Mixed Norm Constrained Sparse APA Algorithm for Satellite and Network Echo Channel Estimation

YINGSONG LI[1,2,3], **(Member, IEEE), ZHENGXIONG JIANG**[1],
OMER MOHAMMED OMER OSMAN[1], **XIAO HAN**[3],
AND JINGWEI YIN[3], **(Member, IEEE)**

[1]College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
[2]Key Laboratory of Microwave Remote Sensing, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China
[3]Acoustic Science and Technology Laboratory, Harbin Engineering University, Harbin 150001, China

Corresponding author: Yingsong Li (liyingsong@ieee.org)

ABSTRACT We propose an improved proportionate affine projection algorithm (PAPA), which is realized by integrating a hybrid-norm constraint into the affine projection algorithm (APA) to estimate the cluster sparse signals that are happened in satellite and network echo channels. The proposed algorithm optimizes the hybrid $l_{2,0}$-norm of the filter coefficients, and the PAPA is a special case of the proposed method. Moreover, an enhanced PAPA is modified for the cluster sparse channel estimations. Various simulation experiments are performed to verify that the proposed algorithms are superior to the APA, PAPA, and related sparse algorithms with different input signals and various parameters.

INDEX TERMS Cluster sparse signal, $l_{2,0}$-norm, proportionate affine projection algorithm, sparse system identification.

I. INTRODUCTION

Adaptive filtering has been widely studied and used for various engineering applications, such as system identification, channel estimation, active noise control and echo cancellation [1], [2]. The normalized least mean square (NLMS) algorithm has been extensively exploited owing to its low complexity and easy implementation [3]–[5]. However, the NLMS suffers from slow convergence for colored inputs [6]. Therefore, the data reusing method has been presented to create the affine projection algorithm (APA) to provide faster convergence when the inputs are colored [6]. Furthermore, natural signals may have sparse structure, and in this case, most of the signal model coefficients are zero or near-zero while only a few coefficients are active. Thus, sparse system identification has attracted substantial attention in the area of adaptive filtering [7]–[11].

In recent years, proportionate-type adaptive filtering has been widely developed and used for different sparse

signal estimations, including channel estimations and echo cancellations [7]–[9]. The first proportionate adaptive filtering is the proportionate NLMS (PNLMS), which updates each filter coefficient by proportionately assigning a different adaption step size to each estimated filter coefficient [8]. The PNLMS converges rapidly in the initial stage and provides good tracking ability in comparison with the NLMS [8], [9]. However, the PNLMS might achieve a worse estimation mean square error (MSE) when it converges. Several improved PNLMSs have been reported and analyzed based on various weighting rules, such as μ-law and ε-law rules, which stimulate iterations for complete convergence [12]–[15]. Moreover, partition block (PB) [16] and block sparse (BS) [17] PNLMSs have been proposed to identify the active and inactive regions in the echo channel response. PB-PNLMS uses different algorithms to track the channels with two partitioned blocks, while BS-PNLMS identifies the block channel impulse response (CIR) using

the block scheme of the l_2-norm and utilizes the sparseness of the l_1-norm penalty [16], [17]. In fact, the presented PB-IPNLMS uses two different adaptive algorithms to implement the time-domain partitioned blocks.
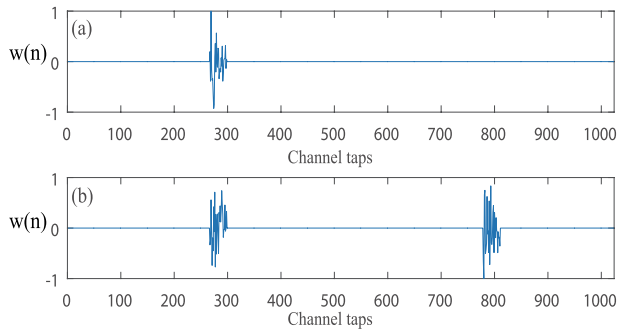

FIGURE 1. Two different cluster sparse systems: (a) single-cluster system; (b) double-cluster system.

Similar to the NLMS, the PNLMS algorithm suffers from slow convergence when the inputs are colored. The proportionate APA (PAPA) and the improved PAPA (IPAPA) have been presented to exploit the sparseness of the echo channel [18]. Furthermore, the memory-improved PAPA method, which is realized by making use of the memory of proportionate taps [19], [20], has been reported to accelerate the convergence rate and to reduce the computational burden. Several improved PAPA methods have been reported and analyzed [19]–[23]. In fact, the PAPA can be obtained from the basis pursuit and convex optimization [24], [25], which renders the PAPA more powerful for practical applications with respect to the modern hardware platform. The network echo path and the satellite channels are usually characterized by a bulk delay caused by the network encoding, loading and jitter buffer delays [25], [26]. Thus, the echoes of sparse systems are divided into two types, as shown in Figure 1. Figure 1(a) is a typical cluster sparse network channel with an active region in the channel ranging from 8 ms to 12 ms, and the other channel taps are close to zero and are denoted as inactive regions. Thus, the CIR of the network echo is usually a single-cluster sparse system, as presented in Figure 1(a). Satellite communication develops so fast that it has attracted increased attention for satellite networks. The CIR from the echoes of satellite-linked communication systems always contains active regions with long flat delays and dispersive active regions [17], [27], [28]. In fact, single-side modulated satellite communication links usually include large far-end echo that has a short delay and a small near-end echo that has a longer delay. Thus, the satellite echo path can be regarded as the two-cluster sparse channel shown in Figure 1(b).

To accurately estimate these sparse channels, the PAPA algorithm can be improved to further exploit the cluster sparse characteristics. In this paper, we propose a cluster sparse (CS) PAPA (CS-PAPA) and CS improved PAPA (CS-IPAPA) to make use of the cluster characteristics of the channels. The proposed CS-PAPA algorithm is realized by incorporating

an $l_{2,0}$-norm into the cost function of the PAPA algorithm, where the l_2-norm divides the clusters and the l_0-norm aims to exploit the sparsity of the channels. The simulation results indicate that the proposed CS-PAPA and CS-IPAPA achieve faster convergence for handling cluster sparse signals under different inputs. Additionally, the basic APA and PAPA are special cases of the proposed cluster sparse PAPAs.

II. REVIEW OF THE PAPA AND IPAPA

In the classic adaptive filtering algorithm, the far-end signal is assumed to be $\mathbf{x}(m) = [x(m), x(m-1), x(m-2), \cdots, x(m-K+1)]^T$, where m is the time index. The impulse response (IR) of the unknown system is $\mathbf{w}(m) = [w_0(m), w_1(m), \cdots, w_{K-1}(m)]^T$, where the length of the coefficient vector is K. Then, the desired signal $d(m)$ containing the output of the echo channel and the near-end signal is given by [5]

$$d(m) = \mathbf{x}^T(m)\mathbf{w}(m) + u(m), \tag{1}$$

where $u(m)$ is the additive noise, which is assumed to be independent from the far-end signal $\mathbf{x}(m)$, and $\mathbf{x}^T(m)$ is the transpose operation of $\mathbf{x}(m)$. The estimated IR vector is denoted to be $\hat{\mathbf{w}}(m-1) = [\hat{w}_0(m-1), \hat{w}_1(m-1), \cdots, \hat{w}_{K-1}(m-1)]^T$. Therefore, the output of the filter is

$$y(m) = \mathbf{x}^T(m)\hat{\mathbf{w}}(m-1). \tag{2}$$

Then, the estimation error is obtained

$$e(m) = d(m) - y(m). \tag{3}$$

A. PAPA

Inspired by the PNLMS algorithm, PAPA introduces a data reusing scheme and proportionate scheme to reuse the input signals to improve the convergence speed and to exploit the sparsity. The input data matrix of the PAPA is defined as [18]

$$\mathbf{X}(m) = [\mathbf{x}(m), \mathbf{x}(m-1), \cdots, \mathbf{x}(m-P+1)], \tag{4}$$

where P is the projection order of the PAPA. The output vector $y(m)$, the expected signal vector $d(m)$ and the estimation error vector $\mathbf{e}(m)$ of the PAPA are given by

$$\mathbf{y}(m) = \mathbf{X}^T(m)\hat{\mathbf{w}}(m-1), \tag{5}$$
$$\mathbf{d}(m) = [d(m), d(m-1), \cdots, d(m-P+1)]^T, \tag{6}$$
$$\mathbf{e}(m) = \mathbf{d}(m) - \mathbf{y}(m) \tag{7}$$

The updated equation of the PAPA is described as [18]

$$\hat{\mathbf{w}}(m) = \hat{\mathbf{w}}(m-1) + \mu\mathbf{G}(m)\mathbf{X}(m)\mathbf{e}(m)$$
$$\times \left(\mathbf{X}^T(m)\mathbf{G}(m)\mathbf{X}(m) + \delta_{\text{PAPA}}\mathbf{I}_P\right)^{-1} \tag{8}$$

where μ is the overall step size, δ_{PAPA} is a regularization parameter, $\mathbf{G}(m)$ is the gain allocation matrix, which is a diagonal matrix, and $\mathbf{I}_P$ is a $P \times P$ identity matrix. The proportionate matrix $\mathbf{G}(m)$ is given by [7], [18]

$$\mathbf{G}(m) = \text{diag}\{g_0(m), g_1(m), \ldots, g_{K-1}(m)\}, \tag{9}$$

where the elements of $\mathbf{G}(m)$ are defined as

$$g_k(m) = \frac{\varphi_k(m)}{\sum_{i=0}^{K-1} \varphi_i(m)}, \quad 0 \le k \le K - 1, \quad (10)$$

where

$$\varphi_k = \max \left\{ a\max \left\{ b, |\hat{w}_0|, |\hat{w}_k|, \cdots, |\hat{w}_{K-1}| \right\}, |\hat{w}_k| \right\}, \quad (11)$$

where the time index is ignored. Parameter $b > 0$ is an initialization parameter that is used to prevent $\hat{\mathbf{w}}(m-1)$ from stalling at the initialization stage when $\hat{\mathbf{w}}(m-1) = \mathbf{0}_{K\times 1}$, while parameter $a > 0$ prevents $\hat{w}_k(m-1)$ from stalling when they are smaller than the largest coefficient [8]. Usually, parameter a ranges from $\frac{1}{K}$ to $\frac{5}{K}$.

B. IPAPA
The IPAPA, which introduces the l_1-norm to reassign the gains to each coefficient, is similar to the IPNLMS [12]. The updating equation of the IPAPA is the same as that of the PAPA and is presented in (8) [9], [29].

$$\varphi_k(m) = (1 - \alpha) \frac{\|\hat{\mathbf{w}}(m)\|_1}{K} + (1 + \alpha) |\hat{w}_k(m)| \quad (12)$$

where $-1 \le \alpha < 1$ and $0 \le k \le K - 1$. Using Equation (12) to replace Equation (11), we obtain

$$g_k(m) = \frac{1-\alpha}{2L} + (1 + \alpha) \frac{|\hat{w}_k(m)|}{2\sum_{i=0}^{K-1} |\hat{w}_i(m)| + \varepsilon}. \quad (13)$$

Similar to the IPNLMS algorithm, the IPAPA turns into APA when $\alpha = -1$. When α is close to 1, the IPAPA becomes the PAPA.

III. PROPOSED CS-PAPA AND CS-IPAPA ALGORITHMS
In this section, we propose the CS-PAPA and CS-IPAPA to exploit the cluster sparse characteristics in the satellite and network echo channels illustrated in Figure 1.

A. THE PROPOSED CS-PAPA ALGORITHM
The proposed CS-PAPA is obtained by minimizing the cost function obtained by adding a mixed $l_{2,0}$-norm to the cost function of the PAPA to fully exploit the cluster sparse characteristics. The CS-PAPA solves the following problem

$$\min \|\hat{\mathbf{w}}(m)\|_{2,0} \quad s.t \ \mathbf{d}(m) - \mathbf{X}^T(m)\hat{\mathbf{w}}(m) = 0 \quad (14)$$

where the $l_{2,0}$-norm is defined as

$$\|\hat{\mathbf{w}}(m)\|_{2,0} = \left\| \begin{bmatrix} \|\hat{\mathbf{w}}_{[1]}\|_2 \\ \|\hat{\mathbf{w}}_{[2]}\|_2 \\ \vdots \\ \|\hat{\mathbf{w}}_{[N]}\|_2 \end{bmatrix} \right\|_0, \quad (15)$$

and N is the number of clusters, which is $N = K/B$, where B is the number of channel coefficients per cluster. Since the

solution of the l_0-norm is an NP-hard problem, it can be approximated by [30]

$$\|\hat{\mathbf{w}}(m)\|_0 \approx \sum_{k=0}^{K-1} (1 - e^{-\beta|\hat{w}_k|}), \quad (16)$$

where $\beta > 0$. From (15) and (16), we obtain

$$\|\hat{\mathbf{w}}(m)\|_{2,0} \approx \sum_{i=1}^{N} (1 - e^{-\beta\|\hat{\mathbf{w}}_{[i]}\|_2}). \quad (17)$$

The PAPA is gotten from the basis pursuit (BP) by solving [24], [31]

$$\min \|\tilde{\mathbf{w}}(n)\|_1 \quad s.t. \ \mathbf{d}(m) = \mathbf{X}^T(m)\tilde{\mathbf{w}}(n) \quad (18)$$

where $\tilde{\mathbf{w}}(n)$ denotes the correction component, which is given by

$$\tilde{\mathbf{w}}(m) = \tilde{\mathbf{G}}(m)\mathbf{X}(m)(\mathbf{X}^T(m)\tilde{\mathbf{G}}(m)\mathbf{X}(m))^{-1}\mathbf{d}(m). \quad (19)$$

If we consider $\tilde{\mathbf{G}}(m) \approx \mathbf{G}(m-1)$, a good approximation of (19) is written as

$$\tilde{\mathbf{w}}(n) = \mathbf{G}(m-1)\mathbf{X}(m) \left(\mathbf{X}^T(m)\mathbf{G}(m-1)\mathbf{X}(m) \right)^{-1} \mathbf{d}(m). \quad (20)$$

Then, a projection matrix R(m) is constructed for $\tilde{\mathbf{w}}(n)$ to span in its nullspace, which is given by

$$\mathbf{R}(m) = \mathbf{I}_K - \mathbf{G}(m-1)\mathbf{X}(m)$$
$$\times \left(\mathbf{X}^T(m)\mathbf{G}(m-1)\mathbf{X}(m) \right)^{-1} \mathbf{X}^T(m). \quad (21)$$

From the BS perspective, the updating equation of the CS-PAPA is

$$\hat{\mathbf{w}}(m) = \mathbf{R}(m)\hat{\mathbf{w}}(m-1) + \tilde{\mathbf{w}}(m-1), \quad (22)$$

which can be rewritten as

$$\hat{\mathbf{w}}(m) = \hat{\mathbf{w}}(m-1) + \mathbf{G}(m-1)\mathbf{X}(m)$$
$$\times \left(\mathbf{X}^T(m)\mathbf{G}(m-1)\mathbf{X}(m) \right)^{-1} \mathbf{e}(m). \quad (23)$$

The convergence factor μ is introduced into (23) to control the increment of the weigh vector during iteration. Moreover, a regularization parameter δ_{CS} is used in the denominator to avoid dividing by zero. Finally, the updating equation of the CS-PAPA is

$$\hat{\mathbf{w}}(m) = \hat{\mathbf{w}}(m-1) + \mu\mathbf{G}(m-1)\mathbf{X}(m)$$
$$\times \left(\mathbf{X}^T(m)\mathbf{G}(m-1)\mathbf{X}(m) + \delta_{CS}\mathbf{I} \right)^{-1} \mathbf{e}(m), \quad (24)$$

where the diagonal matrix $\mathbf{G}(m)$ is given by

$$\mathbf{G}(m-1) = \text{diag} \{g_1(m-1)\mathbf{1}_B, g_2(m-1)\mathbf{1}_B,$$
$$\cdots, g_N(m-1)\mathbf{1}_B\}, \quad (25)$$

where $\mathbf{1}_B$ is a row vector with B elements, and all of its elements are ones. Moreover, $g_s(m-1)$ is defined as

$$g_s(m-1) = \frac{\varphi_s(m-1)}{\sum_{i=1}^{N} \varphi_i(m-1)}, \quad 1 \le s \le N, \quad (26)$$

Algorithm 1 CS-PAPA Algorithm

Input: $K, P, B, \mu, \beta, \rho, \sigma$, in

Output: $\hat{\mathbf{w}}$ out

    Initialization :

1: $\hat{\mathbf{w}} = zeros(K, 1)$;

    Loop Process :

2: **for** $m=1, 2,...$ **do**

3:    $\mathbf{e}(m) = \mathbf{d}(m) - \mathbf{y}(m)$;

4:    **for** $i=1, 2,..., N$ **do**

5:       $\mathbf{q}_i(\hat{\mathbf{w}}) = 1 - e^{-\beta \|\hat{\mathbf{w}}_{[i]}\|_2}$;

6:    **end for**

7:    **for** $s=1, 2,..., N$ **do**

8:       $\varphi_s(m-1) = $
         $\max \left\{ \rho \max \left\{ \sigma, \mathbf{q}_1(\hat{\mathbf{w}}), \cdots, \mathbf{q}_N(\hat{\mathbf{w}}) \right\}, \mathbf{q}_s(\hat{\mathbf{w}}) \right\}$;

9:    **end for**

10:   **for** $j=1, 2,..., N$ **do**

11:      $g_j(m-1) = \dfrac{\varphi_j(m-1)}{\sum\limits_{i=1}^{N} \varphi_i(m-1)}$;

12:   **end for**

13:    $\mathbf{G}(m-1) = diag$;
     $\{g_1(m-1)\mathbf{1}_B, \cdots, g_N(m-1)\mathbf{1}_B\}$
     $\hat{\mathbf{w}}(m) = \hat{\mathbf{w}}(m-1)+$

14:    $\mu\mathbf{G}(m-1)\mathbf{X}(m)\left(\mathbf{X}^T(m)\mathbf{G}(m-1)\mathbf{X}(m) + \delta_{CS}\mathbf{I}\right)^{-1}$;
     $\mathbf{e}(m)$

15: **end for**

16: **return** $\hat{\mathbf{w}}$

---

where

$$\varphi_s(m-1) = \max \left\{ \rho \max \left\{ \sigma, \mathbf{q}_1(\hat{\mathbf{w}}), \mathbf{q}_2(\hat{\mathbf{w}}), \right.\right.$$
$$\left.\left. \cdots, \mathbf{q}_N(\hat{\mathbf{w}}) \right\}, \mathbf{q}_s(\hat{\mathbf{w}}) \right\}, \quad (27)$$

and

$$\mathbf{q}_i(\hat{\mathbf{w}}) = 1 - e^{-\beta \|\hat{\mathbf{w}}_{[i]}\|_2}, \quad (28)$$

where the time index is ignored in (27). The proposed CS-PAPA is summarized as **Algorithm 1**.

B. THE PROPOSED CS-IPAPA ALGORITHM

The CS-IPAPA is proposed by incorporating the $l_{2,0}$-norm into the cost function of the IPAPA. The updating equation of the CS-IPAPA is the same as that of CS-PAPA, and the difference is the gain assign matrix, which is

$$g_s(m) = \frac{1-\alpha}{2L} + (1+\alpha) \frac{1 - e^{-\beta \|\hat{\mathbf{w}}_{[s]}\|_2}}{2 \sum\limits_{i=1}^{N} (1 - e^{-\beta \|\hat{\mathbf{w}}_{[i]}\|_2}) + \varepsilon}. \quad (29)$$

In contrast to the PAPA and IPAPA, the proposed CS-PAPA and CS-IPAPA utilize only the priori information to define the cluster known for network echo and satellite communication channels. Furthermore, it is worth noting that the APA and PAPA are special cases of the proposed scheme.
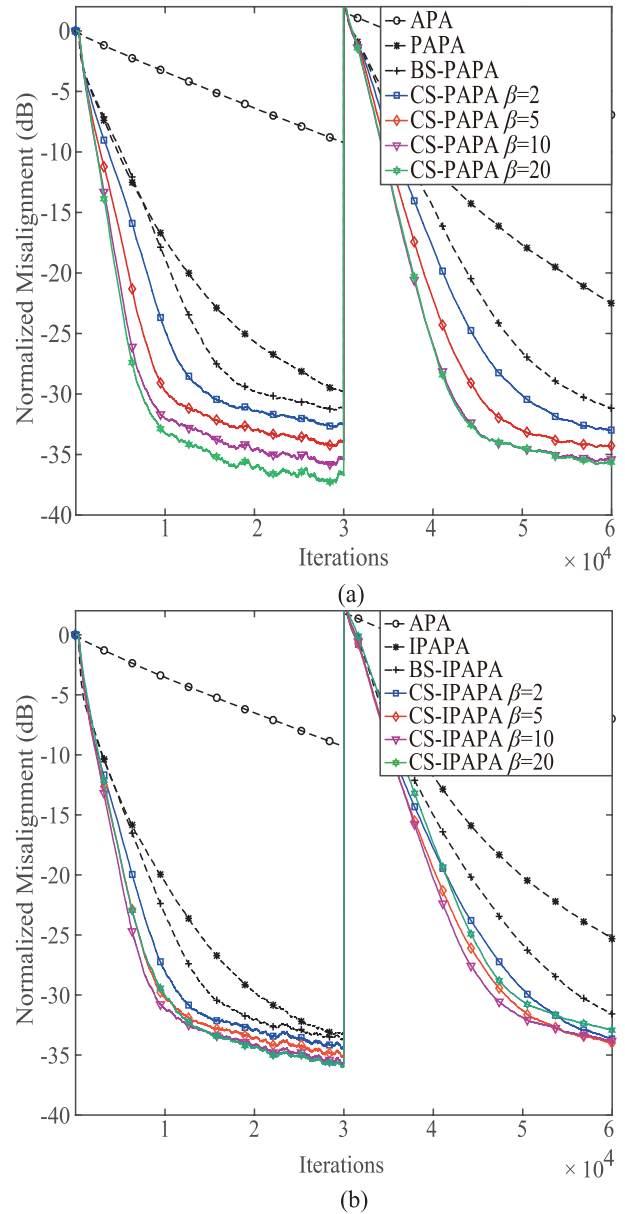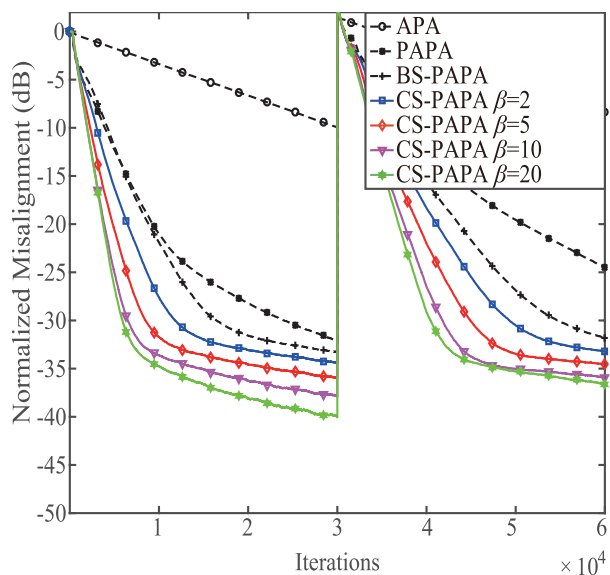

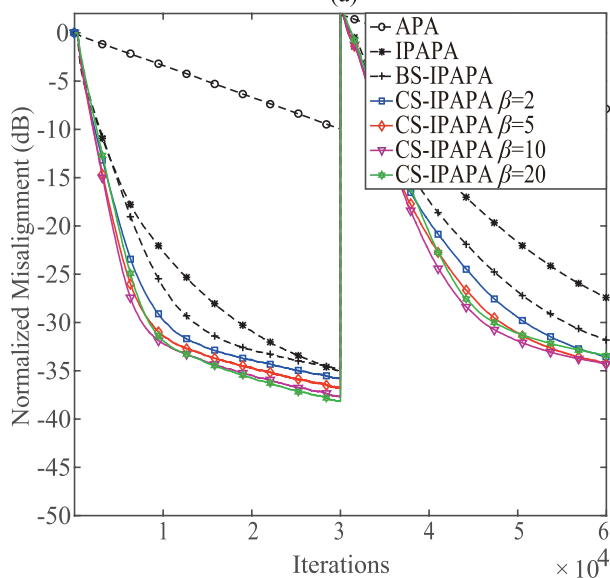**FIGURE 2.** Parameter β effects on the CS-PAPA and CS-IPAPA with colored input.

IV. SIMULATION RESULTS

To verify the performance of the CS-PAPA and CS-IPAPA, various experiments are conducted to assess their tracking behaviors, which are compared with the APA, PAPA, BS-PAPA and BS-IPAPA in the context of system identification. In all the experiments, the length of the adaptive filter is $K = 1024$, and all the mentioned algorithms are investigated in the presence of white Gaussian noise (WGN), colored noise and speech signals. The colored noise is obtained from WGN through a first-order system with a pole of 0.8, and the sampling frequency of the colored noise and speech signals is 8 kHz. Noise is added to the background of the unknown system at a signal-to-noise ratio (SNR) of 30 dB.
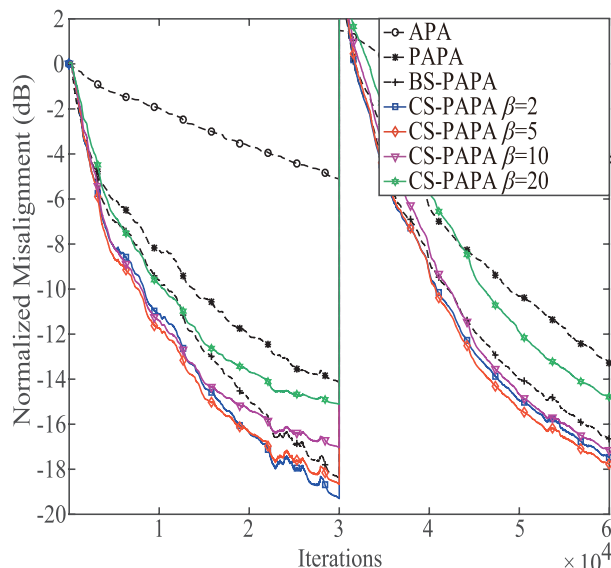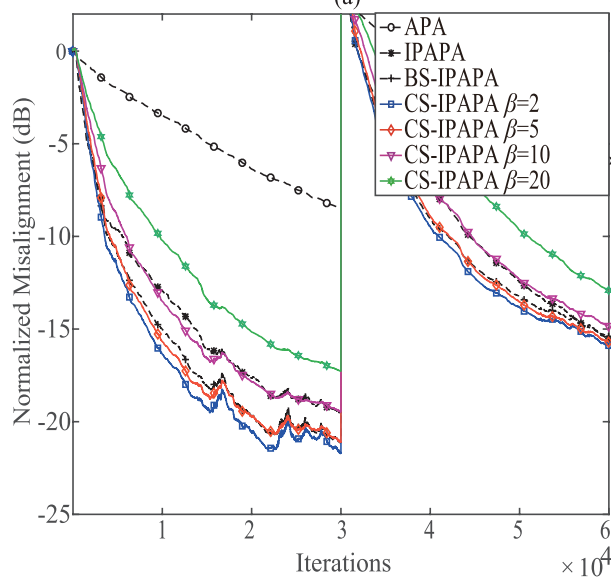
FIGURE 3. Parameter β effects on the CS-PAPA and CS-IPAPA with WGN input.

FIGURE 4. Parameter β effects on the CS-PAPA and CS-IPAPA with speech input.

We use normalized misalignment (NM) to measure the performance, and the NM is defined as $10 \lg(\|\mathbf{w} - \hat{\mathbf{w}}\|_2^2 / \|\mathbf{w}\|_2^2)$. The two cluster-sparse IR systems shown in Figure 1(a) and Figure 1(b) are considered. The first IR in Figure 1(a) has 32 nonzero taps and a single cluster with active coefficients distributed in [267], [298]. The second IR in Figure 1(b) has two clusters with active taps distributed at [267], [298] and [779], [810], which contain 64 nonzero taps. In all the experiments, the step sizes of the discussed algorithms are set to $\mu = 0.01$ when the input signal is WGN or colored noise, and the step size is set to $\mu = 0.02$ for speech inputs. $\varepsilon = 0.01$ and $a = 5/K$ are used in the related algorithms. The regularization parameters are $\delta_{\text{APA}} = 0.01$ and

TABLE 1. Simulation parameters.

	a	b	α	ρ	σ
PAPA	0.01	0.01	-	-	-
IPAPA	-	-	0	-	-
BS-PAPA	0.01	0.01	-	-	-
CS-PAPA	-	-	-	0.01	0.01
CS-IPAPA	-	-	0.5	-	-

$\delta_{\text{IPAPA}} = \frac{1-\alpha}{2} \delta_{\text{PAPA}} = \frac{1-\alpha}{2L} \delta_{\text{APA}}$ [32]. Other simulation parameters are shown in Table 1. In all the experiments, the first 30,000 iterations represent the performance of the single-cluster system, and the second 30,000 iterations indicate the performance of the double-cluster system.
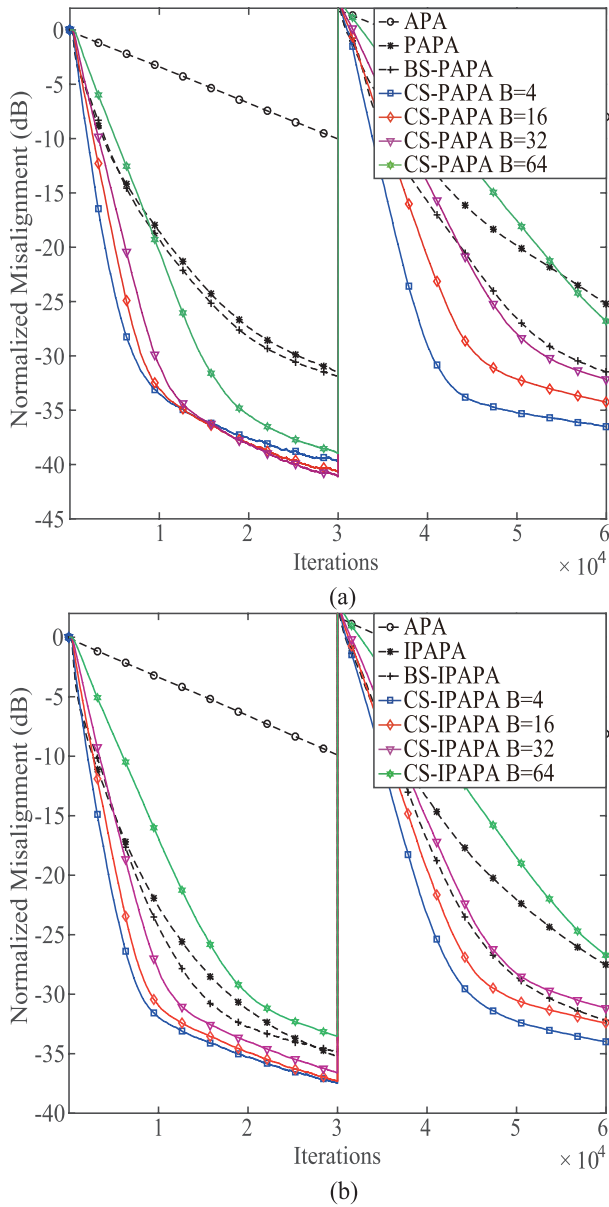
FIGURE 5. Parameter *B* effects on the CS-PAPA and CS-IPAPA with WGN input.

FIGURE 6. Parameter *B* effects on the CS-PAPA and CS-IPAPA with colored input.

A. PERFORMANCE OF THE CS-PAPA AND CS-IPAPA WITH DIFFERENT β

The effect of β on the performance of the CS-PAPA and CS-IPAPA is investigated with different inputs, namely, WGN, colored noise and speech signals. Herein, the cluster size is $B = 4$, and α is set to 0.5 for CS-IPAPA. The results with β values of 2, 5, 10 and 20 are presented in Figures 2, 3 and 4 for different inputs.

The experimental results presented in Figure 3(a) indicate that the NM of the CS-PAPA is reduced with the increase of β for both single-cluster and double-cluster systems when the input is a WGN signal. The CS-IPAPA algorithm shown in Figure 3(b) achieves the smallest NM for single-cluster systems when β is equal to 20, while the NM deteriorates when β is greater than 20. For the colored noise input signal
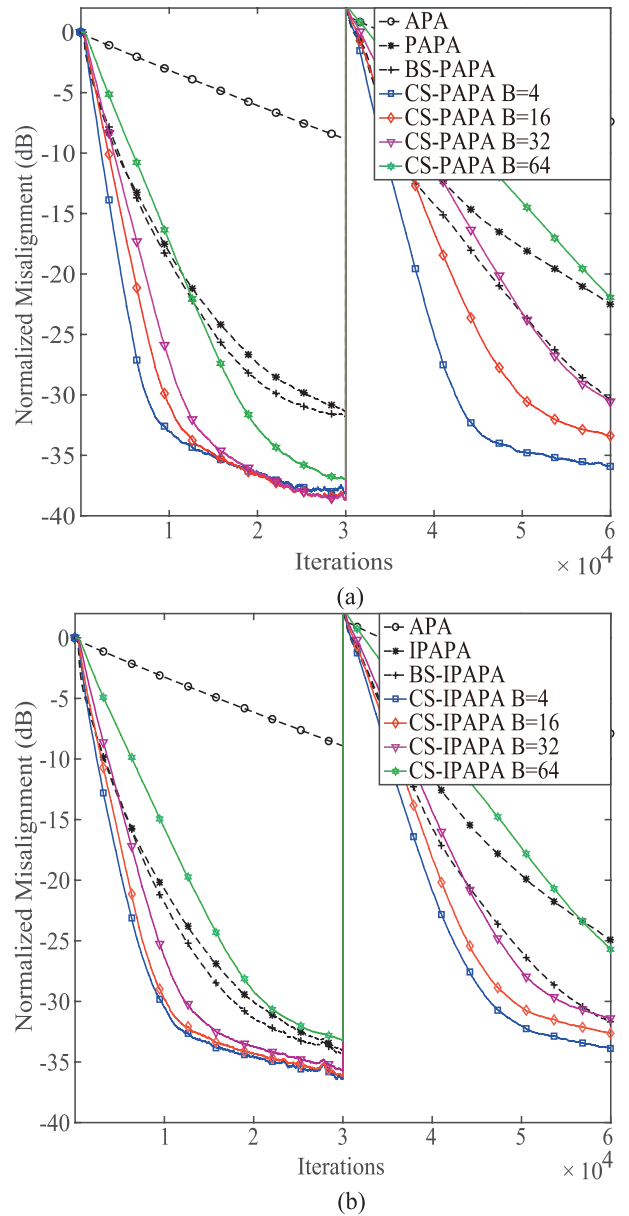
given in Figure 2, the CS-PAPA has the best NM for both the single- and double-cluster systems for $\beta = 20$. However, the CS-IPAPA has the lowest mean square error when β is equal to 10. If the inputs is a speech signal, the CS-PAPA and CS-IPAPA achieve the best performance with respect to the NM when β is small, as shown in Figure 4.

B. PERFORMANCE OF CS-PAPA AND CS-IPAPA WITH DIFFERENT B

The investigation of β showed that β has important effects on the NM of the CS-PAPA and CS-IPAPA algorithms. Another important parameter is B, which is used to determine the block size used in the proposed CS algorithms. Thus, four different block sizes with B=4, 16, 32, 64 are selected to
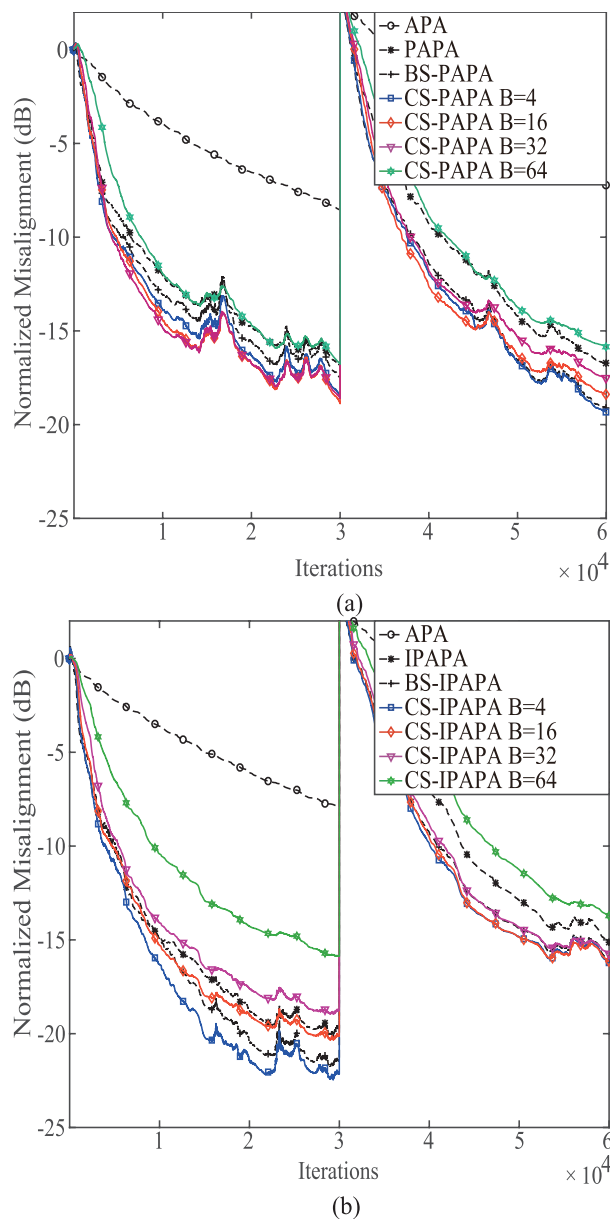
FIGURE 7. Parameter *B* effects on the CS-PAPA and CS-IPAPA with speech input.

evaluate the performance of the CS-PAPA and CS-IPAPA. In these experiments, β is set to 20 when the input signal is WGN or colored noise, and β is set to 2 when the input signal is a speech signal. The simulation results are given in Figures 5, 6 and 7. When the input signal is WGN and colored noise, the convergence and the NM of the CS-PAPA and CS-IPAPA deteriorate with increasing *B* for both the single- and double-cluster systems, as shown in Figures 5 and 6. When *B* = 4, the CS-PAPA and CS-IPAPA achieve the best performance. When the speech signal is the inputs, the CS-PAPA achieves the fastest convergence and lowest NM for *B* = 16, as shown in Figure 7(a). Figure 7(b) shows that the CS-IPAPA has smallest NM for *B* = 4. Thus, β and *B* should be carefully selected and optimized to achieve

superior performance with respect to the convergence and NM for both the CS-PAPA and CS-IPAPA. However, the proposed CS-PAPA and CS-IPAPA outperform the APA, PAPA and BS-IPAPA [23] for estimating single- and double-cluster channels.

V. CONCLUSIONS

In this paper, the CS-PAPA and CS-IPAPA have been proposed for cluster-sparse satellite and network echo channel estimation applications. The devised algorithms are implemented by using the mixed $l_{2,0}$-norm in the PAPA and IPAPA to make use of the prior cluster sparse information. Different β, *B* and inputs are taken into account to investigate the performance of the CS-PAPA and CS-IPAPA for estimating single- and double-cluster systems. The simulation results confirm that the proposed CS-PAPA and CS-IPAPA are superior to the APA, PAPA and the BS-IPAPA in various experiments.

REFERENCES

[1] E. Hansler and G. Schmidt, "Topics in acoustic echo and noise control," in *Signals and Communication Technology*. Berlin, Germany: Springer-Verlag, 2006.

[2] J. Benesty, T. Gansler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer, 2001.

[3] I. Song and P. Park, "A normalized least-mean-square algorithm based on variable-step-size recursion with innovative input data," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 817–820, Dec. 2012.

[4] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. New York, NY, USA: Springer, 2008.

[5] A. H. Sayed, *Fundamentals of Adaptive Fltering*. Hoboken, NJ, USA: Wiley, 2003.

[6] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 132–135, Feb. 2004.

[7] K. T. Wagner and M. I. Doroslovacki, "Complex proportionate-type normalized least mean square algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Kyoto, Japan, Jul. 2013, pp. 3285–3288.

[8] 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.

[9] H. Deng and M. Doroslovacki, "Proportionate adaptive algorithms for network echo cancellation," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1794–1803, May 2006.

[10] Y. Li, Y. Wang, and T. Jiang, "Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation," *Signal Process.*, vol. 128, pp. 243–251, 2016.

[11] Y. Li, Y. Wang, and T. Jiang, "Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation," *AEU-Int. J. Electron. Commun.*, vol. 70, no. 7, pp. 895–902, 2016.

[12] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Orlando, FL, USA, May 2002, pp. 1881–1884.

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

[14] L. Liu, M. Fukumoto, and S. Saiki, "An improved mu-law proportionate NLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Las Vegas, NV, USA, Mar. 2008, pp. 3797–3800.

[15] K. Wagner, M. Doroslovacki, and H. Deng, "Convergence of proportionate-type LMS adaptive filters and choice of gain matrix," in *Proc. 14th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Oct. 2006, pp. 243–247.

[16] P. Loganathan, E. A. P. Habets, and P. A. Naylor, "A partitioned block proportionate adaptive algorithm for acoustic echo cancellation," in *Proc. Asia–Pacific Signal Inf. Process. Assoc.*, Singapore, Dec. 2010, pp. 14–17.

[17] J. Liu and S. L. Grant, "Proportionate adaptive filtering for block-sparse system identification," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 4, pp. 623–630, Apr. 2016.

[18] T. Gansler, J. Benesty, S. L. Gay, and M. M. Sondhi, "A robust proportionate affine projection algorithm for network echo cancellation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Istanbul, Turkey, Jun. 2000, pp. 5–9.

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

[20] H. Zhao, Y. Yu, X. Zeng, and Z. He, "Memory proportionate APA with individual activation factors for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 22, no. 6, pp. 1047–1055, Jun. 2014.

[21] S. Werner, J. A. Apolinario, Jr., and P. S. R. Diniz, "Set-membership proportionate affine projection algorithms," *EURASIP J. Audio, Speech, Music Process.*, vol. 2007, no. 1, pp. 1–10, 2007.

[22] H. B. Yang, "A sparseness-controlled improved proportionate affine projection algorithm," in *Proc. 7th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Wuhan, China, Sep. 2011, pp. 1–4.

[23] J. Liu and S. L. Grant, "Proportionate affine projection algorithms for block-sparse system identification," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Shanghai, China, May 2016, pp. 529–533.

[24] J. Benesty, C. Paleologu, and S. Ciochina, "Proportionate adaptive filters Grom a basis pursuit perspective," *IEEE Signal Process. Lett.*, vol. 17, no. 12, pp. 985–988, Dec. 2010.

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

[26] T. Dragan, "Echo cancellation device with mobile crosstalk control," WO Patent 2 004 054 127 A1, Jun. 24, 2004.

[27] P. A. C. Marques, F. M. G. Sousa, and J. M. N. Leitao, "A DSP based long distance echo canceller using short length centered adaptive filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Munich, Germany, Apr. 1997, pp. 1885–1888.

[28] P. J. McLane, P. H. Wittke, P. K.-M. Ho, and C. Loo, "PSK and DPSK trellis codes for fast fading, shadowed mobile satellite communication channels," *IEEE Trans. Commun.*, vol. COMM-36, no. 11, pp. 1242–1246, Nov. 1988.

[29] K. Sakhnov, "An improved proportionate affine projection algorithm for network echo cancellation," in *Proc. 15th Int. Conf. Syst., Signals Image Process.*, Bratislava, Slovakia, Jun. 2008, pp. 125–128.

[30] Y. Gu, J. Jin, and S. Mei, "ℓ_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.

[31] C. Paleologu and J. Benesty, "Proportionate affine projection algorithms from a basis pursuit perspective," in *Proc. IEEE Int. Symp. Circuits Syst.*, Seoul, South Korea, May 2012, pp. 2757–2760.

[32] J. Benesty, C. Paleologu, and S. Ciochina, "On regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 6, pp. 1734–1742, Aug. 2011.

ZHENGXIONG JIANG received the B.S. degree in electrical and information engineering from Harbin Engineering University, Harbin, China, in 2016, where he is currently pursuing the M.E. degree in information and communication engineering.

OMER MOHAMMED OMER OSMAN was born in El-Dweem, Sudan, in 1976. He received the B.S. degree in mathematics and computer science from Al-Neelain University in 2000, the master's degree in computer science in 2006, and the master's degree in data and communication networks. He is currently pursuing the Ph.D. degree with Harbin Engineering University, and his project is sparse signal processing. He has been an Engineer at Al-neelain University since 2006, and a Lecturer at many universities in Sudan.

XIAO HAN received the B.S., M.S., and Ph.D. degrees in underwater acoustic engineering from Harbin Engineering University, Harbin, China, in 2011, 2014, and 2016, respectively, and the Ph.D. degree from Harbin Engineering University in 2016. In 2018, he joined the ninth Xuelong Arctic Expedition. He is currently an Associate Professor with the College of Underwater Acoustic Engineering, Harbin Engineering University. He has authored numerous research articles published in related international conference proceedings and journal. He is the holder of nearly 10 patents. His research interests include underwater acoustic signal propagation, ocean acoustics, digital communication theory acoustic communication in under-ice environment, and embedded systems. Most of his work focuses on channel estimation, tracking and equalization in time domain based on passive time reversal and decision feedback equalizer in underwater environment. Currently, his research is mainly concentrated in under-ice acoustics communication in shallow water.

YINGSONG LI (M'14) received the B.S. degree in electrical and information engineering, and the M.S. degree in electromagnetic field and microwave technology from Harbin Engineering University, in 2006 and 2011, respectively, and the Ph.D. degree from both the Kochi University of Technology (KUT), Japan, and Harbin Engineering University, China, in 2014. He has been a Full Professor with the Harbin Engineering University since 2014. He is a Visiting Scholar with the University of California at Davis, Davis, CA, USA, from 2016 to 2017, and the University of York, U.K., in 2018. He is also a Visiting Professor at the Far Eastern Federal University and KUT. He is a Senior Member of the Chinese Institute of Electronics. He is an Associate Editor of the IEEE Access and the *Applied Computational Electromagnetics Society Journal*. He is also an Area Editor of the *AEÜ-International Journal of Electronics and Communications*.

Recently, his recent research interests are mainly in remote sensing, underwater communications, signal processing, radar, SAR imaging, compressed sensing, and antennas. He also serves as a reviewer for over 20 journals.

JINGWEI YIN received the B.S., M.S., and Ph.D. degrees in underwater acoustic engineering from Harbin Engineering University, China, in 1999, 2006, and 2007, respectively. He is a Visiting Professor at the Russian Far Eastern Federal University. He is the Deputy Director of the Water Acoustic Branch of the China Acoustic Society. He is also the Director of the Key Laboratory of Marine Information Acquisition and Security, Ministry of Industry and Information Technology, Harbin Engineering University. He is currently a Professor and also the Dean of the College of Underwater Acoustic Engineering, Harbin Engineering University. He has published a monograph, over 100 papers, and more than 20 inventions. His current research interests include signal processing for acoustic, underwater acoustic engineering, underwater acoustic communication, and polar acoustics.

● ● ●