence behavior of the proportionate APSA through geometrical illustration. First consider the NLMS algorithm

$$e(k) = y(k) - \mathbf{x}^T(k)\mathbf{w}(k) \tag{12}$$

and

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{x}(k) + \delta]^{-1}e(k). \tag{13}$$

When $\mu = 1$ and $\delta = 0$, the NLMS can be shown to be a projection algorithm. Inserting (12) into (13) with $\mu = 1$ and $\delta = 0$ yields

$$\begin{aligned} \mathbf{w}(k+1) &= \{\mathbf{I} - \mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{x}(k)]^{-1}\mathbf{x}^T(k)\}\mathbf{w}(k) \\ &\quad +\mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{x}(k)]^{-1}y(k). \end{aligned} \tag{14}$$

Recall that $y(k) = \mathbf{x}^T(k)\mathbf{h} + z(k) + v(k)$ and define the *projection matrix*,

$$\mathbf{P}(k) = \mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{x}(k)]^{-1}\mathbf{x}^T(k). \tag{15}$$

We can write (14) in terms of $\mathbf{P}(k)$ as

$$\begin{aligned} \mathbf{w}(k+1) &= [\mathbf{I} - \mathbf{P}(k)]\mathbf{w}(k) + \mathbf{P}(k)\mathbf{h} \\ &\quad +\mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{x}(k)]^{-1}[z(k) + v(k)]. \end{aligned} \tag{16}$$

Consider the case where there are no near-end signals, $z(k) + v(k) = 0$, then

$$\mathbf{w}(k+1) = [\mathbf{I} - \mathbf{P}(k)]\mathbf{w}(k) + \mathbf{P}(k)\mathbf{h}. \tag{17}$$

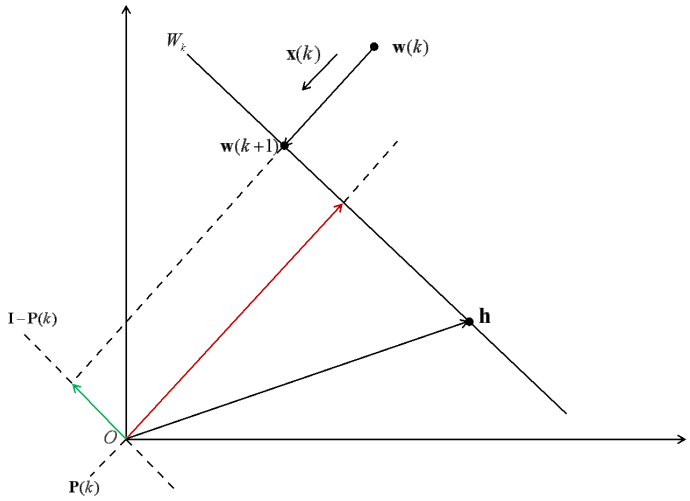


Figure 15. Two dimensional weight vector space observation for NLMS.

Since $[\mathbf{I}-\mathbf{P}(k)]$ is a projection matrix complementary to $\mathbf{P}(k)$, the coefficient update described by (17) is a classical *affine projection* of $\mathbf{w}(k)$ onto the *affine subspace* defined by $\mathbf{P}(k)$ and $\mathbf{h}$ to obtain the new coefficient vector $\mathbf{w}(k+1)$. This is shown in Fig. 15 where $\mathbf{w}(k+1)$ is found by projecting $\mathbf{w}(k)$ onto the subspace defined by $[\mathbf{I}-\mathbf{P}(k)]$ and then adding the projection of $\mathbf{h}$ onto the subspace defined by $\mathbf{P}(k)$. Let us define that particular affine subspace as W_k.

Note that the true echo path vector, $\mathbf{h}$, is always in the affine subspace W_k since replacing $\mathbf{w}(k)$ with $\mathbf{h}$ in (17) results in $\mathbf{w}(k+1)=\mathbf{h}$. We further note from (13) that the direction of the update is determined by the input vector $\mathbf{x}(k)$ and the sign of $e(k)$.

The vectors, $\mathbf{w}(k)$ through $\mathbf{w}(k+3)$, in Fig. 16a shows the coefficient update over several time periods. As the adaptive coefficient vector projects onto the subspaces W_k, W_{k+1}, and W_{k+2}, it keeps getting closer to the true coefficient vector $\mathbf{h}$. The dotted vectors, $\mathbf{w}'(k)$ through $\mathbf{w}'(k+3)$, indicate the progression of the coefficients when the step size μ is less than 1. These are called *relaxed* projections and the step size μ is also the relaxation parameter. In relaxed projections, the update

only goes part way to the affine subspace W_k. Clearly relaxed projections converge more slowly than pure projections.

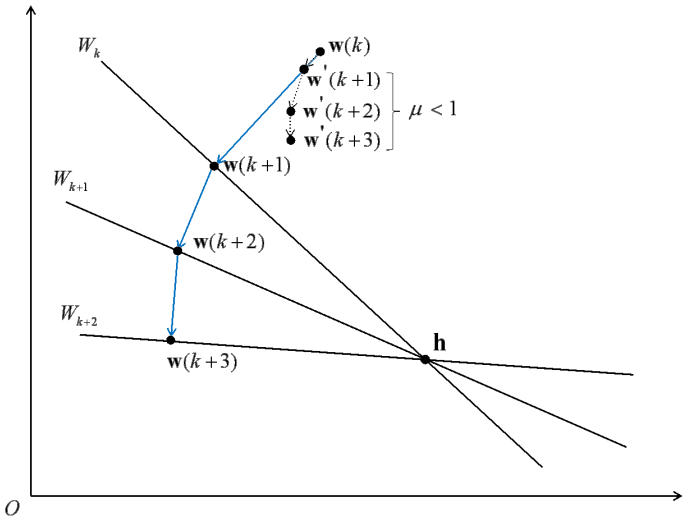
Now we consider the case when there are near-end signals, that is when $z(k)+v(k) \neq 0$. The last two terms of equation (16) are the near-end speech and background noise terms of the update. The direction of the update is determined by $\mathbf{x}(k)$ and the sign of $[z(k)+v(k)]$. This effect is geometrically illustrated in Fig. 16b, where the affine projection subspaces at the various sample periods are offset by the double-talk and background noise terms. The "noisy" affine subspaces, denoted $\hat{W}_k$, $\hat{W}_{k+1}$ and $\hat{W}_{k+2}$, are shown in dashed lines. The NLMS algorithm will project onto the noisy subspaces rather than the true ones. Unfortunately, the noisy subspaces no longer all meet at the true NIR vector, $\mathbf{h}$. However, they do "almost" intersect in an area whose size is a function of the near-end speech variance, σ_z^2, and the noise variance, σ_v^2. When the adaptive coefficients $\mathbf{w}(k)$'s get close to the true coefficients $\mathbf{h}$, they enter this zone of "almost intersection". The smaller the relaxation value μ is, the less the $\mathbf{w}(k)$'s "chase after" the noisy $\hat{W}_k$, and the closer to the true coefficient vector they remain. Thus, graphically, we have described the usual trade-off with the relaxation parameter. For large μ (close to one) there is fast convergence, but large steady-state mean squared error, and with small μ (closer to zero) we have slower convergence but lower steady-state mean squared error.

Now consider the PNLMS algorithm

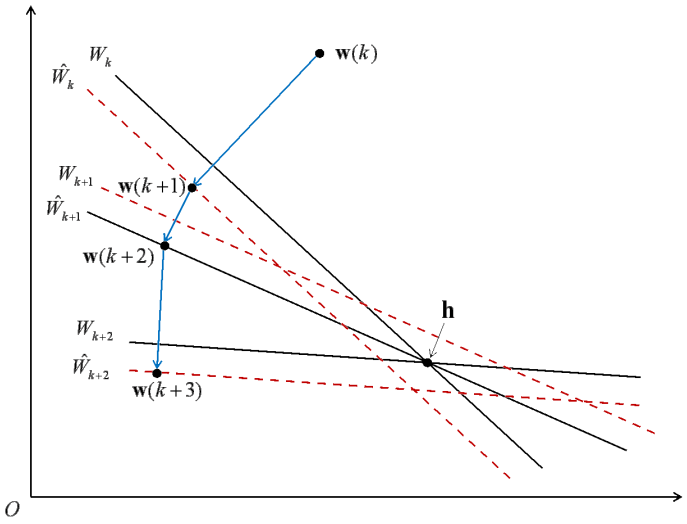
$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{G}(k)\mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \delta]^{-1} e(k). \tag{18}$$

This can also be shown to be a projection algorithm. To do this we make the change of variables,

$$\mathbf{w}(k) = \sqrt{\mathbf{G}(k)}\mathbf{b}(k) \tag{19}$$



(a)



(b)

Figure 16. Geometrical illustration of NLMS weight updating process. (a) $z(k) + v(k) = 0$. (b) $z(k) + v(k) \neq 0$.

where $\sqrt{\mathbf{G}(k)}$ =diag$\{\sqrt{g_0(k)}, \ldots, \sqrt{g_{L-1}(k)}\}$. Then (12) becomes

$$e(k) = y(k) - \mathbf{x}^T(k)\sqrt{\mathbf{G}(k)}\mathbf{b}(k) \tag{20}$$

and (18) becomes,

$$\begin{aligned}\sqrt{\mathbf{G}(k)}\mathbf{b}(k+1) &= \sqrt{\mathbf{G}(k)}\mathbf{b}(k) \\ &\quad +\mu\mathbf{G}(k)\mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)+\delta]^{-1}e(k). \qquad (21)\end{aligned}$$

Multiplying both sides from the left by $[\sqrt{\mathbf{G}(k)}]^{-1}$

$$\mathbf{b}(k+1) = \mathbf{b}(k) + \mu\sqrt{\mathbf{G}(k)}\mathbf{x}(k)[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)+\delta]^{-1}e(k). \qquad (22)$$

We can define

$$\mathbf{s}(k) = \sqrt{\mathbf{G}(k)}\mathbf{x}(k), \qquad (23)$$

then (20) and (22) become,

$$e(k) = y(k) - \mathbf{s}^T(k)\mathbf{b}(k) \qquad (24)$$

and

$$\mathbf{b}(k+1) = \mathbf{b}(k) + \mu\mathbf{s}(k)[\mathbf{s}^T(k)\mathbf{s}(k)+\delta]^{-1}e(k). \qquad (25)$$

Note that (24) and (25) have exactly the same form as (12) and (13) thus, under the same conditions as described for NLMS ($\mu = 1$ and $\delta = 0$), PNLMS is an affine projection algorithm with behavior similar to that of NLMS, but with a twist. With NLMS we saw that $\mathbf{h}$ was always in the affine subspaces defined in each sample period and that accordingly the convergence of the algorithm was toward the intersections of these affine subspaces, $\mathbf{h}$. With PNLMS, however, it is $[\sqrt{\mathbf{G}(k)}]^{-1}\mathbf{h}$ that is in each sample period's affine subspace. Since $\mathbf{G}(k)$ is time varying, the affine subspaces will not all intersect at a single point until $\mathbf{w}(k)$ (or equivalently, $\mathbf{b}(k)$) reaches steady state. In fact, since $[\sqrt{\mathbf{G}(k)}]^{-1}\mathbf{h}$ is a moving target, it speeds the convergence of PNLMS for sparse $\mathbf{h}$'s in that it tends to keep the $\mathbf{w}(k)$'s sparse and

thus searches out a smaller space than otherwise the NLMS would search the entire L-dimensional space.

Now consider the new proportionate APSA algorithm with the coefficient update for projection order $M = 1$ being

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{x}_{gs}(k)[\mathbf{x}_{gs}^T(k)\mathbf{x}_{gs}(k) + \delta]^{-1/2} \tag{26}$$

where

$$\mathbf{x}_{gs}(k) = \mathbf{G}(k)\mathbf{x}(k)\text{sgn}[e(k)]. \tag{27}$$

Unlike the NLMS and PNLMS cases, this coefficient update cannot be manipulated into the form of an affine projection operation. However, the update still has some interesting properties. One is that the term in the denominator of (26) does not in general converge to zero as $\mathbf{w}(k)$ approaches $\mathbf{h}$ because of the sign operator in (27). That is, $\text{sgn}[e(k)] = \pm 1$ despite convergence. The only time $\mathbf{x}_{gs}^T(k)\mathbf{x}_{gs}(k)$ gets small is when $\mathbf{G}(k)\mathbf{x}(k)$ has a small magnitude. If $\mathbf{G}(k)\mathbf{x}(k)$ has a large enough magnitude that we may ignore the small positive number δ, then the coefficient updating equation becomes

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{x}_{gs}(k)/\|\mathbf{x}_{gs}(k)\|_2. \tag{28}$$

Note that the update vector, $\mu \mathbf{x}_{gs}(k)/\|\mathbf{x}_{gs}(k)\|_2$, always has a magnitude of μ. The direction of the coefficient update is always given by $\mathbf{x}_{gs}(k)$.

Though (26) through (28) do not represent an affine projection, we may consider the update as a relaxed affine projection if μ is sufficiently small. The relaxation factor in this case is not equal to the step size μ and it changes every sample period. The length of the update remains constant each sample period. In contrast, with NLMS and PNLMS, the relaxation factor was constant, but the length of the update generally change each sample period.

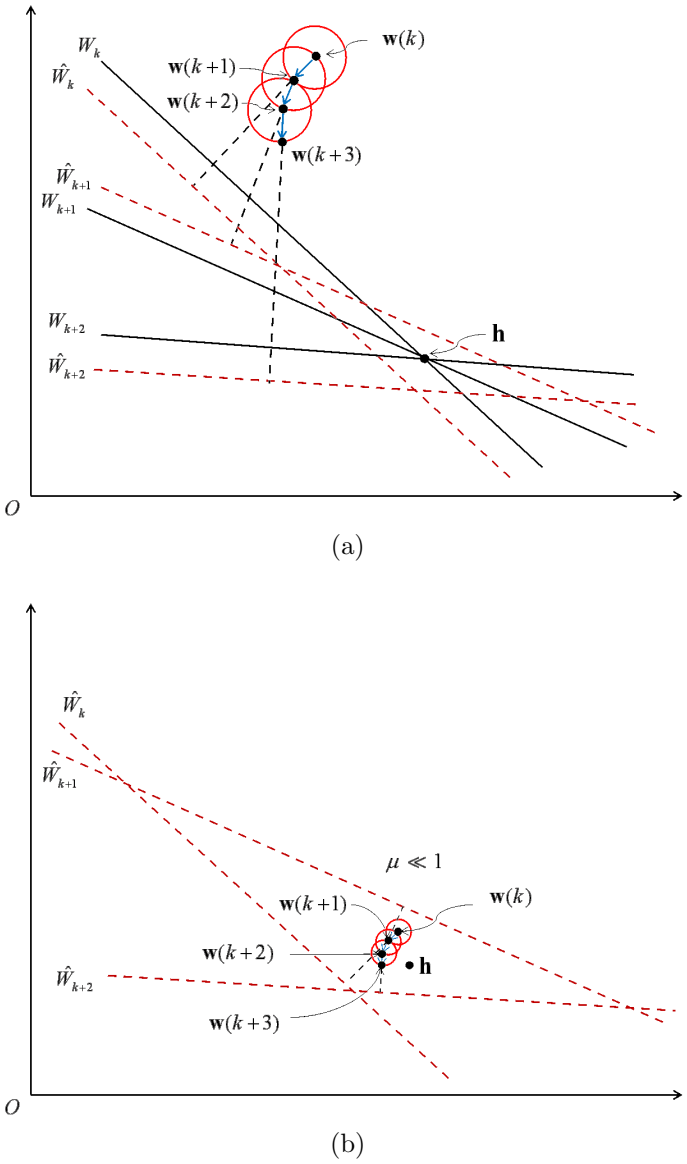


Figure 17. Geometrical illustration of APSA weight updating process. (a) Large μ, far from steady state. (b) Small μ, closed to steady state.

Figure 17a shows the update of two sample periods using the new algorithms when the $\mathbf{w}(k)$s are far from their steady-state values. The size of the update vector is always μ regardless of how far the coefficient vector is from the affine subspace of the projection. Fig. 17b shows the case where the adaptive coefficients are in the "almost intersection" region. Since the updates are always the same size, the coefficients are

less likely to "chase after" noisy $\hat{W}_k$s than conventional l_2-norm algorithms. This accounts for the algorithm's good steady-state performance and low minimum mean squared error.

7 ACKNOWLEDGEMENT

The authors wish to thank Dr. Jacob Benesty for suggesting the extension of APSA to proportionate APSA. The authors would also like to thank Ms. Tiange Shao for her help in the initial stage of the algorithm development.

8 REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, New Jersey, 07458.: MPrentice Hall, 2002.

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

[3] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Signals, Syst. Comput., 1998. 32th Asilomar Conf.*, vol. 1, 1-4 1998, pp. 394 –398.

[4] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, 2002. (ICASSP '02)*, vol. 2, 2002, pp. 1881 –1884.

[5] P. Loganathan, A. Khong, and P. Naylor, "A class of sparseness-controlled algorithms for echo cancellation," *IEEE Trans. Audio, Speech, Language Processing*, vol. 17, no. 8, pp. 1591 –1601, Nov. 2009.

[6] J. Arenas-Garcia and A. Figueiras-Vidal, "Adaptive combination of proportionate filters for sparse echo cancellation," *IEEE Trans. Audio, Speech, Language Processing*, vol. 17, no. 6, pp. 1087 –1098, Aug. 2009.

[7] 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, Mar. 2005.

[8] K. Sakhnov, "An improved proportionate affine projection algorithm for network echo cancellation," in *Syst., Signals and Image Processing, 2008. IWSSIP 2008. 15th International Conference*, 25-28 2008, pp. 125 –128.

[9] C. Paleologu, S. Ciochina, and J. Benesty, "An efficient proportionate affine projection algorithm for echo cancellation," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 165 –168, Feb. 2010.

[10] T. Gansler, S. Gay, M. Sondhi, and J. Benesty, "Double-talk robust fast converging algorithms for network echo cancellation," *IEEE Trans. Speech, Audio Processing*, vol. 8, no. 6, pp. 656 –663, Nov. 2000.

[11] J. Benesty, H. Rey, L. Vega, and S. Tressens, "A nonparametric VSS NLMS algorithm," *IEEE Signal Processing Lett.*, vol. 13, no. 10, pp. 581 –584, Oct. 2006.

[12] C. Paleologu, J. Benesty, and S. Ciochina, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Processing*, vol. 16, no. 8, pp. 1466 –1478, Nov. 2008.

[13] T. Shao, Y. R. Zheng, and J. Benesty, "An affine projection sign algorithm robust against impulsive interferences," *IEEE Signal Processing Lett.*, vol. 17, no. 4, pp. 327–330, Apr. 2010.

[14] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO Signal Processing (Signals and Communication Technology Series)*. Springer, 2006.

[15] L. R. Vega, H. Rey, J. Benesty, and S. Tressens, "A new robust variable step-size NLMS algorithm," *IEEE Trans. Signal Processing*, vol. 56, no. 5, pp. 1878 –1893, May 2008.

[16] D. Duttweiler, "A twelve-channel digital echo canceler," *IEEE Trans. Communications*, vol. 26, no. 5, pp. 647 – 653, May 1978.

[17] H.-C. Shin and A. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 90–102, Jan. 2004.

[18] N. Bershad, "Analysis of the normalized LMS algorithm with gaussian inputs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 4, pp. 793 – 806, Aug. 1986.

[19] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Processing*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[20] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters with error nonlinearities," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 653–663, Mar. 2003.

[21] S. G. Sankaran and A. A. L. Beex, "Convergence behavior of affine projection algorithms," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1086–1096, Apr. 2000.

[22] V. Mathews and S. C., "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 4, pp. 450–454, Apr. 1987.

[23] V. J. Mathews, "Performance analysis of adaptive filters equipped with the dual sign algorithm," *IEEE Trans. Signal Processing*, vol. 39, no. 1, pp. 85–91, Jan. 1991.

[24] S. C. B., S. Ann, and I. Song, "Performance analysis of the dual sign algorithm for additive contaminated-gaussian noise," *IEEE Signal Processing Lett.*, vol. 1, no. 12, pp. 196–198, Dec. 1994.

[25] E. Masry and F. Bullo, "Convergence analysis of the sign algorithm for adaptive filtering," *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 489–495, Mar. 1995.

[26] J. A. Bucklew, T. G. Kurtz, and W. A. Sethares, "Weak convergence and local stability properties of fixed step size recursive algorithms," *IEEE Trans. Information Theory*, vol. 39, no. 3, pp. 966 –978, May 1993.

[27] E. Eweda, "Convergence analysis of the sign algorithm without the independence and gaussian assumptions," *IEEE Trans. Signal Processing*, vol. 48, no. 9, pp. 2535 –2544, Sep. 2000.

[28] H.-F. Chen and G. Yin, "Asymptotic properties of sign algorithms for adaptive filtering," *IEEE Trans. Automatic Control*, vol. 48, no. 9, pp. 1545 – 1556, Sep. 2003.

II. A COMPARATIVE STUDY OF COMPRESSED SENSING APPROACHES FOR 3-D SYNTHETIC APERTURE RADAR IMAGE RECONSTRUCTION

Zengli Yang, and Yahong Rosa Zheng

ABSTRACT—This paper investigates two compressed sensing (CS) approaches that can be used to reconstruct 3-D synthetic aperture radar (SAR) images with undersampled measurements. Combining CS with the range migration algorithm (RMA), using either Stolt transform or non-uniform fast Fourier transform (NUFFT), yields two different approaches: Stolt-CS and NUFFT-CS. These approaches can decrease the load of data acquisition while recovering satisfactory 3-D SAR images through l_1-norm optimization. A simulated image is used as the ground truth to facilitate the comparative study. The 2-D structured similarity (SSIM) index is extended to 3-D to assess the quality of the reconstructed images. Both the simulation and the experimental reconstruction results demonstrate that the Stolt-CS contributes little to image quality improvement or computational complexity reduction due to the inaccuracy of the Stolt transform. In contrast, the NUFFT-CS achieves a good tradeoff between the reconstruction quality and the computational costs.

1 INTRODUCTION

Wideband 3-D synthetic aperture radar (SAR) imaging has important applications in the area of nondestructive testing and evaluation (NDT&E) [1]. This is due to its feasibility to acquire high-resolution holographic images of specimen under test (SUT). Microwave and millimeter wave as interrogating signals can not only penetrate dielectric materials but also interact with their inner structure to render a comprehensive image for inspection. Microwave and millimeter wave NDT&E techniques have been applied to diverse applications (i.e., the detection and evaluation of corrosion under paint and composite laminates, the detection and sizing of fatigue cracks in metal surfaces, and the characterization of dielectric material).

Several high-resolution 3-D SAR imaging systems were developed at the Applied Microwave Nondestructive Testing Laboratory (AMNTL) at Missouri University of Science and Technology (Missouri S&T). The imaging system uses stepped frequencies in the 20–60 GHz band and raster/rotary scanners with step size on the order of a millimeter. These systems can achieve a volumetric resolution on the order of millimeter. They can also assist in the detection and quantification of either small flaws or targets in a specimen under test (SUT). However, the speed of data acquisition for these wideband 3-D SAR imaging systems hinders their practical applications due to the slowness of mechanical scanning. For example, uniform raster scanning requires approximately one hour to scan a 120×180 mm^2 area at 2 mm spacing. In contrast, random undersampling can reduce the acquisition time when the antenna probe is placed at only a fractional number of positions on the uniform grid. Advanced sparse methods are required to reconstruct images when undersampled measurements are used. Our experiments have demonstrated that sampling 20% – 40% of the original uniform grid can reduce the data acquisition time by 70% – 60% [2, 3, 4]. Note that

the percentage of the measurement points kept is referred to as the undersampling rate. Meanwhile, the reconstructed images can achieve a quality that is comparable to those reconstructed from full sampling when the compressed sensing (CS) approach is used [6, 7, 8, 9, 5, 10].

In the literature [5] [11] [12], CS has been applied for remote sensing with 3-D SAR imaging where fewer 3-D Fourier (k-space) measurements were obtained at the airborne radar from multiple elevation passes. Unlike our CS problem formulation, the works in [5] [11] [12] use the $\omega - k$ algorithm [13] with nearest-neighbor Stolt interpolation as a pre-processing step. After the one-time interpolation, the raw measurements were firstly converted into a uniformly sampled k-space data, in which an iterative recovery algorithm is applied to reconstruct the 3-D images. This method results in computational savings but the quality of the reconstructed image is compromised. For our CS application, the backscatter data are acquired on a planar aperture, and one-time preprocessing yields unsatisfactory image quality. Therefore, we focus on involving the complete nonlinear forward and reverse SAR transforms during the reconstruction process via l_1 optimization. The procedure needed to convert the raw measurements into the 3-D image is known as forward SAR transform, the reverse procedure is known as reverse SAR transform [10]. The CS approach applies both the forward and the reverse SAR transforms during each iteration of the minimization. The forward SAR transform is used for image reconstruction, and the reverse SAR transform is used to generate the estimated measurement from the reconstructed image. The CS approach emphasizes the consistency between the estimated measurement and the actual raw measurement. Thus, it is a truly underdetermined, compressed sensing system which exhibits a larger computational cost than the one-time preprocessing approach due to the complicated nonlinear SAR transforms in each iteration.

When the backscatter data are acquired on a planar aperture, the conventional 3-D SAR transform will typically adopt the range migration algorithm (RMA) [14] with either Stolt transform [15] or non-uniform fast Fourier transform (NUFFT) [16] [17]. In this paper, the two conventional RMAs are referred to as Stolt-RMA and NUFFT-RMA, respectively. The image recovered by the NUFFT-RMA usually has fewer artifacts and better resolution than that using the Stolt-RMA. However, the NUFFT-RMA has a much higher computational cost than does the Stolt-RMA. With undersampled measurements, the CS principle can be combined with the two RMAs, thus yielding two different approaches. The two CS approaches are denoted here as Stolt-CS and NUFFT-CS, respectively. Previous works investigated these approaches using experimental data, demonstrating their performance separately [2, 3, 10, 4]. Therefore, comparing the overall tradeoff between the image quality and computational costs for the two approaches is quite interesting.

In this paper, the performance of the two CS approaches are first compared using a synthesized SUT and its simulated measurements, and then by experimental measurements of real SUT and millimeter wave imaging system. Choosing simulation provides access to the ground truth image, thus enabling a fair comparison. Meanwhile, experimental results were provided to verify the characteristics of the two CS approaches in practice. The results of the experimental system verified that the CS approaches can indeed save data acquisition time while achieving satisfactory image quality. The structured similarity (SSIM) [18] index rather than the peak signal-to-noise ratio (PSNR) or mean square error (MSE) was adopted to evaluate the image qualities, because SSIM is considered to be more consistent with human eye perception. The original 2-D SSIM discussed in [18] was extended into a 3-D SSIM in this paper to assess the similarity between the reconstructed images and the ground truth image. The 2-D split Bregman framework [19] [20] was also extended here to 3-D to solve the l_1-regularized image reconstruction problems. The split Bregman algorithm

was adopted because it can achieve faster convergence and better numerical stability than traditional continuation methods [21] [22]. Both the numerical simulations and the experimental results demonstrate that the Stolt-CS approach yields low image quality, albeit with a low computational complexity. In contrast, the NUFFT-CS exhibited a higher image quality than the Stolt-CS at all undersampling rates. The overall computational complexity of the NUFFT-CS is affordable for an undersampling rate that is greater than 25%.

2 CONVENTIONAL 3-D SAR IMAGE RECONSTRUCTION WITH FULL SAMPLING

Consider a wideband monostatic stripmap 3-D SAR imaging system in which raw data is collected by a probe that scans the SUT over a 2-D plane in the 3-D Cartesian space. The dimension perpendicular to the 2-D XY plane is defined as the Z-dimension; its positive direction points away from the surface of the SUT. A point target is characterized by its reflectivity function $g(x', y', z')$. Because the probe is within the near-field of the SUT, the wavefront curvature is no longer negligible. The received spherical waveform at position (x, y), with a temporal angular frequency of ω, is then given by

$$\begin{aligned} r(x, y, \omega) &= \int\int\int g(x', y', z') \\ &\quad \times e^{-j2k\sqrt{(x-x')^2+(y-y')^2+z'^2}}\, \mathrm{d}x'\, \mathrm{d}y'\, \mathrm{d}z' \end{aligned} \tag{1}$$

where $j = \sqrt{-1}$, $\omega = 2\pi f$ (with f being the temporal frequency), and $k = \omega/c$ is the wavenumber (with c being the propagation speed of the microwave). By using the plane wave decomposition for spherical wave and Stolt transform [15], the 3-D reflectivity function of the SUT is given by [23]

$$g(x, y, z) = \mathcal{F}_{3\mathrm{D}}^{-1}\left\{\boldsymbol{\Theta}\{\mathcal{F}_{2\mathrm{D}}\left[r(x, y, \omega)\right] e^{-jk_z z}\}\right\}, \tag{2}$$

which is known as either the 3-D Stolt-RMA [14] or the $\omega - k$ algorithm [13] (in its 2-D version). Here, $\mathcal{F}_{2\mathrm{D}}$ denotes a 2-D cross-range fast Fourier transform (FFT) along the XY-plane, $\mathcal{F}_{3\mathrm{D}}^{-1}$ denotes the 3-D inverse FFT (IFFT), and $\boldsymbol{\Theta}$ denotes the 1-D Stolt transform with the nearest neighbor interpolation. Note that the distinction

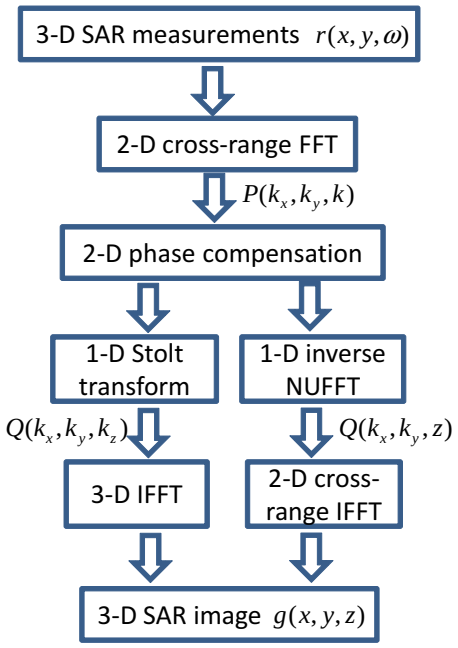


Figure 1. Conventional 3-D SAR image reconstruction methods: Stolt-RMA (on the left) and NUFFT-RMA (on the right). The 2-D phase compensation step accounts for the $e^{-jk_z z}$ term and is used to adjust the focal plane.

between the primed and the unprimed coordinate systems is now dropped because the coordinate systems coincide after the FFT and IFFT operations.

Let $P(k_x, k_y, k)$ represent the 2-D cross-range Fourier transform of $r(x, y, \omega)$, and $Q(k_x, k_y, k_z)$ represent the 3-D Fourier transform of the reflectivity function $g(x, y, z)$, where k_x, k_y, and k_z are the wavenumbers in the x, y, and z dimensions, respectively. Curvature in near-field wave propagation creates $k_z = \sqrt{4k^2 - k_x^2 - k_y^2}$. The direct interpolation is known as Stolt transform $\mathbf{\Theta}$ to obtain equispaced $Q(k_x, k_y, k_z)$ in k_z for the 3-D IFFT operation. Alternatively, nonuniform spaced $Q(k_x, k_y, k_z)$ in k_z can be transformed to the reflectivity image by applying nonuniform FFT (NUFFT) [16] [17],

$$g(x, y, z) = \mathcal{F}_{\text{2D}}^{-1} \left\{ \mathcal{F}_{\text{NUFFT}}^{-1} \{ \mathcal{F}_{\text{2D}} \left[r(x, y, \omega) \right] e^{-jk_z z} \} \right\}, \tag{3}$$

which is known as the 3-D NUFFT-RMA [17]. Here, $\mathcal{F}_{\mathrm{2D}}^{-1}$ denotes the 2-D cross-range IFFT, and $\mathcal{F}_{\mathrm{NUFFT}}^{-1}$ is the 1-D inverse NUFFT along k_z. With Gaussian-kernel interpolation, the NUFFT-RMA achieves better image quality at a higher computational cost than does the Stolt-RMA. A summary of the complete procedures used for the conventional, fully-sampled 3-D SAR image reconstruction methods is given in Fig. 1.

3 COMPRESSED SENSING APPROACHES FOR 3-D SAR IMAGE RECONSTRUCTION

To take the advantage of the CS approach for 3-D SAR imaging, the radar probe can measure a small percentage of randomly-selected positions on the uniform XY grid. The backscatter data at these positions are collected and saved as the raw data, which is referred to as the undersampled measurements. The ratio of the number of undersampled points over the number of fully sampled points is defined as the undersampling rate α. The CS approaches are used to reconstruct 3-D SAR images from the raw measurement $r(x, y, \omega)$. This approach exploits the sparsity of 3-D SAR images, relying on both the l_1-norm and total variation (TV) minimizations [19]. Either the Stolt-RMA or NUFFT-RMA is applied to yield two different CS approaches: Stolt-CS and NUFFT-CS, respectively.

3.1 TWO COMPRESSED SENSING APPROACHES

The CS approach emphasizes the consistency between the estimated measurements and the actual gathered raw measurements during reconstruction. For 3-D SAR imaging, the CS approach can be interpreted as [4]

$$\min_{\hat{\mathbf{g}}} J(\hat{\mathbf{g}}) \quad \text{subject to} \quad \|\mathbf{\Phi}\hat{\mathbf{g}} - \mathbf{r}\|_2^2 < \sigma^2, \tag{4}$$

where $\|\cdot\|_2$ is the l_2 norm, $\hat{\mathbf{g}} \in \mathbb{C}^{N\times 1}$ is the vectorized estimated 3-D SAR image, σ^2 is the noise variance, and $\mathbf{\Phi} \in \mathbb{C}^{M\times N}$ $(M < N)$ is the measurement matrix that reflects the acquisition of the vectorized raw measurements $\mathbf{r} \in \mathbb{C}^{M\times 1}$. For the Stolt-CS and NUFFT-CS, $\mathbf{\Phi}$ is the reverse Stolt-RMA and NUFFT-RMA, respectively. According

to (2) and (3), the measurement operator $\mathbf{\Phi}$ can be written as

$$\mathbf{\Phi}_{\text{Stolt}} = \mathbf{U}\mathcal{F}_{\text{2D}}^{-1}\left\{\mathbf{\Theta}^{\dagger}[\mathcal{F}_{\text{3D}}(\cdot)]\right\}, \tag{5}$$

$$\mathbf{\Phi}_{\text{NUFFT}} = \mathbf{U}\mathcal{F}_{\text{2D}}^{-1}\left\{\mathcal{F}_{\text{NUFFT}}[\mathcal{F}_{\text{2D}}(\cdot)]\right\}, \tag{6}$$

Note that the phase compensation term is omitted for brevity. Here, $\mathbf{\Theta}^{\dagger}$ represents the pseudoinverse 1-D Stolt transform with the nearest neighbor interpolation, $\mathcal{F}_{\text{3D}}$ denotes the 3-D FFT, and $\mathbf{U}$ denotes the binary matrix that is used to select the random (x, y) positions for random undersampling.

The cost function $J(\hat{\mathbf{g}})$ represents some l_1 regularization term with respect to $\hat{\mathbf{g}}$. In this paper, it is selected as

$$J(\hat{\mathbf{g}}) = \frac{\gamma_2}{2}\|\mathbf{\Psi}\hat{\mathbf{g}}\|_1 + \frac{\lambda}{2}\|\mathbf{D}\hat{\mathbf{g}}\|_1, \tag{7}$$

where $\|\cdot\|_1$ denotes the l_1 norm, $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ is the linear operator that transforms the image from voxel representation into a sparse representation, $\mathbf{D}$ is the discrete 3-D isotropic TV operator, and γ_2 and λ are the weights for the consistency of the l_1 norm and the TV norm, respectively.

Note that this paper uses the split Bregman algorithm, as detailed in Section 3.2, which considers the complex nature of the image and updates the real and imaginary parts jointly in the inner iteration of the CS solver. Therefore, the TV can be successfully applied with the l_1 optimization.

3.2 SPLIT BREGMAN FRAMEWORK FOR 3-D IMAGE RECONSTRUCTION

The split Bregman framework [19,20] is now extended to the 3-D image reconstruction, so that (4) can be solved. The split Bregman algorithms, when compared

to traditional continuation methods [21, 22], can achieve faster convergence and better numerical stability during reconstruction. Equation (4) can be transformed as a sequence of unconstrained problems [19]:

$$\hat{\mathbf{g}}_{i+1} = \arg\min_{\hat{\mathbf{g}}_i} \frac{\gamma_1}{2}\|\mathbf{\Phi}\hat{\mathbf{g}}_i - \mathbf{b}_i\|_2^2 + \frac{\gamma_2}{2}\|\mathbf{\Psi}\hat{\mathbf{g}}_i\|_1 + \frac{\lambda}{2}\|\mathbf{D}\hat{\mathbf{g}}_i\|_1, \quad (8)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_i + \mathbf{r} - \mathbf{\Phi}\hat{\mathbf{g}}_{\mathbf{i+1}}, \quad (9)$$

where i denotes the iteration number of the outer loop, and γ_1 denotes the regularization parameters that determine the trade-off between measurement consistency and sparsity in the $\mathbf{\Psi}$ domain and the finite difference domain. The iterative shrinkage methods [24, 25] provide an efficient way to solve the linear inverse problem with one l_1-norm constraint. Two l_1-norm terms, however, are involved in (8). Therefore, (8) is further relaxed to prepare for further splitting of the two l_1-norm terms:

$$\min_{\hat{\mathbf{g}},\mathbf{d}_x,\mathbf{d}_y,\mathbf{d}_z,\mathbf{w}} \frac{\gamma_1}{2}\|\mathbf{\Phi}\hat{\mathbf{g}} - \mathbf{b}\|_2^2 + \frac{\gamma_2}{2}\|\mathbf{w} - \mathbf{\Psi}\hat{\mathbf{g}} - \mathbf{b}_w\|_2^2 + \|\mathbf{w}\|_1 + \frac{\lambda}{2}\big[\|\mathbf{d}_x - \mathbf{D}_x\hat{\mathbf{g}} - \mathbf{b}_x\|_2^2 + \|\mathbf{d}_y - \mathbf{D}_y\hat{\mathbf{g}} - \mathbf{b}_y\|_2^2 + \|\mathbf{d}_z - \mathbf{D}_z\hat{\mathbf{g}} - \mathbf{b}_z\|_2^2\big] + \|\mathbf{D}\mathbf{d}\|_1, \quad (10)$$

where the subscript i is omitted for brevity, $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z, \mathbf{d}, \mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z, \mathbf{b}_\omega$ are the auxiliary variables with $d = (d_x, d_y, d_z)$ (given the elements of $\mathbf{d}, \mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z$ are d, d_x, d_y, d_z, respectively). Moreover, $\mathbf{D}_x, \mathbf{D}_y, \mathbf{D}_z$ represents the 1-D discrete derivative operator in the X, Y, Z dimension, respectively. The goal is to find the optimal solution pair $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z$, $\mathbf{w}$, and $\hat{\mathbf{g}}$ to (10) in sequence by solving their corresponding subproblems. For example, since $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z$ are only involved in the fourth, fifth, sixth, and seventh terms in (10), it can be formulated as a standard linear inverse problem as shown in (16) in the Appendix. The iterative shrinkage method can then be applied, so as the way to find the solution $\mathbf{w}$. The procedures needed to identify the optimal $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z, \mathbf{w}$ are described in detail in the Appendix.

Once the solution $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z$, and $\mathbf{w}$ to (10) is found, the optimal $\hat{\mathbf{g}}$ can be found by solving the optimization subproblem of (10) as related to $\hat{\mathbf{g}}$. That is,

$$\hat{\mathbf{g}} = \arg\min_{\hat{\mathbf{g}}} \frac{\lambda}{2} \left[\|\mathbf{d}_x - \mathbf{D}_x\hat{\mathbf{g}} - \mathbf{b}_x\|_2^2 + \|\mathbf{d}_y - \mathbf{D}_y\hat{\mathbf{g}} - \mathbf{b}_y\|_2^2 + \|\mathbf{d}_z - \mathbf{D}_z\mathbf{g} - \mathbf{b}_z\|_2^2 \right] + \frac{\gamma_1}{2}\|\mathbf{\Phi}\hat{\mathbf{g}} - \mathbf{b}\|_2^2 + \frac{\gamma_2}{2}\|\mathbf{w} - \mathbf{\Psi}\hat{\mathbf{g}} - \mathbf{b}_w\|_2^2. \quad (11)$$

Note that (11) has only l_2-norm terms, which are differentiable. Thus the solution ($\hat{\mathbf{g}}$) can be found by apply the nonlinear conjugate gradient (CG) algorithm [26]. Setting the first-order derivative of (11) with respect to $\hat{\mathbf{g}}$ as zero produces

$$\mathbf{B}\hat{\mathbf{g}} = \lambda\big[\mathbf{D}_x^T(\mathbf{d}_x - \mathbf{b}_x) + \mathbf{D}_y^T(\mathbf{d}_y - \mathbf{b}_y) + \mathbf{D}_z^T(\mathbf{d}_z - \mathbf{b}_z)\big] + \gamma_1\mathbf{\Phi}^H\mathbf{b} + \gamma_2\mathbf{\Psi}^H(\mathbf{w} - \mathbf{b}_w) \quad (12)$$

where $(\cdot)^T$ represents the transpose, $(\cdot)^H$ represents Hermitian transpose, and

$$\mathbf{B} = \lambda\left(\mathbf{D}_x^T\mathbf{D}_x + \mathbf{D}_y^T\mathbf{D}_y + \mathbf{D}_z^T\mathbf{D}_z\right) + \gamma_1\mathbf{\Phi}^H\mathbf{\Phi} + \gamma_2\mathbf{\Psi}^H\mathbf{\Psi}. \quad (13)$$

Because $\mathbf{B}$ is a large, symmetric, and positive definite matrix, the CG algorithm can be applied to solve (12) efficiently.

Now the solution $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z$, $\mathbf{w}$, and $\hat{\mathbf{g}}$ tin (10) is available for the current iteration. The $\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z, \mathbf{b}_w$ in (10) is updated to make the $\hat{\mathbf{g}}$ converge to the solution in (8), and (10) is solved for a number of iterations (identified here as the inner loop). The procedures used to update $\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z, \mathbf{b}_w$ in (10) for each inner iteration are described in the Appendix. After the inner loop, the suboptimal solution $\hat{\mathbf{g}}_i$ in (8) is obtained, and the outer loop is used to solve the original constrained problem (4). The split Bregman algorithm for the sparse 3-D SAR image reconstruction is summarized in Algorithm 1.

Algorithm 1: Split Bregman Algorithm for sparse 3-D SAR Image Reconstruction

Data: maximum outer iterations N_{outer}
maximum inner iterations N_{inner}
tolerance η

begin
- $\hat{\mathbf{g}} = \mathbf{\Phi}^{-1}\mathbf{r}$
- $\mathbf{b} = \mathbf{r}$
- $\mathbf{b}_x = \mathbf{b}_y = \mathbf{b}_z = \mathbf{b}_w = \mathbf{0}$
- $i = n = 0$
- **while** $\varepsilon = \frac{\|\mathbf{\Phi}\hat{\mathbf{g}} - \mathbf{r}\|_2^2}{\|\mathbf{r}\|_2^2} > \eta$ *or* $i \leq N_{\text{outer}}$ **do**
  - **while** $n \leq N_{\text{inner}}$ **do**
    - Compute the optimal $\hat{\mathbf{g}}$ in (11) by solving (12) with the nonlinear CG algorithm
    - Use (17a) and (18) in the Appendix to compute $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z, \mathbf{w}$
    - Update $\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z, \mathbf{b}_w$ according to (19a) in the Appendix;
    - $n \leftarrow n + 1$
  - **end**
  - $\mathbf{b} \leftarrow \mathbf{b} + \mathbf{r} - \mathbf{\Phi}\hat{\mathbf{g}}$
  - $i \leftarrow i + 1$
- **end**

end

For the image denoising approach, we have $\mathbf{A} = \mathbf{I}_{N\times N}$ and $\mathbf{\Psi}$ is an orthogonal transform. Equation (13) can be simplified as

$$\mathbf{B} = \lambda\left(\mathbf{D}_x^T\mathbf{D}_x + \mathbf{D}_y^T\mathbf{D}_y + \mathbf{D}_z^T\mathbf{D}_z\right) + (\gamma_1 + \gamma_2)\mathbf{I}. \tag{14}$$

4 SIMULATIONS AND EXPERIMENTAL RESULTS

Our previous publications [2, 3, 10, 4] reported experimental results that used undersampled raw data that was measured by the 3-D SAR imaging system. Good image qualities were demonstrated with 30 – 40 % randomly undersampled spatial points. The qualities were compared with the reconstructed image from the fully sampled raw data using the NUFFT-RMA algorithm. Due to limitations of the RMA algorithm, the reconstructed image from 100% spatial points may still deviate significantly from the ground truth. Therefore, in this paper, several targets are simulated as the ground truth image. According to the simulated targets, the raw measurement data is generated. The reconstructed images recovered from either the simulated or the experimental undersampled raw data are compared to demonstrate the differences between the two CS approaches.

4.1 IMAGE QUALITY METRICS AND PARAMETER SELECTIONS

The SSIM [18] is commonly used for 2-D image quality assessment.It is considered more consistent with human eye perception than either the PSNR or the MSE. Hence, the mean SSIM (MSSIM) measure is extended to 3-D images by averaging the SSIM over 3-D windows instead of 2-D windows. The specific form of the SSIM index between signals $\mathbf{x}$ and $\mathbf{y}$ is defined as [18]

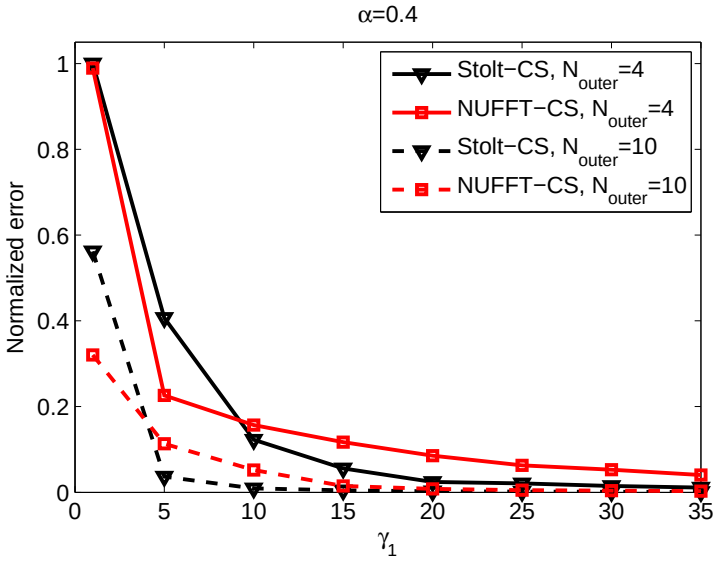
$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{15}$$

where C_1 and C_2 are the auxiliary variables, and $\mu_x(\mu_y), \sigma_x^2(\sigma_y^2)$ and σ_{xy} are the weighted mean, variance, and covariance, respectively, which are computed locally in a cubic window with a weighting function. The cubic window moves pixel-by-pixel over the entire 3-D image. Throughout these simulations, the SSIM measure uses

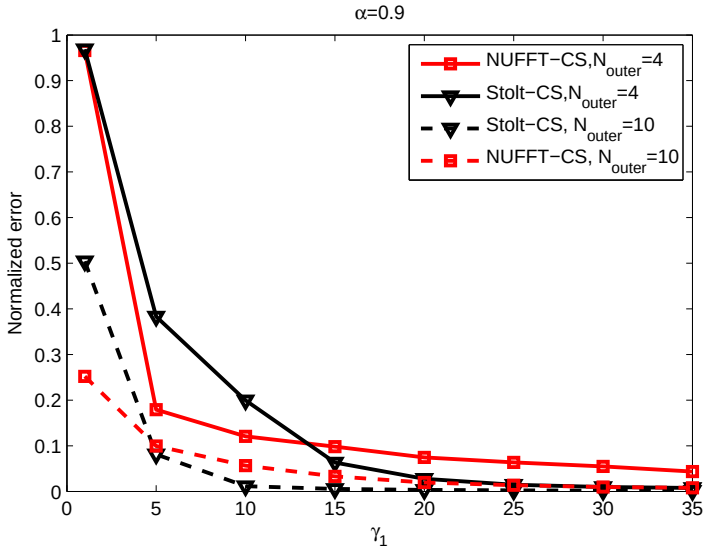
the following typical parameter settings: $C_1 = 10^{-4}, C_2 = 9 \times 10^{-4}$. The weighting function is selected as a spherical-symmetric Gaussian lowpass filter that was $16 \times 16 \times 16$. Its standard deviation is 4 samples that is normalized to unit sum. If the reconstructed image is exactly the same as the ground truth image, then the MSSIM is 1. Thus, a larger MSSIM index corresponds to better reconstruction quality, and vice versa. For all of these simulation results, the MSSIMs were averaged over the results of 8 independent trials.

For the CS approaches, when solving, we adopted The split Bregman framework [19] was adopted to achieve fast convergence and better numerical stability for the CS approaches when solving (4). The sparse transform $\mathbf{\Psi}$ was selected as the Haar wavelet transform in (7). For the NDT&E application, the SUT typically exhibited the nature of sparsity in the wavelet domain [1]. Choosing $\lambda = \gamma_1/2$ and $\gamma_2 = \gamma_1/100$ ensures good convergence. Also, we set $N_{\text{inner}} = 10$ and the sparse transform $\mathbf{\Psi}$ to be 3-D Haar wavelet transform. The parameters for the CG algorithm [26] embedded in the Split Bregman framework were taken as: the CG tolerance $\eta_{cg} = 0.01$ and the iteration limit $N_{\text{cg}} = 20$. The iteration limit is set to be small to keep the overall reconstruction complexity acceptable because it is unnecessary to solve each subproblem entirely to a high numerical precision for SAR imaging.

The parameters γ_1 and N_{outer} were selected according to their relationship with the normalized error $\varepsilon = \|\mathbf{\Phi}\hat{\mathbf{g}} - \mathbf{r}\|_2^2 / \|\mathbf{r}\|_2^2$, as illustrated in Fig. 2. The tolerance was $\eta = 0$. A larger γ_1 indicates a larger measurement consistency. A small γ_1 weakens the denoising function. Both $\gamma_1 = 10$ and $N_{\text{outer}} = 7$ were selected to achieve a good compromise among the data consistency, the denoising capability, and the computational complexity. If $\eta > 0$, the number of outer iteration will always be less than N_{outer}.



(a)



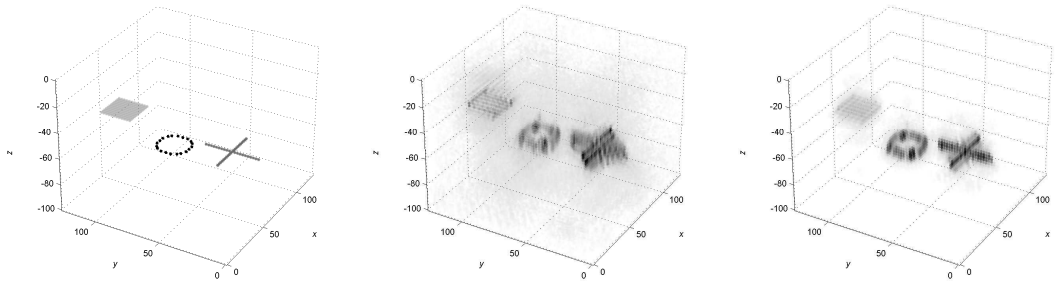
(b)

Figure 2. Normalized error vs. γ_1 with $N_{\text{outer}} = 4$ and $N_{\text{outer}} = 10$ for the two CS approaches. $\eta = 0$, $N_{\text{inner}} = 10$, $\lambda = \gamma_1/2$ and $\gamma_2 = \gamma_1/100$. (a) Undersampling rate $\alpha = 0.4$. (b) Undersampling rate $\alpha = 0.9$.

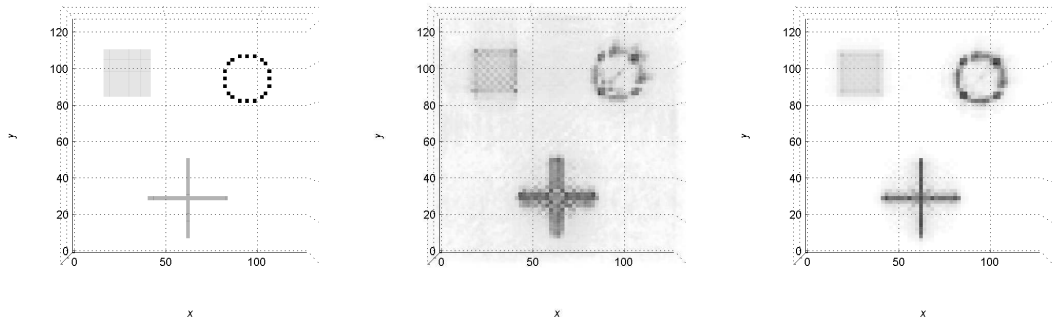
4.2 SIMULATIONS AND RESULTS

The simulated SUT had three objects: a square pad, a cross profile, and a circle profile distributed at depths of -28 mm, -58 mm, and -88 mm, respectively, as illustrated in the leftmost column of Fig. 3. The simulated SUT are represented by complex numbers with a certain magnitude and random phase. The backscatter data were simulated over the square area of 128×128 mm^2 with an additive complex-valued noise. Raw measurements were generated from the SUT according to (1) with a uniform measurement grid of 2 mm in both the X and Y directions. The stepped-frequencies were between from 35.04 GHz and 44.64 GHz (Q-band) with a step-size of $\delta_f = 0.64$ GHz. The maximum depth of the reconstructed image was then $Z_{\max} = c/(4\delta_f) \approx 118$ mm for the benchmark [14]. Spacing along the Z-dimension was set at 2 mm, so that the data cube of the 3-D image had a dimension of $64 \times 64 \times 59$. This is also the dimension of the fully sampled raw data. The undersampled measurements were selected from the fully-sampled raw measurements at random XY positions. All frequency points were kept because reducing the number of frequencies saved little on the acquisition time. The undersampling rate was defined as α being the ratio of the number of points in the undersampled measurements over the number of samples in fully sampled raw data.

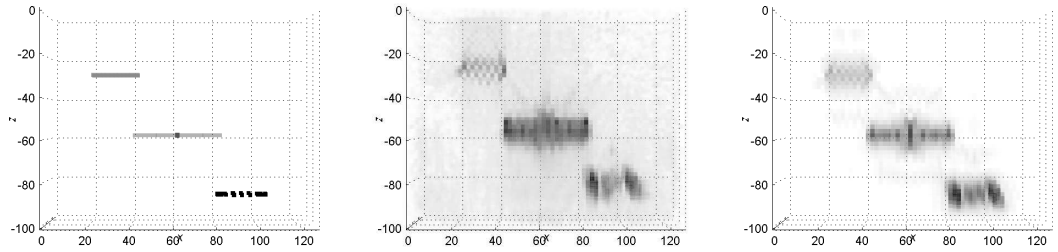
4.2.1 Reconstruction Performance. Figure 3 illustrates a the comparison of the ground truth image, a reconstructed image from 100% raw data using the conventional methods in the noiseless environment. The shadow of the targets and artifacts created by the RMAs and random undersampling are dominant in the 3-D images, as illustrated in the right two columns of Fig. 3. More specifically, the Stolt-RMA blurred the SUT to a large extent. This image indicates that more severe background artifacts are present in the image recovered by the Stolt-RMA than they are in the image recovered by the NUFFT-RMA.



(a) 3-D view.



(b) Top view.

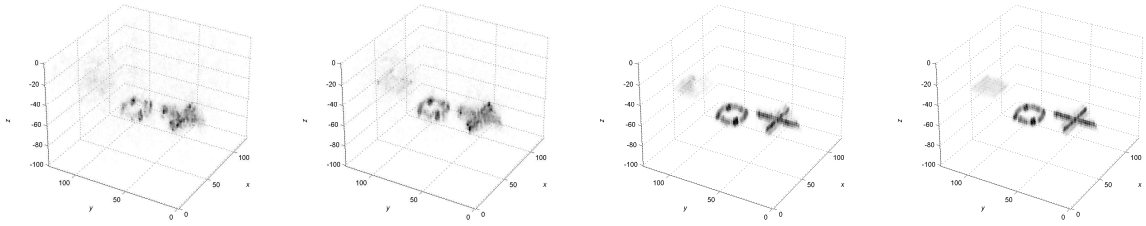


(c) Side view.

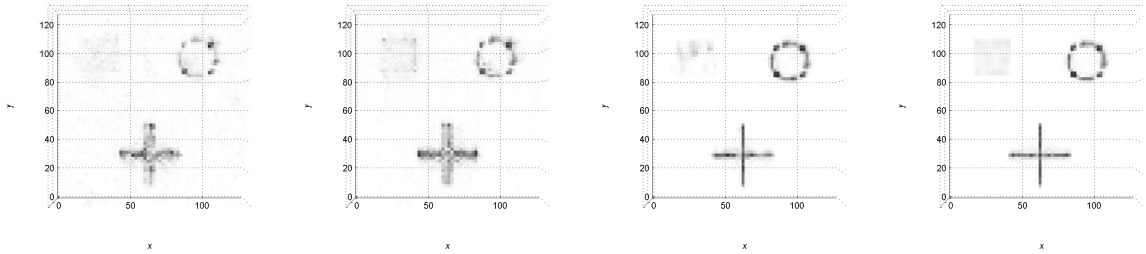
Figure 3. Ground truth image, and reconstructed image from 100% raw data using conventional methods. The three columns (from left to right) represent the ground truth image, Stolt-RMA with 100% data, and NUFFT-RMA with 100% data, respectively.

The reconstructed images using the two CS approaches are given in Fig. 4, with $\eta = 0.25$, 20% or 40% undersampled measurements in the noiseless environment. Not surprisingly, the CS approach based on the NUFFT recovered the 3-D images with better resolution and less background noise than did the CS approach based on the Stolt transform. More specifically, the NUFFT-CS recovered the image

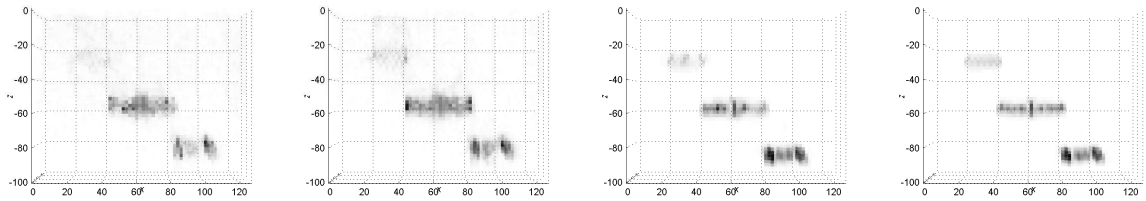
with fewer horizontal shadows of both the cross profile and the circle profile. When the undersampling rate was reduced from 40% and 20%, both the Stolt-CS and the NUFFT-CS suffered little degradation. The reflectivity of the square pad at the lower undersampling rate is weakened. Additionally, both CS approaches failed to remove the vertical blurriness of the targets. Because the vertical resolution of the 3-D SAR image depends on the bandwidth of the imaging system [23], the CS approaches were unable to make up the bandwidth deficiency.



(a) 3-D view.



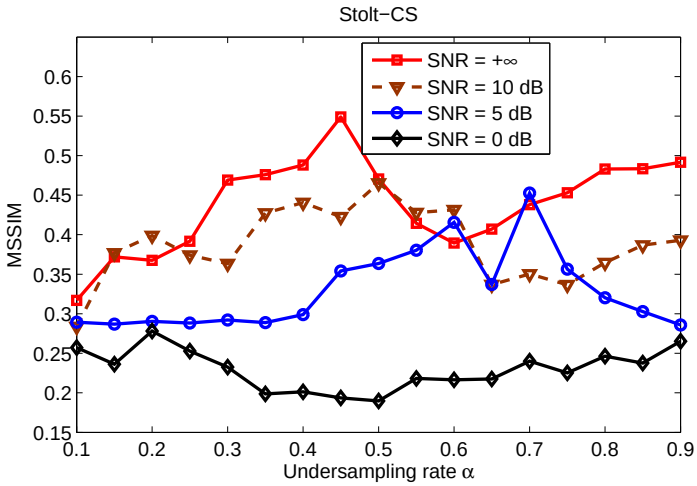
(b) Top view.



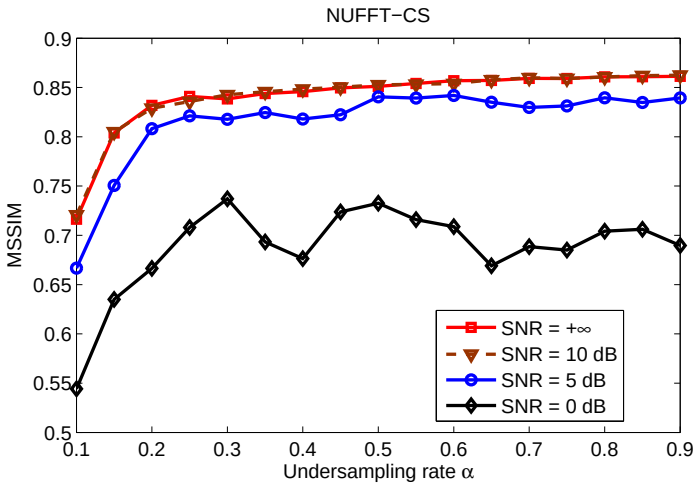
(c) Side view.

Figure 4. Reconstructed images from undersampled simulated measurements using the CS approaches. The four columns (from left to right) represents Stolt-CS with 20% data, Stolt-CS with 40% data, NUFFT-CS with 20% data, and NUFFT-CS with 40% data, respectively. $\eta = 0.25$, SNR $= \infty$.

Figure 5 illustrates the averaged MSSIM between the ground truth image and the reconstructed 3-D SAR images with varying undersampling rates and SNRs. The NUFFT-CS had a much higher MSSIM than did the Stolt-CS at all SNR scenarios, a finding that is consistent with the data illustrated in Fig. 4. This high performance gain is achieved at the cost of high computational complexity.



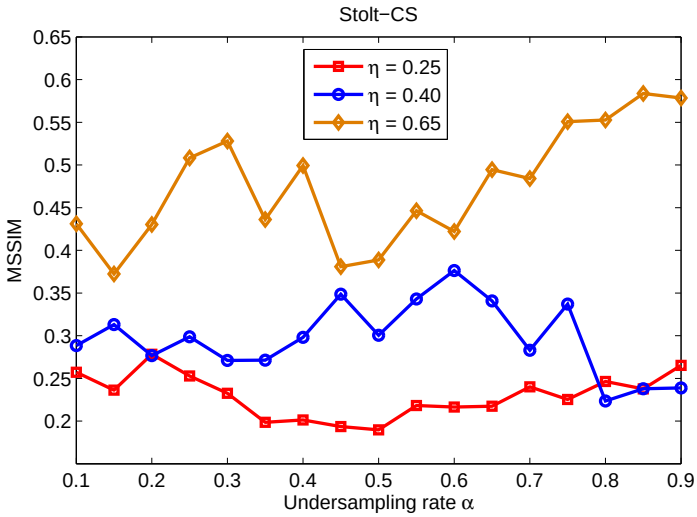
(a)



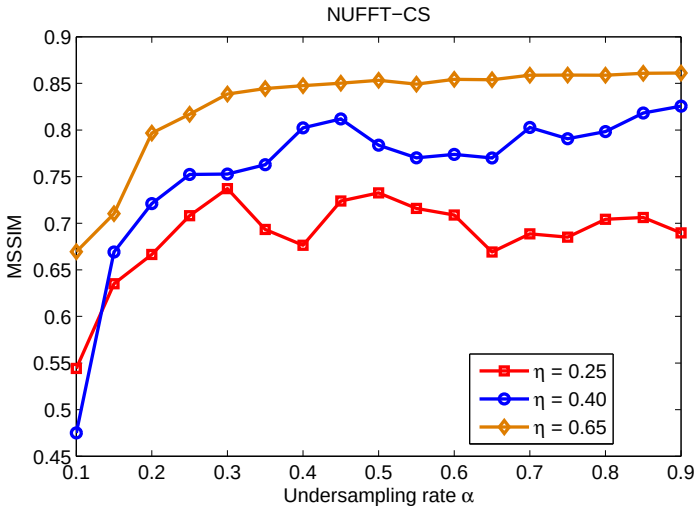
(b)

Figure 5. Averaged MSSIM of the reconstructed 3-D SAR images, using the CS approaches, with varying undersampling rates and SNRs. $\eta = 0.25$. (a) Stolt-CS. (b) NUFFT-CS.

The tolerance (η) also affects the MSSIMs of the reconstructed images, as illustrated in Fig. 6. A larger η results in a larger MSSIM when SNR = 0 dB. For the NUFFT-CS, MSSIM degradation caused by a small η is relatively small when compared to the Stolt-CS. Therefore, the NUFFT-CS can handle noise variations with fixed parameter settings better than the Stolt-CS.



(a)



(b)

Figure 6. Averaged MSSIM of the reconstructed 3-D SAR images using the CS approaches with varying undersampling rates and η. SNR = 0 dB. (a) Stolt-CS. (b) NUFFT-CS.

4.2.2 Computational Complexity. All of the conventional methods and CS approaches were implemented by MATLAB R2011a (x86) on a computer with Intel(R) Core(TM)2 Quad CPU Q9400 at 2.66 GHz and 8.00 GB RAM. The typical running times for conventional Stolt-RMA and NUFFT-RMA were 0.08 s and 0.40 s, respectively. This is regardless of the undersampling rate. The CS approaches have a much higher computational complexity than do the conventional methods, as presented in Fig. 7. The running times were averaged over the results of 8 independent trials for each undersampling rate. The running time of the NUFFT-CS was comparable to that of the NUFFT-DN when the undersampling rate $\alpha > 0.6$, $\eta = 0.25$, and SNR = 10 dB. Interestingly, when SNR = ∞ dB, the running time for the NUFFT-CS was comparable to that of the Stolt-CS at a low undersampling rate and lower than that of the Stolt-CS at a high undersampling rate. This occurred because the NUFFT has better accuracy. Thus, the NUFFT-CS required a smaller number of iterations to reach the stopping criteria. Therefore, the low-complexity Stolt transform cannot decrease the overall computational load for the sparse methods due to its inherent inaccuracy. When SNR = 10 dB, both approaches required a smaller running time than did the zero-noise cases, because it was easier to reach the tolerance criterion. Overall, the NUFFT-CS provided significant improvement on image quality for all cases when $\alpha > 0.25$ at an affordable computational complexity.

4.3 EXPERIMENTS AND RESULTS

The experimental SUT consisted of three layers of construction foam taped together yielding dimensions $120 \times 180 \times 80$ mm^3 [10]. On each layer of foam, three round rubber pads of 5 mm diameter and 2 mm height were embedded at different locations, as illustrated in Fig. 8. The distance between the aperture and the surface of the SUT (standoff distance) was 34 mm. The SNR was approximately 30 dB during the experiment. The same frequency range, step size, and 3-D grids were

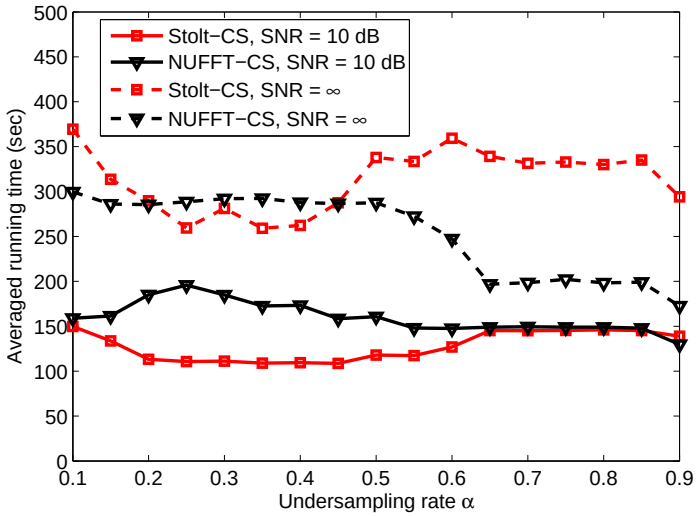
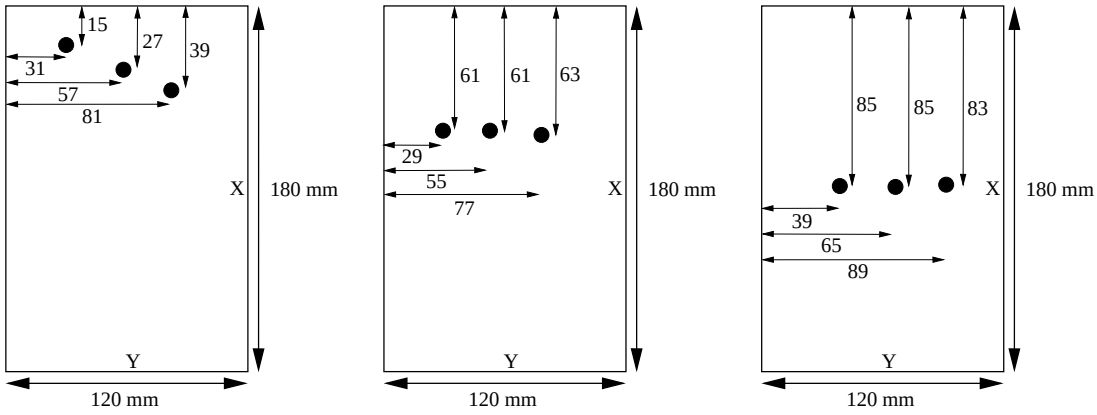


Figure 7. Averaged running time for the two CS approaches with varying undersampling rate. $\eta = 0.25$.

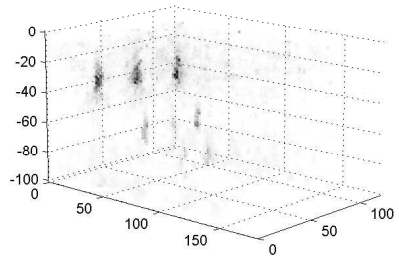


(a) $z = 34$mm (b) $z = 65$mm (c) $z = 81$mm

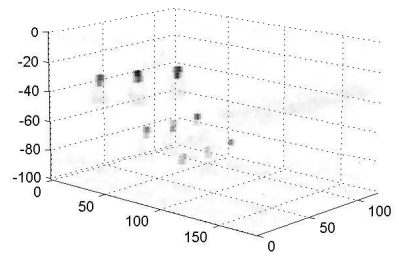
Figure 8. Schematic of the rubber pads in the scanned area of SUT, where unit in figure is mm, and z is the distance from the probe (not to scale).

adopted as those used in the simulations discussed in Section 4.2. Our experiments on a 120×180 SUT show that by using the optimum scanning method, we can reduce the acquisition time from 50 min for full-set measurement of a to 17 min for a typical random sampling of 20% of the full uniform grid [10].

Figure 9 illustrates the reconstructed images from undersampled experimental measurements, using the Stolt-CS and NUFFT-CS approaches with 40% data, respectively. The 3-D image recovered by the NUFFT-CS was more focused than that recovered by the Stolt-CS, while the running time of the Stolt-CS algorithms was approximately 2/3 of that of the NUFFT-CS using the same PC as in Section 4.2. Therefore, the experimental results verified that the characteristics of the two CS approaches in practice are similar to the simulation results.



(a)



(b)

Figure 9. The reconstructed images from undersampled experimental measurements, using the CS approaches, with 40% data. Unit in figure is mm and $\eta = 0.25$. (a) Stolt-CS, the running time is 218 s. (b) NUFFT-CS, the running time is 340 s.

5 CONCLUSION

Two CS approaches for 3-D SAR image reconstruction have been compared in terms of their reconstructive quality and computational complexity. Both numerical simulations and experimental results demonstrate that the CS approach based on the Stolt transform helps little on image quality improvement and computational complexity reduction, albeit the low complexity property of Stolt transform. The NUFFT-CS method significantly improves image quality for an undersampling rate greater than 25% with affordable computational complexity.

6 APPENDIX: PROCEDURES FOR UPDATING AUXILIARY VARIABLES IN THE SPLIT-BREGMAN ALGORITHM FOR 3-D SAR IMAGE RECONSTRUCTION

The procedures of finding optimal $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z, \mathbf{w}$ and updating $\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z, \mathbf{b}_w$ for (10) are described as follows. To obtain the optimal $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z$ in (10), we must iteratively solve the subproblem [19],

$$\{\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z\} = \arg \min_{\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_z} \frac{\lambda}{2}\big[\|\mathbf{d}_x - \mathbf{D}_x\hat{\mathbf{g}} - \mathbf{b}_x\|_2^2 + \|\mathbf{d}_y - \mathbf{D}_y\hat{\mathbf{g}} - \mathbf{b}_y\|_2^2 + \|\mathbf{d}_z - \mathbf{D}_z\hat{\mathbf{g}} - \mathbf{b}_z\|_2^2\big] + \|\mathbf{D}\mathbf{d}\|_1. \tag{16}$$

Specifically, (16) can be solved by the generalized shrinkage formula [24, 25],

$$(\mathbf{d}_x)_l = \max(s - 1/\lambda, 0)\frac{(\mathbf{D}_x\hat{\mathbf{g}} + \mathbf{b}_x)_l}{s}, \tag{17a}$$

$$(\mathbf{d}_y)_l = \max(s - 1/\lambda, 0)\frac{(\mathbf{D}_y\hat{\mathbf{g}} + \mathbf{b}_y)_l}{s}, \tag{17b}$$

$$(\mathbf{d}_z)_l = \max(s - 1/\lambda, 0)\frac{(\mathbf{D}_z\hat{\mathbf{g}} + \mathbf{b}_z)_l}{s}, \tag{17c}$$

where $s = \big\|(\mathbf{D}_x\hat{\mathbf{g}} + \mathbf{b}_x)_l, (\mathbf{D}_y\hat{\mathbf{g}} + \mathbf{b}_y)_l, (\mathbf{D}_z\hat{\mathbf{g}} + \mathbf{b}_z)_l\big\|_2$, and $(\cdot)_l$ denotes the l^{th} element of a vector. Similarly, we use the standard shrinkage formula to compute the optimal $\mathbf{w}$,

$$(\mathbf{w})_l = \max(|(\mathbf{\Psi}\hat{\mathbf{g}} + \mathbf{b}_w)_l| - 1/\gamma_2, 0)\frac{(\mathbf{\Psi}\hat{\mathbf{g}} + \mathbf{b}_w)_l}{|(\mathbf{\Psi}\hat{\mathbf{g}} + \mathbf{b}_w)_l|}. \tag{18}$$

Then we update $\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_z, \mathbf{b}_w$ in the similar way as that in (9),

$$\mathbf{b}_{x,n+1} = \mathbf{b}_{x,n} + \mathbf{D}_x\hat{\mathbf{g}} - \mathbf{d}_{x,n+1}, \tag{19a}$$

$$\mathbf{b}_{y,n+1} = \mathbf{b}_{y,n} + \mathbf{D}_y\hat{\mathbf{g}} - \mathbf{d}_{y,n+1}, \tag{19b}$$

$$\mathbf{b}_{z,n+1} = \mathbf{b}_{z,n} + \mathbf{D}_z\hat{\mathbf{g}} - \mathbf{d}_{z,n+1}, \tag{19c}$$

$$\mathbf{b}_{w,n+1} = \mathbf{b}_{w,n} + \mathbf{\Psi}\hat{\mathbf{g}} - \mathbf{w}_{n+1}. \tag{19d}$$

7 REFERENCES

[1] S. Kharkovsky and R. Zoughi, "Microwave and millimeter wave nondestructive testing and evaluation - overview and recent advances," *IEEE Instrum. Meas. Mag.*, vol. 10, no. 2, pp. 26–38, Apr. 2007.

[2] H. Kajbaf, J. Case, Y. Zheng, S. Kharkovsky, and R. Zoughi, "Quantitative and qualitative comparison of SAR images from incomplete measurements using compressed sensing and nonuniform FFT," in *IEEE Radar Conf. (Radar'11)*, May 2011, pp. 592–596.

[3] H. Kajbaf, Y. R. Zheng, and R. Zoughi, "Improving efficiency of microwave wideband imaging using compressed sensing techniques," *Material Evaluation, ASNT*, Dec. 2012.

[4] Z. Yang and Y. R. Zheng, "Near-field 3-D synthetic aperture radar imaging via compressed sensing," in *IEEE Int'l Conf. Acoustics, Speech, Signal Processing (ICASSP'12)*, Mar. 2012, pp. 2513–2516.

[5] C. Austin, E. Ertin, and R. Moses, "Sparse signal methods for 3-D radar imaging," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 408–423, Jun. 2011.

[6] L. Potter, E. Ertin, J. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006–1020, Jun. 2010.

[7] V. M. Patel, G. R. Easley, D. M. Healy, and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 244–254, Apr. 2010.

[8] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: from theory to applications," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[9] L. Poli, G. Oliveri, and A. Massa, "Microwave imaging within the first-order Born approximation by means of the contrast-field Bayesian compressive sensing," *IEEE Trans. Antennas Propag.*, vol. 60, no. 6, pp. 2865–2879, Jun. 2012.

[10] H. Kajbaf, J. T. Case, Z. Yang, and Y. R. Zheng, "Compressed sensing for SAR-based wideband 3D microwave imaging system using nonuniform FFT," *IET Radar, Sonar, Navig.*, vol. 7, no. 6, pp. 658–670, Jul. 2013.

[11] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by L_1-norm regularization - the compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, 2010.

[12] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-dimensional SAR focusing from multipass signals using compressive sampling," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 488–499, 2011.

[13] R. Bamler, "A comparison of range-Doppler and wavenumber domain SAR focusing algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 4, pp. 706–713, Jul. 1992.

[14] J. M. Lopez-Sanchez and J. Fortuny-Guasch, "3-D radar imaging using range migration techniques," *IEEE Trans. Antennas Propag.*, vol. 48, no. 5, pp. 728–737, May 2000.

[15] R. H. Stolt, "Migration by Fourier transform," *Geophysics*, vol. 43, no. 1, pp. 23–48, 1978.

[16] J. Fessler and B. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.

[17] J. T. Case, M. T. Ghasr, and R. Zoughi, "Optimum 2-D nonuniform spatial sampling for microwave SAR-based NDE imaging systems," *IEEE Trans. Instrum. and Meas.*, vol. 61, no. 11, pp. 3072–3083, Nov. 2012.

[18] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[19] T. Goldstein and S. Osher, "The split Bregman algorithm for L1 regularized problems," *UCLA CAM Report*, pp. 08–29, 2008.

[20] P. Purkait and B. Chanda, "Super resolution image reconstruction through Bregman iteration using morphologic regularization," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4029–4039, Sep. 2012.

[21] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l_1-regularized least squares," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 606–617, Dec. 2007.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[23] D. Sheen, D. McMakin, and T. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.

[24] J. Tropp and S. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, Jun. 2010.

[25] I. Daubechies, M. Defrise, and C. De Mol, "An iteratie thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. LVII, pp. 1413–1457, 2004.

[26] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, Aug. 2006.

III. ROBUST ADAPTIVE CHANNEL ESTIMATION IN MIMO UNDERWATER ACOUSTIC COMMUNICATIONS

Zengli Yang, and Yahong Rosa Zheng

ABSTRACT—For underwater acoustic (UWA) communications, a robust iterative channel estimation based minimum mean-square-error (MMSE) adaptive turbo equalizer is proposed and studied for multiple-input multiple-output (MIMO) detection. Rather than the classical MMSE or normalized least mean squares (NLMS) estimation algorithms, the improved proportionate NLMS (IPNLMS) is adopted for the iterative MIMO channel estimator. The MIMO channel estimation is performed jointly with the MMSE equalizer and the maximum *a posteriori* probability (MAP) decoder. With inter-block interference removed, the MIMO MMSE equalization is performed with overlapped information subblocks without guard intervals, thus a high transmission efficiency is guaranteed and performance degradation is prevented. The proposed MIMO detection scheme has been tested by experimental data and proved to be robust against tough MIMO channels. The experimental results for the undersea 2008 Surface Processes and Acoustic Communications Experiment (SPACE08) are reported.

1 INTRODUCTION

Accurate channel estimation is crucial in achieving satisfactory performance for multiple-input multiple-output (MIMO) underwater acoustic (UWA) communications [1]. Two of the main challenges for UWA channel estimation in the time domain are the extremely long delay spread and the time-varying nature of the channel [2], which often lead to high computational complexity and low estimation accuracy. The traditional block-based least squares (LS) or minimum mean-square-error (MMSE) channel estimation techniques often need the periodic training sequence and the inversion of large matrix [1]. In comparison, the iterative channel estimation based on the adaptive algorithms, such as least mean squares (LMS) or recursive least square (RLS) algorithms [3], can better track the time-varying frequency-selective channel [4]. As the required adaptive filter lengths grows, the conventional normalized LMS (NLMS) [3] algorithm suffers a slow convergence rate, thus requiring long training sequence. However, it is desireable to use short training sequence to reduce the overhead and increase the data transmission efficiency. With short training sequence, this slow convergence rate degrades the accuracy of the channel estimation. The improved proportionate NLMS (IPNLMS) [5] has been designed to ameliorate this situation by exploiting the sparse nature of the channel impulse response (CIR). By updating the coefficient in the weight vector proportionate to its magnitude, the IPNLMS can estimate the UWA channel better than NLMS and the conventional proportionate NLMS (PNLMS) [6], whatever the sparsity of the impulse response is [7] [8].

When adopting turbo equalization [9] at the receiver, accurate channel estimation is necessary to ensure efficient and effective soft information exchange between

the equalizer and the decoder. In the training mode, the proposed MIMO detector employs iterative channel estimation using pilot symbols. In the decision-directed (DD) mode, both the previous and current detected symbols are used to assist channel estimation. To boost the efficiency of turbo equalizer and ease the overall computational costs, the channel estimation is performed jointly with the MMSE equalizer and the maximum *a posteriori* probability (MAP) decoder. More specifically, in the DD mode, these detected symbols are obtained by performing hard decoding and mapping on the extrinsic log-likelihood ratios (LLRs) from the MAP decoder.

In this paper, the proposed MIMO detection scheme using iterative IPNLMS channel estimation and turbo MMSE equalization is tested by field trial data collected in the undersea 2008 Surface Processes and Acoustic Communications Experiment (SPACE08) conducted at Martha's Vineyard, Edgartown, MA, in October 2008. Also, the performance of the proposed turbo MIMO detector is compared and evaluated with the conventional MMSE channel estimation [10] and the NLMS channel estimation [4] [11]. The experimental results demonstrate that the MIMO detector using IPNLMS estimation algorithm took much less turbo iterations than that using MMSE or NLMS estimation algorithms to reach zero BER. Under the time-varying frequency-selective MIMO UWA channel, the proposed MIMO detector can achieve better performance at lower overall computational costs.

2 SIGNALING AND DATA STRUCTURE

Consider an $N \times M$ MIMO underwater acoustic communication system, where N and M are the numbers of transmit transducers and receive hydrophones, respectively. At the transmitter side, each bit stream is independently encoded, interleaved, modulated and then transmitted by a transducers in designated form. Fig. 1 depicts the signalling process on the n-th transmit branch, with $b_{n,p}$, $c_{n,k'}$, $c_{n,k}$, and $s_{n,k}$ being the information bit, the encoded bit, the interleaved bit and the modulation symbol, respectively.

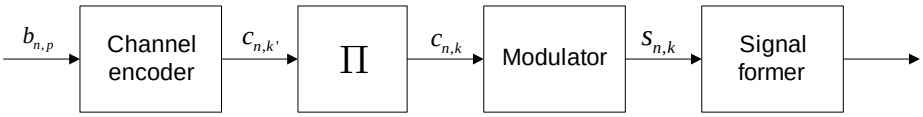


Figure 1. Signaling process on nth transmit branch.

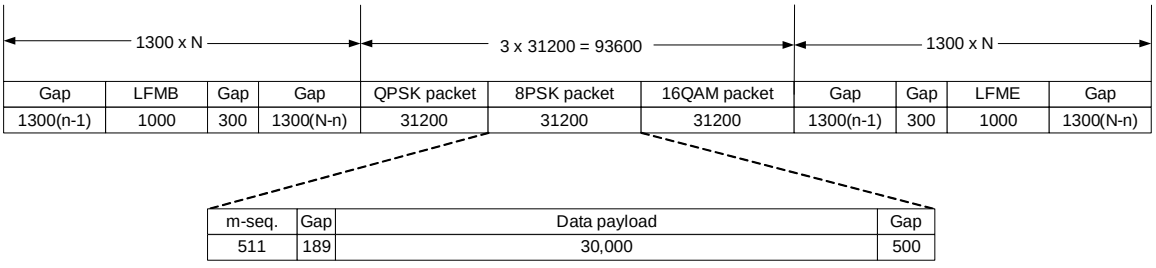


Figure 2. The burst structure of the nth transmit branch in the SPACE08 experiment.

A rate-1/2 non-systematic convolutional channel encoder with generator polynomial $[G_1, G_2] = [17, 13]_{\text{oct}}$ and a random interleaver (Π) are used. The modulator

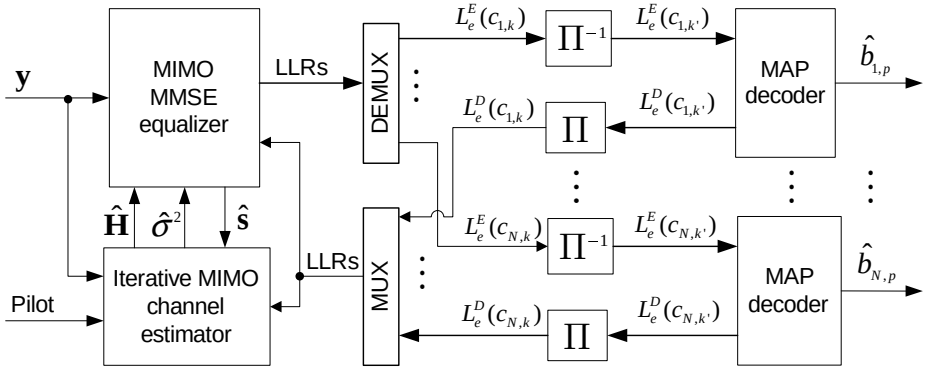


Figure 3. MIMO detector using iterative channel estimation and turbo MMSE equalization.

employs QPSK, 8PSK and 16QAM modulation schemes, with the constellation sizes being 4, 8 and 16, respectively. For a given constellation set $\mathcal{S} = \{\chi_q\}_{q=1}^{Q}$ of size Q, every $\log_2 Q$ interleaved code bits are mapped onto one modulation symbol, *i.e.*, the group of bits, $\{c_{n,(k-1)\log_2 Q+p}\}_{p=1}^{\log_2 Q}$, are mapped to the modulation symbol $s_{n,k}$. The signal former unit assembles the input modulation symbols with auxiliary signals, and then sends out transmission bursts in specific format.

Without loss of generality, a transmission burst of the SPACE08 experiment is illustrated in Fig. 2. The burst begins with a head linear frequency modulation (LFM) signal named LFMB, followed by three packets with QPSK, 8PSK and 16QAM modulations, and ends with a tail LFM signal named LFME. On the receiver side, the LFM signals (LFMB and LFME) will serve multiple purposes including coarse synchronization, Doppler shift estimation and channel length measurement, attributed to their unique correlation properties. Each packet includes a m-sequence of length 511, and a data payload consisting of 30,000 modulation symbols. The m-sequence can be used for evaluating the channel scattering function, for its sensitivity to the Doppler spread [10].

3 ITERATIVE MIMO CHANNEL ESTIMATION BASED TURBO MMSE EQUALIZATION

Figure 3 depicts the structure of the proposed iterative MIMO channel estimation and turbo MMSE equalization. In the training mode, the MIMO detector uses pilot symbols for channel estimation. In the DD mode at the zeroth turbo iteration, the previously detected symbols from the equalizer are used for channel estimation. In the DD mode at the first and subsequent turbo iterations, the detected symbols used for channel estimation are obtained by performing hard decoding and mapping on the extrinsic LLRs of the encoded bits from the interleaver (Π). The estimated CIR $\hat{\mathbf{H}}$ and error variance $\hat{\sigma}^2$ from the channel estimator are then fed into the MIMO MMSE equalizer. The equalizer exchanges the soft information (extrinsic LLRs) on the encoded bits with MAP decoder for each branch. With more reliable soft information from the MAP decoders, this estimated CIR shall show better accuracy over the iterations.

Due to the time-varying nature of the UWA channel, the data payload needs to be partitioned into blocks with length of N_b, as illustrated in Fig. 4. Denote the length of the data payload as N_d, then there are $P = \lceil \frac{N_d}{N_b} \rceil$ blocks in one data payload. The pilot with length of N_p is inserted at the head of each block. Each block is partitioned into subblocks, and each subblock has the length of N_{sb}. The pilot, or the previous and current N_p detected symbols from the equalizer or decoder, are referred to here as the training sequence for the iterative MIMO channel estimator. It is noted that the inter-block interference was removed in the MIMO MMSE equalizer during block processing.

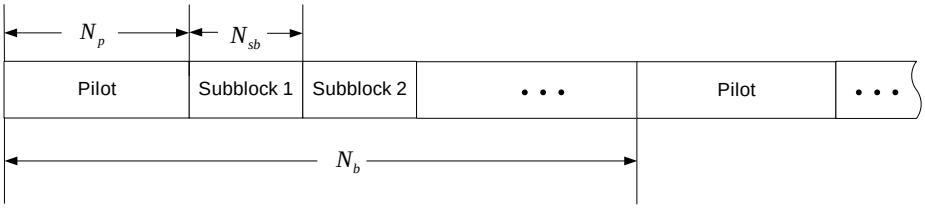


Figure 4. The partition of the transmitted pilot and data payload.

3.1 SYSTEM MODEL

Consider the detection of one subblock, as shown in the Fig. 4, the training sequences with length N_p from N branches are adopted for channel estimation. When the durations of the training sequence is less than the channel coherence time, the channel coefficients can be approximated as quasi time invariant. Denote the training sequences from the N branches as $\{x_{n,k}(0 \leq k \leq N_p - 1)\}_{n=1}^{N}$, and the CIR from n-th transmitter transducer to m-th receiver hydrophone as $\mathbf{h}_{m,n} = [h_{m,n}(0), h_{m,n}(1), \ldots, h_{m,n}(L-1)]^T$, where L is the length of the SISO channel and $(\cdot)^T$ represents transpose, then the signal received at the m-th hydrophone can be approximately expressed, in the matrix form, as

$$\begin{aligned} \mathbf{y}_m^{ce} &\approx \sum_{n=1}^{N} \mathbf{X}_n \mathbf{h}_{m,n} + \mathbf{w}_m \\ &\approx \mathbf{X}\mathbf{h}_m + \mathbf{w}_m. \end{aligned} \tag{1}$$

Here, $\mathbf{y}_m^{ce} \triangleq \left[y_{m,L-1}, y_{m,L}, \cdots, y_{m,N_p-1}\right]^T \in \mathcal{C}^{(N_p-L+1)\times 1}$, and the noise vector $\mathbf{w}_m \triangleq \left[w_{m,L-1}, w_{m,L}, \cdots, w_{m,N_p-1}\right]^T \in \mathcal{C}^{(N_p-L+1)\times 1}$. The matrix $\mathbf{X}_n \in \mathcal{C}^{(N_p-L+1)\times L}$ is defined as

$$\mathbf{X}_n \triangleq \begin{bmatrix} x_{n,L-1} & \cdots & x_{n,1} & x_{n,0} \\ x_{n,L} & \cdots & x_{n,2} & x_{n,1} \\ \vdots & \ddots & \ddots & \vdots \\ x_{n,N_p-1} & \cdots & x_{n,N_p-L+1} & x_{n,N_p-L} \end{bmatrix} \tag{2}$$

which is the matrix of training sequence from n-th transducer. The matrix $\mathbf{X}$ and the vector $\mathbf{h}_m$ are given, respectively, as $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_N] \in \mathcal{C}^{(N_p-L+1)\times NL}$ and $\mathbf{h}_m = [\mathbf{h}_{m,1}^T, \mathbf{h}_{m,2}^T, \cdots, \mathbf{h}_{m,N}^T]^T \in \mathcal{C}^{NL\times 1}$.

3.2 ITERATIVE MIMO UWA CHANNEL ESTIMATION

The UWA channel estimation can be performed by using the nonadaptive algorithms, *i.e.* MMSE [10], or the adaptive algorithms, *i.e.* NLMS [4] [11]. The nonadaptive algorithms perform the channel estimation in block-wise, thus ignoring the time-variant nature of the UWA channel. Instead, the adaptive algorithms have the nature of channel tracking.

3.2.1 MMSE Nonadaptive Channel Estimation. Based on (1), the MMSE estimation of $\mathbf{h}_m$ is obtained as [10]

$$\hat{\mathbf{h}}_m = (\mathbf{X}^H\mathbf{X} + \sigma_w^2\mathbf{I}_{NL})^{-1}\mathbf{X}^H\mathbf{y}_m^{ce}, \tag{3}$$

where the noise variance σ_w^2 at the hydrophones has to be estimated in silent period and is assumed unchanged during the period of data payload. The estimation in (3) is performed on each of the M hydrophones to obtain the MIMO UWA channel estimation. It is noted that the training sequence matrix $\mathbf{X}$ is changed over the subblocks

and turbo iterations. To guarantee the system equation in (1) not underdetermined, the length of the training sequence shall satisfy $N_p \geq (N+1)L - 1$ [10].

3.2.2 IPNLMS Adaptive Channel Estimation. The structure of the adaptive MIMO channel estimator using adaptive algorithms is shown in Fig. 5. The training sequences from N branches and the received signal vector $\mathbf{y}$ at M receive hydrophones are used to identify the MIMO channel. The MIMO channel is modeled as MN FIR filters, where each one has length of L and the coefficients are updated according to the residual error vector ($\mathbf{e}$). With certain adaptive algorithm and proper parameters, the adaptive filters shall converge toward the steady state. Also, the time-averaged error variances are fed back into the MIMO MMSE equalizer.

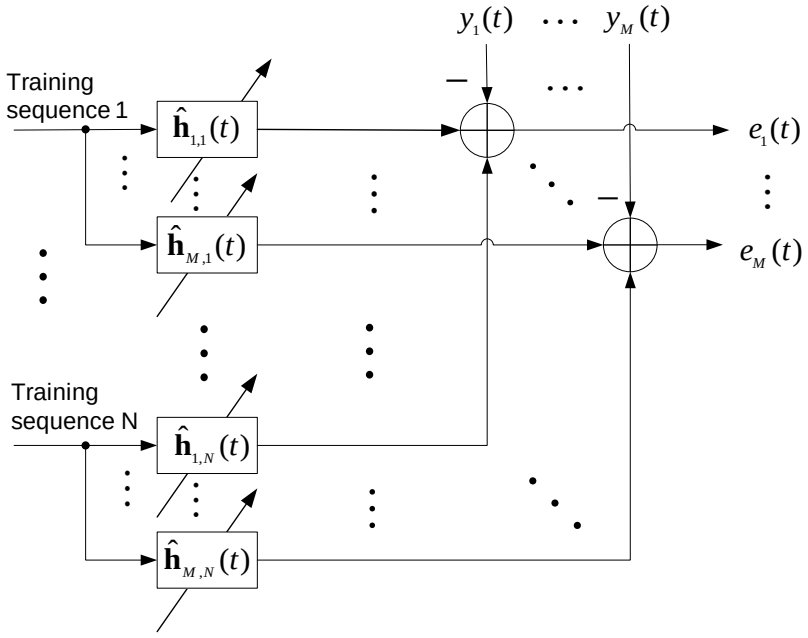


Figure 5. Structure of $N \times M$ MIMO adaptive channel estimator.

With classical NLMS estimation algorithms, $\mathbf{h}_m$ is estimated and updated as [3]

$$\hat{\mathbf{h}}_m(t+1) = \hat{\mathbf{h}}_m(t) + \frac{\mu e_m^*(t)\mathbf{x}(t)}{\mathbf{x}^H(t)\mathbf{x}(t) + \delta_{\text{NLMS}}} \tag{4}$$

where t is the index of the training symbols, μ is the step size, δ_{NLMS} is a small positive parameter to avoid possible division by zero, $\mathbf{x}(t) \triangleq \left[\mathbf{x}_1^T(t), \mathbf{x}_2^T(t), \ldots, \mathbf{x}_N^T(t)\right]^T$, $\mathbf{x}_n(t) \triangleq [x_{n,t}, x_{n,t-1}, \ldots, x_{n,t-L+1}]^T$, and the error at the mth receiver $e_m(t) = y_m(t) - \mathbf{x}^H(t)\hat{\mathbf{h}}_m(t)$. It is noted that the training sequence, as shown in Fig. 4, shall be reused with decreasing step size for proper convergence.

To utilize the sparse nature of the UWA channel, we proposed to adapt the coefficients of $\hat{\mathbf{h}}_m$ proportionately by pre-multiplying the update vector with a proportionate matrix. Then, (4) can be rewritten as

$$\hat{\mathbf{h}}_m(t+1) = \hat{\mathbf{h}}_m(t) + \frac{\mu e_m^*(t)\mathbf{G}_m(t)\mathbf{x}(t)}{\mathbf{x}^H(t)\mathbf{G}_m(t)\mathbf{x}(t) + \delta_{\text{IPNLMS}}} \tag{5}$$

where $(\cdot)^*$ represents conjugate, δ_{IPNLMS} is a small positive parameter to avoid possible division by zero, $\mathbf{G}_m(k)$ is a diagonal proportionate matrix whose elements may be selected according to [6] [5]. Specifically, $\mathbf{G}_m(t) = \text{diag}\{\mathbf{g}_{m,1}(t), \mathbf{g}_{m,2}(t), \ldots, \mathbf{g}_{m,N}(t)\}$ where $\mathbf{g}_{m,n}(t) = [g_{m,n}(t,0), g_{m,n}(t,1), \ldots, g_{m,n}(t,L-1)]$. To make the estimation algorithm robust against various sparsity of the UWA channel, the IPNLMS chooses the diagonal elements of $\mathbf{G}_m(t)$ as [5]

$$g_{m,n}(t,l) = \frac{1-\alpha}{2L} + (1+\alpha)\frac{|\hat{h}_{m,n}(t,l)|}{2\|\hat{\mathbf{h}}_m(t)\|_1 + \varepsilon}, \tag{6}$$

where ϵ regularizes the updating when all coefficients are zero at initialization. For $\alpha = -1$, the IPNLMS reduces to NLMS. For α close to 1, the IPNLMS behaves like the PNLMS [6]. Compared to the MMSE nonadaptive channel estimation, adaptive

channel estimation algorithms has no strict requirement for the length of the training sequence. Also, this iteratively coefficients updating can track the channel changes within current subblock.

3.3 LOW-COMPLEXITY MIMO MMSE TURBO EQUALIZATION

To present the low-complexity (LC) MIMO MMSE equalization algorithm, we define K_1 and K_2 as the length of the noncausal and the causal part of the estimator filter, respectively, and K is the overall filter length, that is, $K = K_1 + K_2 + 1$. Also, we denote the observation for estimating $s_{n,k}$ (the k-th transmitted symbol at n-th transmitter) as $\mathbf{y}^k \triangleq [\mathbf{y}_{k-K_2}^T, \mathbf{y}_{k-K_2+1}^T, \ldots, \mathbf{y}_{k+K1}^T]^T$, where $\mathbf{y}_k \triangleq [y_{1,k}, y_{2,k}, \ldots, y_{M,k}]^T$. The mean and variance of $s_{n,k}$ is denoted as $\bar{s}_{n,k}$ and $v_{n,k}$, respectively, and the mean of $v_{n,k}$ within the subblock is denoted as $\bar{v}$. Both $\bar{s}_{n,k}$ and $\bar{v}$ are obtained based on the *a priori* LLRs of the encoded bits. The details on the calculation are discussed in [12] and omitted here for brevity. With the estimated MIMO channel and the error or noise variance, the LC linear MMSE estimate of the symbol $s_{n,k}$ is given by [11] [12]

$$\begin{aligned} \hat{s}_{n,k} &= \mathbf{f}_n^H(\mathbf{y}^k - \hat{\mathbf{H}}\bar{\mathbf{s}}_n^k) \\ \mathbf{f}_n &= (\sigma_w^2\mathbf{I} + \bar{v}\hat{\mathbf{H}}\hat{\mathbf{H}}^H)\tilde{\mathbf{h}}_n \end{aligned} \tag{7}$$

where

$$\begin{aligned} \bar{\mathbf{s}}_n^k &= [\bar{\mathbf{s}}_{k-K_2-L+1}^T, \ldots, \bar{\mathbf{s}}_{k-1}^T, \tilde{\mathbf{s}}_{n,k}^T, \bar{\mathbf{s}}_{k+1}^T, \ldots, \bar{\mathbf{s}}_{k+K_1}^T]^T \\ \bar{\mathbf{s}}_k &= [\bar{s}_{1,k}, \bar{s}_{2,k}, \ldots, \bar{s}_{N,k}]^T \\ \tilde{\mathbf{s}}_{n,k} &= [\bar{s}_{1,k}, \ldots, \bar{s}_{n-1,k}, 0, \bar{s}_{n+1,k}, \ldots, \bar{s}_{N,k}]^T. \end{aligned}$$

Also, the estimated channel matrix is defined as

$$\hat{\mathbf{H}} \triangleq \begin{bmatrix} \hat{\mathbf{H}}_{L-1} & \cdots & \hat{\mathbf{H}}_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \hat{\mathbf{H}}_{L-1} & \cdots & \hat{\mathbf{H}}_0 \end{bmatrix} \in \mathcal{C}^{MK \times N(K+L-1)}$$

$$\hat{\mathbf{H}}_l \triangleq \begin{bmatrix} \hat{h}_{1,1}(l) & \hat{h}_{1,2}(l) & \cdots & \hat{h}_{1,N}(l) \\ \vdots & \ddots & \ddots & \vdots \\ \hat{h}_{M,1}(l) & \hat{h}_{M,2}(l) & \cdots & \hat{h}_{M,N}(l) \end{bmatrix} \in \mathcal{C}^{M \times N}$$

and $\tilde{\mathbf{h}}_n$ is the $(N(K_2 + L - 1) + n)$-th column of the $\hat{\mathbf{H}}$.

In turbo equalization, the demapped bits of the estimated symbols $\hat{s}_{n,k}$ are converted into the extrinsic LLRs. After the deinterleaver, the extrinsic LLRs are fed into the MAP decoder as the *a priori* information of the encoded bits. Then extrinsic LLRs from MAP decoder after the interleaver are then fed back into the iterative channel estimator for all N branches.

4 EXPERIMENTAL RESULTS

The proposed MIMO detector using the iterative channel estimation and MMSE turbo equalization has been tested by one undersea trial of UWA communications. The trial named SPACE08 was conducted at the coast of Martha's Vineyard, Edgartown, MA, in October 2008. In this experiment, QPSK, 8PSK and 16QAM modulations were used with a symbol period of 0.1024 milliseconds (ms). The carrier frequency was $f_c = 13$ kHz. The transmit filter was a square-root raised cosine filter with roll-off factor $\beta = 0.2$, thus the occupied channel bandwidth was 11.7188 kHz. The transmit equipment consisted of four transducers, numbered 0 through 3. Transducer 0 was fixed on a stationary tripod, and was about 4 meters (m) above the sea bottom. Transducer 1 to 3 were evenly mounted on a vertical array with the inter-transducer spacing being 50 centimeters (cm). The top transducer in the array was about 3 m above the sea bottom. There were six sets of receiving hydrophone arrays placed at six different locations. A detailed description about the six hydrophone arrays, is given in Table 1. It is noted that for the two cross arrays S1 and S2, each

Table 1. Description On The Hydrophone Arrays

Array Name/Type	Range (m)	Orientation	Number of hydrophones	Hydrophone spacing (cm)
S1/Cross	60	Southeast	16	3.75
S2/Cross	60	Southwest	16	3.75
S3/Vertical	200	Southeast	24	5
S4/Vertical	200	Southwest	24	5
S5/Vertical	1000	Southeast	12	12
S6/Vertical	1000	Southwest	12	12

"leg" of the cross consisted of 16 hydrophones. The top hydrophone of each array was approximately 3.3 m above the sea bottom. Finally, the water depth of this experiment was about 15 m. Most of the channel energy is concentrated within 10 ms, corresponding to a channel with approximated length of 100 in terms of the symbol period $T_s = 0.1024$ ms.

An example of the estimated UWA channels is shown in Fig. 6, for a two-transducer 200 m transmission. Clearly, the CIRs are sparse, nonhomogeneous, and some are non-minimum phase. The characteristics of the UWA channels make the channel equalization very difficult.

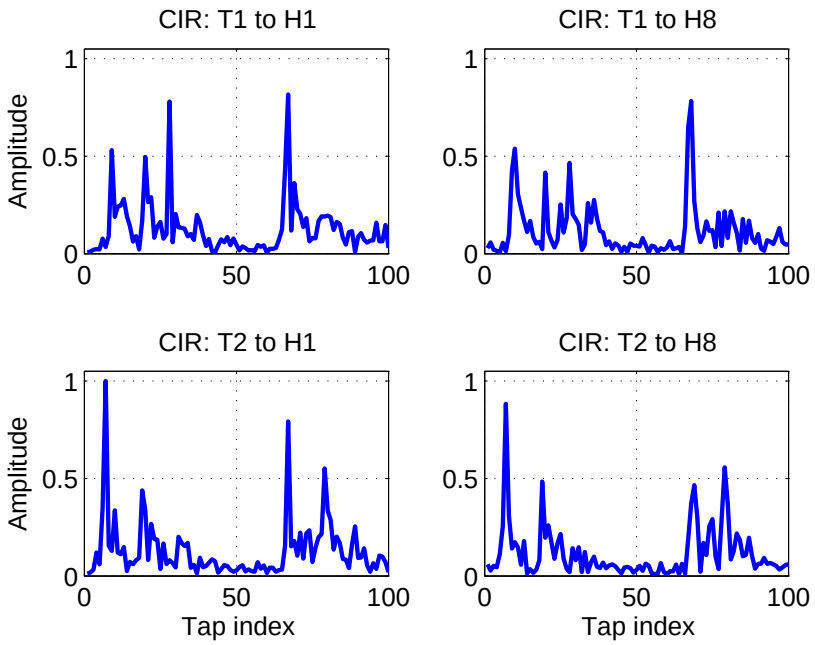


Figure 6. An example of the estimated channel impulse responses in the SPACE08 experiment ('T' denotes transducer, and 'H' denotes hydrophone).

For the 200 m transmission, 30 S3 files and 15 S4 files were recorded in two days, during the SPACE08 experiment. All 45 2×6 packets with QPSK modulation

have been processed with $N_b = 4500$, $N_p = 600$, $N_{sb} = 200$, thus incurring 14% pilot overhead of here. For the LC MMSE equalizer, $K_1 = 100$ and $K_2 = 50$ were chosen. The conventional MMSE and NLMS channel estimator and the proposed IPNLMS channel estimator are studied and compared for this MIMO detector. The training sequences are reused for the adaptive channel estimation algorithms, i.e., NLMS, IPNLMS, with decreasing step size for five times. The lengths of the MN SISO channels are set as $L = 100$. More specifically, other parameters for the adaptive algorithms were set as,

Initial step-size: $\mu = 1$;

Exponential decay factor for data reuse: $\beta = 0.3$;

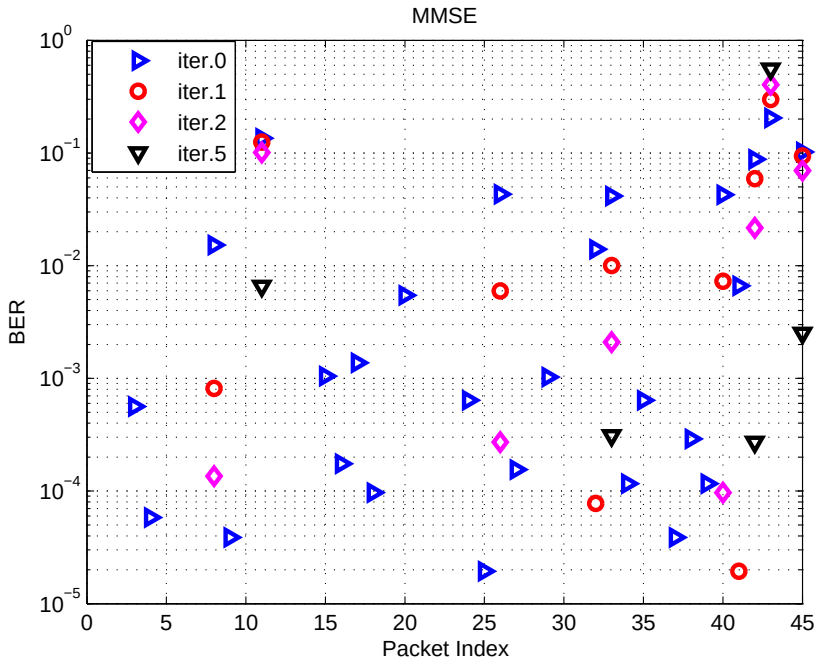
$\delta_{\text{NLMS}} = 0.01, \quad \epsilon = 0.01, \quad \alpha = 0, \quad \delta_{\text{IPNLMS}} = 5 \times 10^{-5}$.

The experimental results using the proposed MIMO detector are shown in Fig. 7 for turbo iterations $0, 1, 2, 5$. The MIMO detector using MMSE channel estimator cannot improve the BER for the tough packets, i.e. packet 11, 44, 45, in the initial several iterations. This is caused by the fast time-varying nature of these tough channels, and the MMSE estimation algorithm failed to track the changes within sub-block. Also, the incorrect estimated symbols cause error propagation within block in the DD mode. The QPSK packet detection results are also listed in Table 2 in terms of the number of iterations to achieve zero BER. From the table, all 45 QPSK packets have achieved zero BER with no more than two iterations when adopting the proposed iterative IPNLMS channel estimator. In comparison, to make 40 QPSK packets achieve zero BER, the MMSE channel estimator took five iterations, and the NLMS channel estimator took four iterations. In [10], MMSE channel estimation and turbo block decision-feedback equalization (BDFE) were adopted, and the results of 2×12 MIMO transmission (200 m) were reported in its Table II. In contrast, with

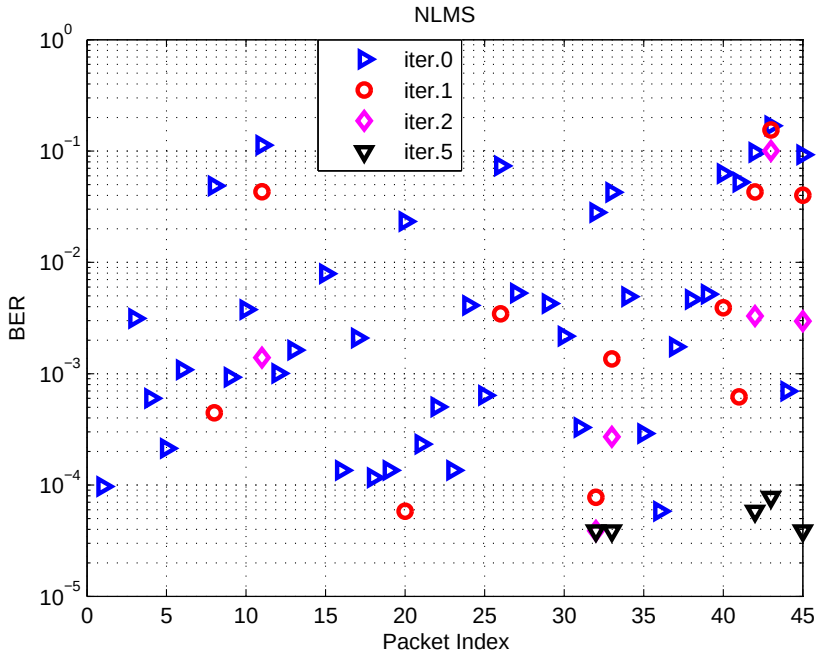
a even lower computational complexity, the proposed MIMO detector using iterative IPNLMS channel estimation and MMSE turbo equalization takes two iterations less to make all QPSK packets achieve zero BER.

Table 2. Results of 2×6 MIMO Transmission (200 m, QPSK modulation)

Number of iterations to achieve zero BER	Number of packets (MMSE)	Number of packets (NLMS)	Number of packets (IPNLMS)
0	18	4	21
1	17	30	20
2	2	5	4
3	2	0	-
4	0	1	-
5	1	0	-

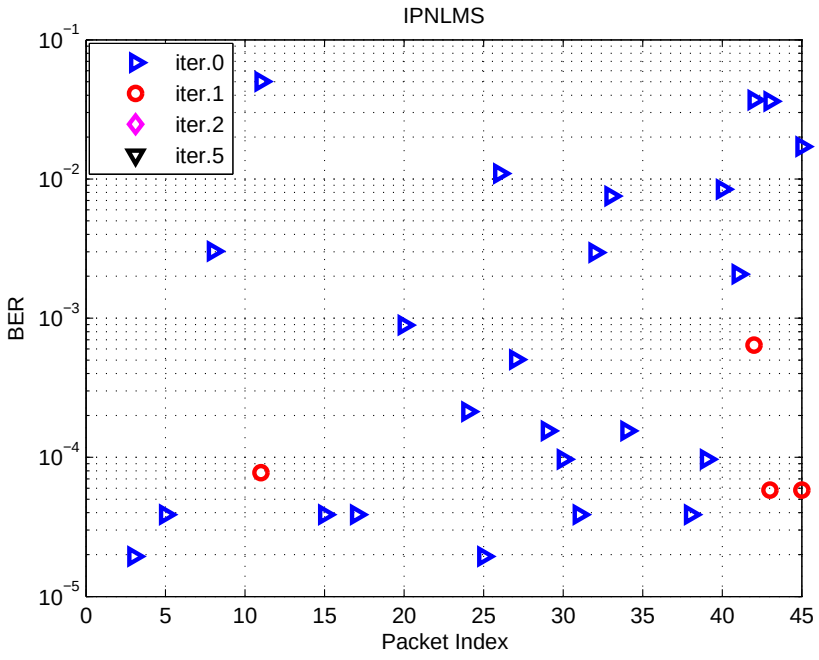


(a)



(b)

Figure 7. Experimental results with iterative MIMO channel estimator and MMSE turbo equalization. QPSK modulation, 200 m transmission, 5 times training sequence reuse for the adaptive channel estimation algorithms, i.e., NLMS, IPNLMS. Those packet indices without corresponding BER value shown for certain iteration indicate zero BERs have been achieved. (a) MMSE block channel estimation. (b) NLMS adaptive channel estimation.



(c)

Figure 7. (c) IPNLMS adaptive channel estimation. (cont.)

5 CONCLUSION

The experimental results demonstrate that the proposed MIMO turbo detector using iterative IPNLMS channel estimation outperforms that using conventional MMSE or NLMS estimation algorithms. By utilizing the sparse nature of the CIRs, the IPNLMS can well track the time-varying frequency-selective UWA channel. With more accurate channel estimation, fewer turbo iterations are needed to achieve zero BER, thus saving much computational costs at no performance loss.

6 ACKNOWLEDGEMENT

This work was supported by the National Science Foundation under Grant ECCS-0846486.

7 REFERENCES

[1] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[2] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, Jan. 2009.

[3] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, New Jersey, 07458.: Prentice Hall, 2002.

[4] R. Otnes and M. Tuchler, "Iterative channel estimation for turbo equalization of time-varying frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1918–1923, Nov. 2004.

[5] J. Benesty and S. L. Gay, "An improved pnlms algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2002. (ICASSP '02)*, vol. 2, 2002, pp. 1881 –1884.

[6] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech, Audio Process.*, vol. 8, no. 5, pp. 508 –518, Sep. 2000.

[7] Z. Yang, Y. R. Zheng, and S. L. Grant, "Proportionate affine projection sign algorithms for network echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 8, pp. 2273–2284, Nov. 2011.

[8] Z. Yang, Y. R. Zheng, and S. L. Grant, "Proportionate affine projection sign algorithms for sparse system identification in impulsive interference," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP'11)*, 2011, pp. 4068–4071.

[9] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.

[10] J. Tao, Y. R. Zheng, C. Xiao, and T. Yang, "Robust MIMO underwater acoustic communications using turbo block decision-feedback equalization," *IEEE J. Ocean. Eng.*, vol. 35, no. 4, pp. 948–960, Oct. 2010.

[11] J. W. Choi, T. J. Riedl, K. Kim, A. C. Singer, and J. C. Preisig, "Adaptive linear turbo equalization over doubly selective channels," *IEEE J. Ocean. Eng.*, vol. 36, no. 4, pp. 473–489, Oct. 2011.

[12] M. Tuchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using *a priori* information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673–683, Mar. 2002.

SECTION

2 CONCLUSIONS

This dissertation proposed three classes of new sparse nonlinear optimization methodology to the network echo cancellation (NEC), 3-D synthetic aperture radar (SAR) image reconstruction, and adaptive turbo equalization in multiple-input multiple-output (MIMO) underwater acoustic (UWA) communications, respectively.

For NEC, two proportionate affine projection sign algorithm (APSA) have been proposed for the identification of real-coefficient, sparse systems. With a modest increase in computational complexity over that of the original APSA, the proportionate APSAs can achieve faster convergence rate and lower in a steady-state misalignment in a sparse network echo path, colored input, and impulsive interference environment. The computational complexity of the two proportionate APSAs is lower than the affine projection algorithm (APA) family due to elimination of the matrix inversion. Especially, the real-coefficient proportionate APSA (RIP-APSA) also exhibits good robustness in all types of network impulse responses (NIRs) without the need to change parameters or estimate the sparseness of the NIRs. Numerical results also demonstrate that the proportionate APSAs exhibit asymptotic convergence for all step size $\mu > 0$, but practically, a small step size of $\mu < 0.1$ is required to achieve excess mean square error (MSE) less than 0 dB.

For 3-D SAR image reconstruction, the proposed sparse methods are compared in the area of imaging denoising and compressed sensing (CS) due to its large-scale and static feature. The CS technique has been successfully simulated, implemented and investigated in a near-field wideband 3-D SAR imaging system. Two CS approaches for 3-D SAR image reconstruction have been compared in terms of their reconstructive quality and computational complexity. Both numerical simulations and

experimental results demonstrate that the CS approach based on the Stolt transform helps little on image quality improvement and computational complexity reduction, albeit the low complexity property of Stolt transform. The non-uniform fast Fourier transform (NUFFT)-based CS approach significantly improves image quality for an undersampling rate greater than 25% with affordable computational complexity.

For MIMO UWA communications, the experimental results demonstrate that the proposed MIMO turbo detector using iterative improved normalized least mean squares (IPNLMS) channel estimation outperforms that using conventional minimumu MSE (MMSE) or normalized least mean squares (NLMS) estimation algorithms. By utilizing the sparse nature of the channel impulse responses (CIRs), the IPNLMS can well track the time-varying frequency-selective UWA channel. With more accurate channel estimation, fewer turbo iterations are needed to achieve zero bit-error-rate (BER), thus saving much computational costs at no performance loss.

The contributions of my PhD research work are summarized in five journal papers and seven conference papers, among which, two journal papers and one conference paper are included in this dissertation.

3 PUBLICATIONS

[1] Z. Yang, B. Han, and Y. R. Zheng,"DSP implementation of DS/CDMA underwater acoustic modems," *MTS/IEEE OCEANS 2014*, St. John's, Newfoundland, Canada, Sep. 14–19, 2014. [To be submitted]

[2] Z. Yang, and Y. R. Zheng,"Robust iterative channel estimation for turbo equalization in MIMO underwater acoustic communications," *IEEE J. Ocean. Eng.*, May 2014. [To be submitted]

[3] Z. Yang, and Y. R. Zheng,"Robust adaptive channel estimation in MIMO underwater acoustic communications," *MTS/IEEE OCEANS 2014*, Taipei, Taiwan, Apr. 7–10, 2014, pp. 1–6. [Accepted]

[4] Z. Yang, and Y. R. Zheng,"A comparative study of compressed sensing approaches for 3-D synthetic aperture radar image reconstruction," *Elsevier Digital Signal Process.*, pp. 1–28, Mar., 2014. [Accepted]

[5] B. Han, Z. Yang, and Y. R. Zheng,"Efficient implementation of iterative MIMO-OFDM receiver using MMSE interference cancelation," *IET Commun.*, pp. 1–10, Dec. 2013. [To be published]

[6] Y. R. Zheng, Z. Yang, J. Hao, and P. Han, "Hardware implementation of underwater acoustic localization system for bridge scour monitoring," *MTS/IEEE OCEANS 2013*, San Diego, CA, Spet. 23–26, 2013, pp.1–6.

[7] B. Han, Z. Yang, and Y. R. Zheng, "FPGA implementation of QR decomposition for MIMO-OFDM using four CORDIC cores," *IEEE Int. Conf. Commun. (ICC'13)*, Budapest, Hungary, Jun. 9–13, 2013, pp. 4556–4560.

[8] H. Kajbaf, J. T. Case, Z. Yang, and Y. R. Zheng,"Compressed sensing for SAR-based wideband 3D microwave imaging system using nonuniform FFT," *IET Radar, Sonar, Navig.*, vol. 7, no. 6, pp. 658–670, Jul., 2013.

[9] Z. Yang, and Y. R. Zheng,"Near-field 3-D synthetic aperture radar imaging via compressed sensing," *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP'12)*, Kyoto, Japan, Mar. 25–30, 2012, pp. 2513–2516.

[10] M. Yang, Z. Yang, B. Hirst, Y. R. Zheng, A. Singh, and L. Ma,"Determination of systolic and diastolic functions of mouser heart using Cine-MRI with compressed sensing," *In Vivo MR Gordon Research Conf.*, Waterville, ME, Jul. 29–Aug. 3, 2012.

[11] Z. Yang, Y. R. Zheng, and S. L. Grant,"Proportionate affine projection sign algorithms for network echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 8, pp. 2273–2284, Nov. 2011.

[12] Z. Yang, Y. R. Zheng, and S. L. Grant,"Proportionate affine projection sign algorithms for sparse system identification in impulsive interference," *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP'11)*, Prague, Czech Republic, May 22–27, 2011, pp. 4068–4071.

BIBLIOGRAPHY

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[2] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, Aug. 2006.

[3] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, New Jersey, 07458.: MPrentice Hall, 2002.

[4] C. Austin, E. Ertin, and R. Moses, "Sparse signal methods for 3-D radar imaging," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 408–423, Jun. 2011.

[5] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.

[6] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech, Audio Processing*, vol. 8, no. 5, pp. 508 –518, Sep. 2000.

[7] T. Gansler, S. Gay, M. Sondhi, and J. Benesty, "Double-talk robust fast converging algorithms for network echo cancellation," *IEEE Trans. Speech, Audio Processing*, vol. 8, no. 6, pp. 656 –663, Nov. 2000.

[8] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, 2002. (ICASSP '02)*, vol. 2, 2002, pp. 1881 –1884.

VITA

Zengli Yang was born in September, 1987 in Suizhou, Hubei, China. He received his B.E. degree in Automation from Huazhong University of Science and Technology, Wuhan, Hubei, China, in 2009. He began his Ph.D. study in August 2009 at the Department of Electrical and Computer Engineering at Missouri University of Science and Technology. His research interests include adaptive signal processing, numerical optimization, compressed sensing, wireless communications, and hardware implementation. He received his Ph.D. degree in Electrical Engineering from Missouri University of Science and Technology in May 2014.