Noise Robust Frequency-Domain Adaptive Blind Multichannel Identification With ℓ_p-Norm Constraint

Hongsen He, *Member, IEEE*, Jingdong Chen, *Senior Member, IEEE*, Jacob Benesty, and Tao Yang

***Abstract*—Blind multichannel identification is a challenging problem in many domains. The normalized multichannel frequency-domain least-mean-square (NMCFLMS) algorithm was developed to blindly identify a single-input multiple-output acoustic system, which can yield good performance in noise-free environments. However, the robustness of this algorithm to noise has been shown to be problematic. One way to improve the robustness is by applying a constraint on the spectral flatness of the channel impulse responses, which led to the development of the so-called robust normalized multichannel frequency-domain least-mean-square (RNMCFLMS) algorithm. This spectral flatness constraint, however, may not be always proper or reasonable in realistic acoustic environments. In this paper, we develop an ℓ_p-norm constraint based robust normalized multichannel frequency-domain least-mean-square (ℓ_p-RNMCFLMS) algorithm. The ℓ_p-norm constraint is introduced into the NMCFLMS algorithm to control the effect of different ℓ_p-norm penalties on the adaptive filter for the impulse responses with different degrees of sparseness. Numerical and realistic experiments justify the effectiveness of the proposed ℓ_p-RNMCFLMS algorithm.**

***Index Terms*—Blind multichannel identification, frequency-domain adaptive filtering, ℓ_p-norm penalty, sparsity, SIMO system, robustness.**

I. Introduction

BLIND multichannel identification (BMCI), which aims at estimating the channel impulse responses of an unknown system based only on the output signals, plays an important role in many application domains, such as multimedia signal processing [1], geophysical exploration [2], communications [3], etc. Over the last few decades, the BMCI problem has drawn a significant amount of research attention and many algorithms have been developed, such as the subspace algorithms [4]–[6], cross-relation algorithms [7], [8], sub-channel matching methods [8], [9], higher-order statistics approaches [10], [11], maximum likelihood algorithms [12]–[14], error-property based identification methods [15], [16], and the normalized multichannel frequency-domain least-mean-square (NMCFLMS) algorithm [17].

Among those algorithms, NMCFLMS is especially attractive for real-world applications since it exploits the fast Fourier transform (FFT) to adaptively identify the impulse responses of a single-input multiple-output (SIMO) system in the frequency domain and, therefore, is computationally very efficient. However, this algorithm was found not to be robust to additive noise. An improved version of the NMCFLMS algorithm, called robust NMCFLMS (RNMCFLMS), was then developed [18], [19], which introduces a logarithmic penalty on the spectra of the channel impulse responses into the NMCFLMS algorithm to make it immune to additive noise. The logarithmic function as a penalty term is effective to promote the flatness of the spectra of the sparse impulse responses. In most of practical acoustic environments, however, the spectra of the acoustic channel impulse responses are not necessarily flat and, as a result, the robustness of the RNMCFLMS algorithm to noise is far from being satisfactory even though it is better than that of the NMCFLMS algorithm.

In this study, we propose an ℓ_p-norm based noise robust normalized multichannel frequency-domain least-mean-square (ℓ_p-RNMCFLMS) algorithm. Unlike the RNMCFLMS algorithm that uses a logarithmic function as the penalty on the spectra of the acoustic channel impulse responses, the presented algorithm employs the ℓ_p-norm as the penalty term in the cost function of the multichannel frequency-domain adaptive filter. The update equations of the frequency-domain adaptive filter are rigorously derived upon the basis of Newton's method. Furthermore, we will also investigate how different ℓ_p-norm penalties affect the performance of the proposed adaptive filter in acoustic environments with different degrees of reverberation, which affects the level of sparsity of the acoustic channel impulse responses. The effect of different ℓ_p-norm penalties on BMCI will be verified through numerical simulations as well as realistic experiments.

Manuscript received November 27, 2017; revised March 23, 2018; accepted May 7, 2018. Date of publication May 11, 2018; date of current version May 25, 2018. The work of H. He was supported in part by the National Science Foundation of China under Grant 61571376 and in part by the Open Foundation of the State Key Laboratory of Acoustics, Chinese Academy of Sciences, under Grant SKLA201712. The work of J. Chen was supported in part by the NSFC "Distinguished Young Scientists Fund" under Grant 61425005 and in part by the NSFC-ISF joint research program under Grant 61761146001. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Huseyin Hacihabiboglu. *(Corresponding author: Hongsen He.)*

H. He is with the School of Information Engineering and Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, Southwest University of Science and Technology, Mianyang 621010, China, and also with the State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China (e-mail: hongsenhe@gmail.com).

J. Chen is with the Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: jingdongchen@ieee.org).

J. Benesty is with the INRS-EMT, University of Quebec, Montreal QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

T. Yang is with the School of Information Engineering and Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, Southwest University of Science and Technology, Mianyang 621010, China (e-mail: yangtao98@tsinghua.org.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASLP.2018.2835729

The rest of this paper is organized as follows. In Section II, we formulate the basic problem and briefly review the existing solutions. Section III first derives the ℓ_p-RNMCFLMS algorithm, then analyzes the effect of different norm penalties on the optimization of the frequency-domain adaptive filter. In Section IV, we examine the convergence behavior of the ℓ_p-RNMCFLMS algorithm by simulations. Section V evaluates the effectiveness of the ℓ_p-RNMCFLMS algorithm through experiments in realistic acoustic environments. Finally, some conclusions are drawn in Section VI.

II. Problem Formulation and the Existing Solutions

A. Notation

For ease of reading, we first give the notation used in this paper. Lowercase and uppercase bold letters denote, respectively, time-domain vectors and matrices. Vectors and matrices in the frequency domain are represented, respectively, by underlined lowercase bold italic and uppercase calligraphic letters. Some illustrative examples are given below:

- $\mathbf{x}$: vector in the time domain (bold, lowercase),
- $\mathbf{X}$: matrix in the time domain (bold, uppercase),
- $\underline{\boldsymbol{x}}$: vector in the frequency domain (bold, italic, underlined, lowercase),
- $\boldsymbol{\mathcal{X}}$: matrix in the frequency domain (calligraphic, bold, uppercase).

To connect a time-domain vector with its frequency-domain counterpart, we define $\boldsymbol{\mathcal{F}}_L$ as the Fourier matrix of size $L \times L$, whose (p,q)th element is

$$(\boldsymbol{\mathcal{F}}_L)_{p,q} = e^{-\jmath 2\pi (p-1)(q-1)/L}, \quad p,q = 1,2,\ldots,L, \tag{1}$$

with $\jmath$ being the imaginary unit and $\jmath^2 = -1$. The inverse of $\boldsymbol{\mathcal{F}}_L$ is denoted $\boldsymbol{\mathcal{F}}_L^{-1}$. In implementation, the operators $\boldsymbol{\mathcal{F}}_L$ and $\boldsymbol{\mathcal{F}}_L^{-1}$ are calculated with FFT. The operators $(\cdot)^T$ and $(\cdot)^H$ stand for the transpose and Hermitian transpose of a vector or matrix, respectively, $(\cdot)^*$ and ∇ denote the complex conjugate and gradient operator, respectively, $*$ stands for linear convolution, and $\odot$ denotes element-by-element multiplication of two vectors.

B. Problem Formulation

Assume that an acoustic SIMO system is composed of a sound source and M microphones. The sound signal captured by the ith ($i = 1,2,\ldots,M$) microphone is then written as

$$x_i(n) = s(n) * h_i(n) + v_i(n), \tag{2}$$

where $s(n)$ is the source signal, $h_i(n)$ is the channel impulse response between the sound source and the ith microphone, which is typically modeled with a finite-impulse-response (FIR) filter, and $v_i(n)$ is the additive noise at the ith microphone. It is assumed that all signals are zero mean. If we neglect the noise term in (2), the following relation can be obtained for any pair of microphone signals:

$$\begin{aligned} x_i(n) * h_j(n) &= s(n) * h_i(n) * h_j(n) \\ &= x_j(n) * h_i(n), \ i,j = 1,2,\ldots,M, \ i \neq j, \end{aligned} \tag{3}$$

which can be expressed in a matrix-vector form as

$$\mathbf{x}_i^T(n)\mathbf{h}_j(n) - \mathbf{x}_j^T(n)\mathbf{h}_i(n) = 0, \tag{4}$$

where

$$\mathbf{h}_i(n) = [h_{i,0}(n) \ \ h_{i,1}(n) \ \cdots \ h_{i,L-1}(n)]^T \ \ i = 1,2,\ldots,M \tag{5}$$

are the impulse response vectors of length L, and

$$\mathbf{x}_i(n) = [x_i(n) \ x_i(n-1) \ \cdots \ x_i(n-L+1)]^T \tag{6}$$

is the convolved signal vector at the ith sensor.

When noise is present and/or the estimate of the impulse responses deviates from the true value, the right-hand side of (4) is no longer zero and an *a priori* error signal between the ith and jth channels is produced as

$$e_{ij}(n) = \mathbf{x}_i^T(n)\widehat{\mathbf{h}}_j(n) - \mathbf{x}_j^T(n)\widehat{\mathbf{h}}_i(n), \tag{7}$$

where $\widehat{\mathbf{h}}_i(n)$ is an estimate of $\mathbf{h}_i(n)$ at time n. This error signal can then be used to define a cost function (with some appropriate constraints) that should be minimized to find an optimal estimate of the impulse responses.

C. Existing Solutions

The NMCFLMS algorithm employs the sum of the squared instantaneous errors between different channels to define the cost function in the frequency domain [17]. Using the mth block of the error signal $e_{ij}(n)$, i.e.,

$$\mathbf{e}_{ij}(m) = [e_{ij}(mL) \ e_{ij}(mL+1) \ \cdots \ e_{ij}(mL+L-1)]^T, \tag{8}$$

the (frequency-domain) cost function of the NMCFLMS algorithm is then defined as

$$\mathcal{J}_{\mathrm{F}}(m) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \underline{\boldsymbol{e}}_{ij}^H(m)\underline{\boldsymbol{e}}_{ij}(m), \tag{9}$$

where

$$\begin{aligned} \underline{\boldsymbol{e}}_{ij}(m) &= \boldsymbol{\mathcal{F}}_L \mathbf{e}_{ij}(m) \\ &= \boldsymbol{\mathcal{G}}\left[\boldsymbol{\mathcal{D}}_i(m)\boldsymbol{\mathcal{W}}\underline{\widehat{\boldsymbol{h}}}_j(m) - \boldsymbol{\mathcal{D}}_j(m)\boldsymbol{\mathcal{W}}\underline{\widehat{\boldsymbol{h}}}_i(m)\right] \\ &= \left[\underline{e}_{ij,0}(m) \ \underline{e}_{ij,1}(m) \ \cdots \ \underline{e}_{ij,L-1}(m)\right]^T, \end{aligned} \tag{10}$$

$$\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{F}}_L \begin{bmatrix} \mathbf{O}_L & \mathbf{I}_L \end{bmatrix} \boldsymbol{\mathcal{F}}_{2L}^{-1}, \tag{11}$$

$$\boldsymbol{\mathcal{D}}_i(m) = \mathrm{diag}\left[\boldsymbol{\mathcal{F}}_{2L}\mathbf{x}_i(m)\right], \tag{12}$$

$$\begin{aligned} \mathbf{x}_i(m) = \big[&x_i(mL-L) \ x_i(mL-L+1) \\ &\cdots \ x_i(mL+L-1)\big]^T, \end{aligned} \tag{13}$$

$$\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{F}}_{2L} \begin{bmatrix} \mathbf{I}_L & \mathbf{O}_L \end{bmatrix}^T \boldsymbol{\mathcal{F}}_L^{-1}, \tag{14}$$

$$\widehat{\underline{\mathbf{h}}}_j(m) = \boldsymbol{\mathcal{F}}_L \widehat{\mathbf{h}}_j(m) = \left[\widehat{\underline{h}}_{j,0}(m)\ \widehat{\underline{h}}_{j,1}(m)\ \cdots\ \widehat{\underline{h}}_{j,L-1}(m)\right]^T, \quad (15)$$

$\mathbf{O}_L$ is the null matrix of size $L \times L$, $\mathbf{I}_L$ is the identity matrix of size $L \times L$, and diag$[\cdot]$ denotes a diagonal matrix with indicated vector along the diagonal. According to Newton's iteration method, the update equations of the NMCFLMS algorithm are then derived as [17]

$$\widehat{\underline{\mathbf{h}}}_k^{10}(m+1) = \widehat{\underline{\mathbf{h}}}_k^{10}(m) - \mu_{\mathrm{f}} \boldsymbol{\mathcal{P}}_k^{-1}(m) \sum_{i=1}^{M} \boldsymbol{\mathcal{D}}_i^*(m) \underline{\mathbf{e}}_{ik}^{01}(m), \quad k = 1, 2, \ldots, M, \quad (16)$$

where μ_{f} is the step size, and

$$\widehat{\underline{\mathbf{h}}}_k^{10}(m) = \boldsymbol{\mathcal{F}}_{2L} \widehat{\mathbf{h}}_k^{10}(m), \quad (17)$$

$$\widehat{\mathbf{h}}_k^{10}(m) = \left[\widehat{\mathbf{h}}_k^T(m)\ \mathbf{0}\right]^T, \quad (18)$$

$$\boldsymbol{\mathcal{P}}_k(m) = \sum_{i=1, i\neq k}^{M} \boldsymbol{\mathcal{D}}_i^*(m) \boldsymbol{\mathcal{D}}_i(m), \quad (19)$$

$$\underline{\mathbf{e}}_{ik}^{01}(m) = \boldsymbol{\mathcal{F}}_{2L} \left[\mathbf{0}\ \mathbf{e}_{ik}^T(m)\right]^T, \quad (20)$$

and $\mathbf{0}$ is the null matrix of size $1 \times L$.

The NMCFLMS algorithm can achieve good estimation performance in very high signal-to-noise-ratio (SNR) environments. However, the performance of this algorithm deteriorates significantly in low SNR cases. To cope with this problem, the RNMCFLMS algorithm was proposed [18], [19], which introduces a logarithmic penalty on the fullband spectral energy into the NMCFLMS algorithm. The corresponding cost function is defined as

$$\mathcal{J}_{\mathrm{R}}(m) = \mathcal{J}_{\mathrm{F}}(m) - \rho(m) \sum_{i=1}^{M} \sum_{j=0}^{L-1} \log\left(\left|\widehat{\underline{h}}_{i,j}(m)\right|^2\right), \quad (21)$$

where $\rho(m)$ is a Lagrange multiplier and the second term on the right-hand side of the equation is the spectral penalty function on the acoustic channel impulse responses. Based on NMCFLMS, the update equations of the RNMCFLMS algorithm are as follows [18], [19]:

$$\widehat{\underline{\mathbf{h}}}_k^{10}(m+1) = \widehat{\underline{\mathbf{h}}}_k^{10}(m) - \mu_{\mathrm{f}} \boldsymbol{\mathcal{P}}_k^{-1}(m) \sum_{i=1}^{M} \boldsymbol{\mathcal{D}}_i^*(m) \underline{\mathbf{e}}_{ik}^{01}(m) + \mu_{\mathrm{f}} \rho(m) \nabla \mathcal{J}_k^{10}(m), \quad k = 1, 2, \ldots, M, \quad (22)$$

where $\nabla \mathcal{J}_k^{10}(m)$ is the gradient of the spectral penalty function with respect to the filter coefficient vector. The spectral constraint introduced by the RNMCFLMS algorithm ensures the spectral flatness of the estimated channel impulse responses in the presence of noise [18], [19]. Thus, the RNMCFLMS algorithm is much more immune to noise than the NMCFLMS algorithm, particularly in the scenario where the impulse responses are impulse like and their spectra are flat.

However, if the spectra of the impulse responses are not flat, which is often true in reverberant acoustic environments, the performance of the RNMCFLMS algorithm is still very sensitive to noise. Unlike the NMCFLMS algorithm, the update equations of RNMCFLMS are derived via Newton's method, but through a heuristic manner, which leads to some difficulty in analyzing its performance with mathematical rigor. In this paper, we develop a noise robust frequency-domain blind multichannel identification algorithm based on an ℓ_p-norm constraint. According to the proposed cost function, we give a detailed deduction process of the frequency-domain adaptive filter with Newton's method. We will investigate the effect of different sparse-level penalties on the identification of the acoustic channels in different reverberation conditions.

III. Noise Robust Frequency-Domain Adaptive Blind Multichannel Identification Algorithm

For an acoustic system, the output signal is equal to the input signal convolved with the corresponding impulse response in the time domain. The characteristics of the acoustic channel impulse response are affected by the shape and size of the room as well as the reflection intensity of all the boundaries (walls, floor, ceiling, etc). A room acoustic impulse response may have a long and heavy tail (corresponding to a room with long reverberation time), or a sparse tail (corresponding to a room with short reverberation time). In the frequency domain, the impulse responses with different levels of tails exhibit diverse uniformities of amplitude spectra. This section investigates how to blindly identify the impulse responses of the acoustic channels by taking into account the sparsity of the impulse responses.

A. An ℓ_p-RNMCFLMS Algorithm

The ℓ_0-norm is often viewed as an accurate metric to measure the sparsity. However, the minimization of the corresponding cost function is non-deterministic polynomial-time (NP)-hard. To relax the NP-hard problem, the ℓ_1-norm is often used, which may lead to similar performance as the ℓ_0-norm [20]. For realistic acoustic channels, however, the ℓ_1-norm constraint may not be optimal in every condition as the sparsity of the room impulse responses varies significantly from one application to another. Inspired by the bridge regression theory [21], which introduces an ℓ_p-norm penalty ($1 \leq p < 2$) into the least-squares criterion so as to achieve a compromise between the lasso ($p = 1$) and ridge regression ($p = 2$), we attempt in this work to employ an ℓ_p-norm constraint of acoustic impulse responses and investigate the effect of different ℓ_p-norm penalties on the adaptive filters for the acoustic channels with different levels of reverberation. One straightforward way of doing this is through the following cost function, which is a combination of the cost function defined in (9) and the ℓ_p-norm of the estimate of the time-domain impulse response as follows:

$$\mathcal{J}(m) = \mathcal{J}_{\mathrm{F}}(m) + \eta(m) \mathcal{J}_{\mathrm{TP}}(m) \quad (23)$$

$$\text{s.t.} \left\|\widehat{\mathbf{h}}(m)\right\|_{\ell_2}^2 = 1, \quad (24)$$

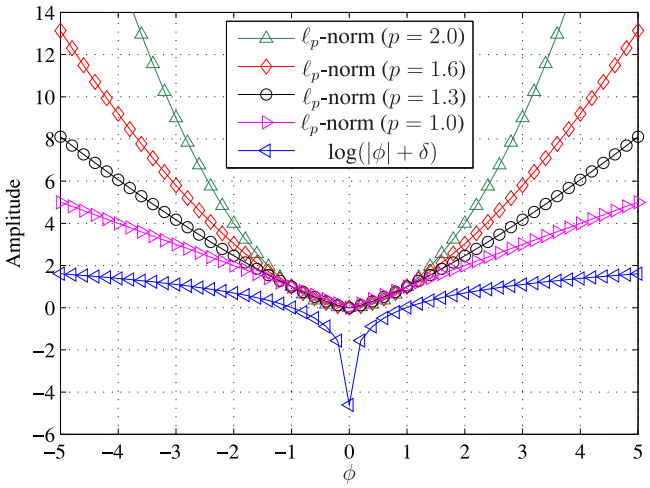

Fig. 1. Comparison of the p-norm ($1 \leq p < 2$) penalty function and logarithmic penalty function ($\delta = 0.01$).

where $\mathcal{J}_{\mathrm{F}}(m)$ is defined in (9), $\eta(m) > 0$ is a Lagrange multiplier,

$$\mathcal{J}_{\mathrm{TP}}(m) = \left\|\widehat{\mathbf{h}}(m)\right\|_{\ell_p}^{p}, \tag{25}$$

$$\widehat{\mathbf{h}}(m) = \left[\widehat{\mathbf{h}}_1^T(m)\ \widehat{\mathbf{h}}_2^T(m)\ \cdots\ \widehat{\mathbf{h}}_M^T(m)\right]^T, \tag{26}$$

$\|\cdot\|_{\ell_p}$ denotes the ℓ_p-norm ($1 \leq p < 2$), which is a penalty constraint, and $\|\cdot\|_{\ell_2}$ denotes the ℓ_2-norm. Equation (24) is a unit length constraint on the time-domain impulse response vector to avoid a trivial solution with all zero elements for the time-domain impulse responses. To simplify, this unit length constraint is imposed on the time-domain impulse responses of acoustic channels after each iteration update. Fig. 1 illustrates the difference between ℓ_p-norms of a scalar variable ϕ with different values of p. For comparison, we also plot the ℓ_2-norm and logarithmic function. From an optimization point of view, the penalty function with a small value of p ($1 \leq p < 2$) has lower emphasis on large magnitude and sharper slope near zero as compared to the case of a large value of p, which can be clearly seen from Fig. 1. According to the connection between Bayes estimator and ℓ_p ($1 \leq p < 2$) penalty [22], this ℓ_p penalty is a generalization of both the lasso ($p = 1$) and ridge regression ($p = 2$). The former corresponds to a Laplacian prior, while the latter corresponds to a Gaussian prior. Therefore, the ℓ_p-norm ($1 \leq p < 2$) can effectively penalize the different sparse levels of acoustic channels that correspond to different distribution characteristics, respectively.

While it is straightforward to follow, the cost function defined in (23) is a mixture of the frequency-domain least-squares error and the time-domain sparsity penalty, which makes it difficult to deduce a computationally efficient frequency-domain adaptive algorithm via Newton's method. One may consider to deduce an adaptive algorithm following the gradient descent method; but the resulting adaptive filter is slow in convergence, particularly in acoustic applications where the filter length is usually very long. To circumvent this issue, we adopt an alternative way to define the cost function, which uses the density of the amplitude spectrum of the acoustic impulse responses as the penalty term, i,e.,

$$\widetilde{\mathcal{J}}(m) = \mathcal{J}_{\mathrm{F}}(m) - \eta(m)\mathcal{J}_{\mathrm{FP}}(m) \tag{27}$$

$$\text{s.t.}\ \left\|\widehat{\mathbf{h}}(m)\right\|_{\ell_2}^2 = 1, \tag{28}$$

where

$$\mathcal{J}_{\mathrm{FP}}(m) = \left\|\widehat{\underline{\boldsymbol{h}}}^{10}(m)\right\|_{\ell_p}^{p}, \tag{29}$$

$$\widehat{\underline{\boldsymbol{h}}}^{10}(m) = \left[\left(\widehat{\underline{\boldsymbol{h}}}_1^{10}(m)\right)^T\ \left(\widehat{\underline{\boldsymbol{h}}}_2^{10}(m)\right)^T\ \cdots\ \left(\widehat{\underline{\boldsymbol{h}}}_M^{10}(m)\right)^T\right]^T, \tag{30}$$

$$\widehat{\underline{\boldsymbol{h}}}_k^{10}(m) = \left[\widehat{\underline{h}}_{k,0}^{10}(m)\ \widehat{\underline{h}}_{k,1}^{10}(m)\ \cdots\ \widehat{\underline{h}}_{k,2L-1}^{10}(m)\right]^T,$$
$$k = 1, 2, \ldots, M. \tag{31}$$

Herein, we impose a penalty on $\widehat{\underline{\boldsymbol{h}}}^{10}(m)$ rather than $\widehat{\underline{\boldsymbol{h}}}(m)$. This is due to the fact that the modeling filter $\widehat{\mathbf{h}}_k(m)$ is padded with trailing zeros to length $2L$, which yields a zoomed spectrum to alleviate picket fence effect and emphasize more spectral components of $\widehat{\mathbf{h}}_k(m)$. Comparing (23) and (27), one can see that minimization of the sparsity penalty in the time domain is now transformed into maximization of the homogenization penalty in the frequency domain, which results in the minus sign in the second term in (27).

Remark 1: It should be pointed out that direct minimization of (27) may lead to negative infinity due to the minus term on the right-hand side of (27). To circumvent this issue so as to effectively optimize the cost function, we apply the following constraints: 1) the number of microphone channels is finite; 2) the sum of the square of the coefficients of the acoustic channel impulse responses is finite, which is true in realistic acoustic environments; 3) the parameter p is bounded as $1 \leq p < 2$; and 4) $\eta(m)$ is a finite value, which is adaptively updated so that during each iteration, the increment of the frequency-domain adaptive filter coefficient vector is not significant. Under the above conditions, the adaptive filter can effectively search for an optimal solution in the constrained solution space.

Remark 2: An alternative scheme to avoid the use of the minus term on the right-hand side of (27) is to introduce an ℓ_q-norm as the spectral penalty on the acoustic channel impulse responses, i.e.,

$$\widetilde{\mathcal{J}}(m) = \mathcal{J}_{\mathrm{F}}(m) + \eta(m)\mathcal{J}_{\mathrm{FQ}}(m), \tag{32}$$

where

$$\mathcal{J}_{\mathrm{FQ}}(m) = \left\|\widehat{\underline{\boldsymbol{h}}}^{10}(m)\right\|_{\ell_q}^{q}, \tag{33}$$

$\|\cdot\|_{\ell_q}$ denotes the ℓ_q-norm, which is defined similarly as the ℓ_p-norm, but the value of q is determined as follows:

$$\frac{1}{q} + \frac{1}{p} = 1, \tag{34}$$

which indicates that

$$q = \frac{p}{p-1}. \tag{35}$$

According to (35), one can see that if $1 \leq p < 2$, the value of q satisfies $2 < q < \infty$. Our experimental investigation shows that the $\|\cdot\|_{\ell_q}^q$ can help improve the performance of the adaptive filter in low SNR environments with speech excitation, but it easily makes the adaptive filter divergent in high SNR conditions. The underlying reason is that in high SNR conditions, the amplitude spectra of microphone signals are less flat and have a large dynamic range and so is the error signal. In this case, large components of the error signal are easily exacerbated by the ℓ_q-norm ($2 < q < \infty$) penalty, resulting divergence of the adaptive filter. Therefore, this approach will not be discussed in detail in this paper.

To deduce the adaptive filtering algorithm, we first express (10) as the following equivalent equation:

$$\underline{\boldsymbol{e}}_{ij}(m) = \boldsymbol{\mathcal{G}}\left[\boldsymbol{\mathcal{D}}_i(m)\widehat{\underline{\boldsymbol{h}}}_j^{10}(m) - \boldsymbol{\mathcal{D}}_j(m)\widehat{\underline{\boldsymbol{h}}}_i^{10}(m)\right], \tag{36}$$

where

$$\widehat{\underline{\boldsymbol{h}}}_j^{10}(m) = \boldsymbol{\mathcal{W}}\widehat{\underline{\boldsymbol{h}}}_j(m). \tag{37}$$

Unlike the previous NMCFLMS-type algorithms which use the gradient of the cost function with respect to $\widehat{\underline{\boldsymbol{h}}}_k^*(m)$, we calculate the gradient of the cost function with respect to $\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*$ to derive the adaptive filtering algorithm. The reason is that by doing so, we can derive the multichannel frequency-domain adaptive filter to directly obtain a generalized diagonal power spectrum matrix, which largely reduces the computational complexity of the inversion of this matrix. According to (27), the gradient of $\widetilde{\mathcal{J}}(m)$ with respect to $\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*$ is written as

$$\begin{aligned}
\nabla\widetilde{\mathcal{J}}(m) &= 2\frac{\partial \mathcal{J}_{\mathrm{F}}(m)}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*} - 2\eta(m)\frac{\partial \mathcal{J}_{\mathrm{FP}}(m)}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*} \\
&= 2\sum_{i=1}^{k-1}\left(\boldsymbol{\mathcal{G}}\boldsymbol{\mathcal{D}}_i(m)\right)^H \underline{\boldsymbol{e}}_{ik}(m) \\
&\quad - 2\sum_{j=k+1}^{M}\left(\boldsymbol{\mathcal{G}}\boldsymbol{\mathcal{D}}_j(m)\right)^H \underline{\boldsymbol{e}}_{kj}(m) \\
&\quad - \eta(m)p\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath\arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\} \\
&= \sum_{i=1}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\underline{\boldsymbol{e}}_{ik}(m) \\
&\quad - \eta(m)p\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath\arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\},
\end{aligned} \tag{38}$$

where $|\cdot|^{p-1}$ is carried out in a component-wise way, and

$$\boldsymbol{\mathcal{R}} = 2\boldsymbol{\mathcal{G}}^H. \tag{39}$$

So, we can get the robust multichannel frequency-domain least mean square (RMCFLMS) algorithm for BMCI as follows:

$$\begin{aligned}
\widehat{\underline{\boldsymbol{h}}}_k^{10}(m+1) &= \widehat{\underline{\boldsymbol{h}}}_k^{10}(m) - \mu_{\mathrm{f}}\sum_{i=1}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\underline{\boldsymbol{e}}_{ik}(m) \\
&\quad + \mu_{\mathrm{f}}\eta(m)p\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath\arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\}, \\
&\quad k = 1, 2, \ldots, M.
\end{aligned} \tag{40}$$

As can be seen from (40) that if we omit the third additive term on the right-hand side of this expression, the update equation degenerates to that of the multichannel frequency-domain least mean square (MCFLMS) algorithm [17]. This indicates that the MCFLMS algorithm can be directly deduced by calculating the gradient of the cost function in the MCFLMS algorithm with respect to $\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*$.

To accelerate the convergence of the adaptive filter and diminish the gradient noise amplification due to large channel outputs [17], we herein employ the iteration scheme of Newton's method [17], [18] to update the adaptive filter coefficient vector. According to Newton's method, we can write the update equations of the channel estimates as

$$\begin{aligned}
\widehat{\underline{\boldsymbol{h}}}_k^{10}(m+1) &= \widehat{\underline{\boldsymbol{h}}}_k^{10}(m) - \mu_{\mathrm{f}}\boldsymbol{\mathcal{S}}_k^{-1}(m)\nabla\widetilde{\mathcal{J}}(m) \\
&\quad k = 1, 2, \ldots, M,
\end{aligned} \tag{41}$$

where the Hessian matrix $\boldsymbol{\mathcal{S}}_k(m)$ can be derived by

$$\begin{aligned}
\boldsymbol{\mathcal{S}}_k(m) &= 2\frac{\partial}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*}[\nabla\widetilde{\mathcal{J}}(m)]^H \\
&= 2\frac{\partial\left[\sum_{i=1,i\neq k}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\underline{\boldsymbol{e}}_{ik}(m)\right]^H}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*} - 2\eta(m)p \\
&\quad \times \frac{\partial\left\{\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath\arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\}\right\}^H}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*}.
\end{aligned} \tag{42}$$

The first partial derivative term in (42) is easily deduced as

$$\begin{aligned}
&\frac{\partial\left[\sum_{i=1,i\neq k}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\underline{\boldsymbol{e}}_{ik}(m)\right]^H}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*} \\
&= \sum_{i=1,i\neq k}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{G}}^H\boldsymbol{\mathcal{R}}^H\boldsymbol{\mathcal{D}}_i(m) \\
&= \sum_{i=1,i\neq k}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{G}}\boldsymbol{\mathcal{D}}_i(m) \\
&= \frac{1}{2}\boldsymbol{\mathcal{P}}_k(m),
\end{aligned} \tag{43}$$

where $\boldsymbol{\mathcal{P}}_k(m)$ is defined in (19), and an approximate relationship:

$$\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{G}} \approx \frac{1}{2}\mathbf{I}_{2L} \tag{44}$$

is used if L is large. For the second partial derivative term in (42), we first write

$$\begin{aligned}
&\left\{\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath \arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\}\right\}^H \\
&= \begin{bmatrix} \left|\widehat{\underline{h}}_{k,0}^{10}(m)\right|^{p-1} \exp\left\{-\jmath \arg\left[\widehat{\underline{h}}_{k,0}^{10}(m)\right]\right\} \\ \left|\widehat{\underline{h}}_{k,1}^{10}(m)\right|^{p-1} \exp\left\{-\jmath \arg\left[\widehat{\underline{h}}_{k,1}^{10}(m)\right]\right\} \\ \vdots \\ \left|\widehat{\underline{h}}_{k,2L-1}^{10}(m)\right|^{p-1} \exp\left\{-\jmath \arg\left[\widehat{\underline{h}}_{k,2L-1}^{10}(m)\right]\right\} \end{bmatrix}^T \\
&= \begin{bmatrix} \frac{\left|\widehat{\underline{h}}_{k,0}^{10}(m)\right|^p}{\widehat{\underline{h}}_{k,0}^{10}(m)} \\ \frac{\left|\widehat{\underline{h}}_{k,1}^{10}(m)\right|^p}{\widehat{\underline{h}}_{k,1}^{10}(m)} \\ \vdots \\ \frac{\left|\widehat{\underline{h}}_{k,2L-1}^{10}(m)\right|^p}{\widehat{\underline{h}}_{k,2L-1}^{10}(m)} \end{bmatrix}^T .
\end{aligned} \tag{45}$$

Since

$$\begin{aligned}
\frac{\partial\left[\frac{\left|\widehat{\underline{h}}_{k,l}^{10}(m)\right|^p}{\widehat{\underline{h}}_{k,l}^{10}(m)}\right]}{\partial\left(\widehat{\underline{h}}_{k,l}^{10}(m)\right)^*} &= \widehat{\underline{h}}_{k,l}^{10}(m)\frac{p\left|\widehat{\underline{h}}_{k,l}^{10}(m)\right|^{p-1}}{\left(\widehat{\underline{h}}_{k,l}^{10}(m)\right)^2}\frac{\widehat{\underline{h}}_{k,l}^{10}(m)}{2\left|\widehat{\underline{h}}_{k,l}^{10}(m)\right|} \\
&= \frac{p\left|\widehat{\underline{h}}_{k,l}^{10}(m)\right|^{p-2}}{2}, l = 0,1,\ldots,2L-1,
\end{aligned} \tag{46}$$

then

$$\begin{aligned}
&\frac{\partial}{\partial\left(\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right)^*}\left\{\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath \arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\}\right\}^H \\
&= \operatorname{diag}\left\{\begin{bmatrix} \frac{p\left|\widehat{\underline{h}}_{k,0}^{10}(m)\right|^{p-2}}{2} \\ \frac{p\left|\widehat{\underline{h}}_{k,1}^{10}(m)\right|^{p-2}}{2} \\ \vdots \\ \frac{p\left|\widehat{\underline{h}}_{k,2L-1}^{10}(m)\right|^{p-2}}{2} \end{bmatrix}\right\}.
\end{aligned} \tag{47}$$

So, the Hessian matrix $\boldsymbol{\mathcal{S}}_k(m)$ is obtained as follows:

$$\boldsymbol{\mathcal{S}}_k(m) = \boldsymbol{\mathcal{P}}_k(m) - \eta(m)p^2\boldsymbol{\mathcal{H}}_k(m), \tag{48}$$

where

$$\boldsymbol{\mathcal{H}}_k(m) = \operatorname{diag}\left\{\begin{bmatrix} \left|\widehat{\underline{h}}_{k,0}^{10}(m)\right|^{p-2} \\ \left|\widehat{\underline{h}}_{k,1}^{10}(m)\right|^{p-2} \\ \vdots \\ \left|\widehat{\underline{h}}_{k,2L-1}^{10}(m)\right|^{p-2} \end{bmatrix}\right\}. \tag{49}$$

As can be seen, $\boldsymbol{\mathcal{S}}_k(m)$ in (48) is a diagonal matrix, so its inversion can be computed efficiently. As compared to the power spectrum matrix $\boldsymbol{\mathcal{P}}_k(m)$ of the NMCFLMS-type algorithms, the matrix $\boldsymbol{\mathcal{S}}_k(m)$ is adjusted by a weighted spectrum of the acoustic channel impulse responses. Thus, we can define this Hessian matrix $\boldsymbol{\mathcal{S}}_k(m)$ as a generalized power spectrum matrix.

Now, substituting (38) and (48) into (41) produces the update equations of the proposed ℓ_p-RNMCFLMS algorithm as follows:

$$\begin{aligned}
\widehat{\underline{\boldsymbol{h}}}_k^{10}(m+1) = \widehat{\underline{\boldsymbol{h}}}_k^{10}(m) &- \mu_{\mathrm{f}}\left[\boldsymbol{\mathcal{P}}_k(m) - \eta(m)p^2\boldsymbol{\mathcal{H}}_k(m)\right]^{-1} \\
&\times \left\{\sum_{i=1}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\underline{\boldsymbol{e}}_{ik}(m) - \eta(m)\right. \\
&\left.\times p\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \odot \exp\left\{\jmath \arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\}\right\}, \\
&k = 1,2,\ldots,M.
\end{aligned} \tag{50}$$

Comparing the update equations of RNMCFLMS [18], [19] with those of the proposed ℓ_p-RNMCFLMS algorithm, one can see that the RNMCFLMS algorithm is an approximate solution without detailed derivation while the ℓ_p-RNMCFLMS algorithm is rigorously derived upon the basis of the robust cost function and Newton's method. Also, (50) can be formulated as a simplified version as follows:

$$\begin{aligned}
\widehat{\underline{\boldsymbol{h}}}_k^{10}(m+1) = \widehat{\underline{\boldsymbol{h}}}_k^{10}(m) - \mu_{\mathrm{f}}\nabla\mathcal{J}_{\mathrm{F},k}(m) + \mu_{\mathrm{f}}\eta(m)\nabla\mathcal{J}_{\mathrm{FP},k}(m), \\
k = 1,2,\ldots,M,
\end{aligned} \tag{51}$$

where

$$\nabla\mathcal{J}_{\mathrm{F},k}(m) = \boldsymbol{\mathcal{S}}_k^{-1}(m)\sum_{i=1}^{M}\boldsymbol{\mathcal{D}}_i^*(m)\boldsymbol{\mathcal{R}}\underline{\boldsymbol{e}}_{ik}(m), \tag{52}$$

$$\begin{aligned}
\nabla\mathcal{J}_{\mathrm{FP},k}(m) = p\boldsymbol{\mathcal{S}}_k^{-1}(m)&\left|\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right|^{p-1} \\
&\odot \exp\left\{\jmath\arg\left[\widehat{\underline{\boldsymbol{h}}}_k^{10}(m)\right]\right\},
\end{aligned} \tag{53}$$

$$\eta(m) = \left|\frac{\left[\nabla\mathcal{J}_{\mathrm{FP}}(m)\right]^H\nabla\mathcal{J}_{\mathrm{F}}(m)}{\left\|\nabla\mathcal{J}_{\mathrm{FP}}(m)\right\|_{\ell_2}^2}\right|, \tag{54}$$

$$\begin{aligned}
\nabla\mathcal{J}_{\mathrm{FP}}(m) = \Big[&\left(\nabla\mathcal{J}_{\mathrm{FP},1}(m)\right)^T \left(\nabla\mathcal{J}_{\mathrm{FP},2}(m)\right)^T \\
&\cdots \left(\nabla\mathcal{J}_{\mathrm{FP},M}(m)\right)^T\Big]^T,
\end{aligned} \tag{55}$$

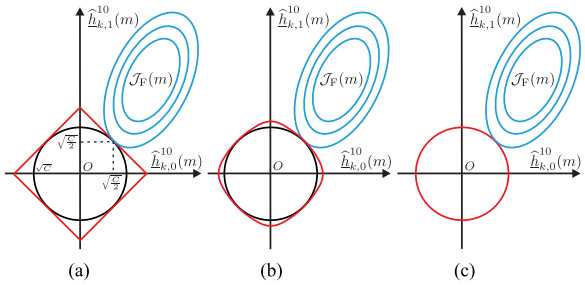

Fig. 2. Effect of different norm penalties on the optimization of the cost function $\widetilde{\mathcal{J}}(m)$, where the concentric ellipses denote the contours of the error curved surface corresponding to $\mathcal{J}_{\mathrm{F}}(m)$, the black circle with the radius of $\sqrt{C}$ stands for the ℓ_2-norm constraint of the amplitude spectrum of the adaptive filter coefficient vector in (59), the diamond in (a) depicts the ℓ_1-norm penalty, the rounded and curved diamond in (b) denotes the ℓ_p-norm ($1 < p < 2$) penalty, and the circle in (c) stands for the ℓ_2-norm penalty (overlapped with the black circle that depicts unit length constraint).

$$\nabla\mathcal{J}_{\mathrm{F}}(m) = \Big[\left(\nabla\mathcal{J}_{\mathrm{F},1}(m)\right)^T \ \left(\nabla\mathcal{J}_{\mathrm{F},2}(m)\right)^T \cdots \left(\nabla\mathcal{J}_{\mathrm{F},M}(m)\right)^T \Big]^T. \tag{56}$$

The Lagrange multiplier $\eta(m)$ in (54) is calculated in such a way that the total update vector is not significant to cause divergence. Note that this multiplier in $\boldsymbol{\mathcal{S}}_k(m)$ is adjusted as $\eta(m-1)$ to calculate $\eta(m)$ in (54).

Finally, the estimated time-domain adaptive filter coefficient vector $\widehat{\mathbf{h}}(m+1)$ is imposed to have a unit length for each update, i.e.,

$$\widehat{\mathbf{h}}(m+1) := \frac{\widehat{\mathbf{h}}(m+1)}{\left\|\widehat{\mathbf{h}}(m+1)\right\|_{\ell_2}}. \tag{57}$$

B. Effect of Different Norm Penalties on the Optimization of the Frequency-Domain Adaptive Filter

The optimization on the ℓ_1-norm as a penalty of the cost function can be intuitively interpreted according to Fig. 2 where we only consider the case of two dimensions. As can be seen from Fig. 2(a), under the constraints of both the ℓ_2-norm and the ℓ_1-norm of the amplitude spectrum of the adaptive filter coefficient vector, it is easy for the frequency-domain adaptive filter to search for a dense solution in which all of elements of the frequency-domain adaptive filter coefficient vector are equal or close, especially for a perfectly homogeneous amplitude spectrum, i.e., anechoic channels.

To analyze the effect of the ℓ_p-norm ($1 < p < 2$) penalty on the frequency-domain adaptive filter, we first introduce a proposition. For convenience, we define a vector:

$$\begin{aligned}\boldsymbol{\gamma}(m) &= [\gamma_1(m)\ \gamma_2(m)\ \cdots\ \gamma_{2ML}(m)]^T \\ &= \Big[\left|\widehat{\underline{h}}^{10}_{1,0}(m)\right| \ \left|\widehat{\underline{h}}^{10}_{1,1}(m)\right| \cdots \left|\widehat{\underline{h}}^{10}_{1,2L-1}(m)\right| \\ &\quad \left|\widehat{\underline{h}}^{10}_{2,0}(m)\right| \ \left|\widehat{\underline{h}}^{10}_{2,1}(m)\right| \cdots \left|\widehat{\underline{h}}^{10}_{2,2L-1}(m)\right| \cdots \\ &\quad \left|\widehat{\underline{h}}^{10}_{M,0}(m)\right| \ \left|\widehat{\underline{h}}^{10}_{M,1}(m)\right| \cdots \left|\widehat{\underline{h}}^{10}_{M,2L-1}(m)\right| \Big]^T,\end{aligned} \tag{58}$$

which indicates $\gamma_i(m) \geq 0$, $i = 1, 2, \ldots, 2ML$. Then, the proposition is given as follows:

Proposition 1: According to the assumption that the impulse response vectors of the M acoustic channels are of a unit length, the square length of their amplitude spectra coefficients can be added up to a constant, i.e.,

$$\gamma_1^2(m) + \gamma_2^2(m) + \cdots + \gamma_{2ML}^2(m) = C, \tag{59}$$

where C is a constant. Then,

$$\mathcal{J}_{\mathrm{FP}}(m) = \gamma_1^p(m) + \gamma_2^p(m) + \cdots + \gamma_{2ML}^p(m) \tag{60}$$

takes the maximum value of $2ML\left(\frac{C}{2ML}\right)^{\frac{p}{2}}$ when $\gamma_i(m) = \sqrt{\frac{C}{2ML}}$, $i = 1, 2, \ldots, 2ML$.

Proof: This proposition can be formulated as a constrained optimization problem as follows:

$$\max\ \mathcal{J}_{\mathrm{FP}}(m) = \gamma_1^p(m) + \gamma_2^p(m) + \cdots + \gamma_{2ML}^p(m) \tag{61}$$

$$\text{s.t. } \gamma_1^2(m) + \gamma_2^2(m) + \cdots + \gamma_{2ML}^2(m) = C. \tag{62}$$

Using the method of Lagrangian multipliers [23] to solve the constrained optimization problem, we rewrite the Lagrangian function of (61) and (62) as

$$\begin{aligned}\mathcal{L}[\gamma_i(m), \lambda] &= \gamma_1^p(m) + \gamma_2^p(m) + \cdots + \gamma_{2ML}^p(m) \\ &\quad + \lambda\left[\gamma_1^2(m) + \gamma_2^2(m) + \cdots + \gamma_{2ML}^2(m) - C\right],\end{aligned} \tag{63}$$

where λ is the Lagrangian multiplier. Let the partial derivative of $\mathcal{L}[\gamma_i(m), \lambda]$ with respect to $\gamma_i(m)$, $i = 1, 2, \ldots, 2ML$, be equal to zero, i.e.,

$$\begin{aligned}\frac{\partial\mathcal{L}[\gamma_i(m), \lambda]}{\partial\gamma_i(m)} &= p\gamma_i^{p-1}(m) + 2\lambda\gamma_i(m) \\ &= \gamma_i^{p-1}(m)\left[p + 2\lambda\gamma_i^{2-p}(m)\right] \\ &= 0.\end{aligned} \tag{64}$$

Since the solution is assumed to be nontrivial, it follows immediately that

$$\gamma_i(m) = \left(\frac{-p}{2\lambda}\right)^{\frac{1}{2-p}},\ i = 1, 2, \ldots, 2ML. \tag{65}$$

Substituting (65) into (62) and with some simple mathematical manipulations, we obtain

$$\lambda = -\frac{p}{2}\left(\frac{C}{2ML}\right)^{-\frac{2-p}{2}}. \tag{66}$$

Substituting (66) into (65), we get

$$\gamma_i(m) = \sqrt{\frac{C}{2ML}},\ i = 1, 2, \ldots, 2ML. \tag{67}$$

Consequently, $\mathcal{J}_{\mathrm{FP}}(m)$ in (61) obtains the maximum value:

$$\begin{aligned}\mathcal{J}_{\mathrm{FP}}(m) &= 2ML\gamma_i^p(m) \\ &= C^{\frac{p}{2}}(2ML)^{1-\frac{p}{2}} \\ &= 2ML\left(\frac{C}{2ML}\right)^{\frac{p}{2}}.\end{aligned} \tag{68}$$

Note that the impulse responses of all the M channels are the unit impulse responses when (67) holds.

Furthermore, according to the property of the Fourier matrix $\boldsymbol{\mathcal{F}}_L^H \boldsymbol{\mathcal{F}}_L = L\mathbf{I}_L$ and the constraint of the unit length of the channel time-domain impulse responses, we deduce that

$$\begin{aligned}\sum_{k=1}^{2ML} \gamma_k^2(m) &= \sum_{k=1}^{M} \left\| \widehat{\underline{\boldsymbol{h}}}_k^{10}(m) \right\|_{\ell_2}^2 \\ &= \sum_{k=1}^{M} \left(\widehat{\mathbf{h}}_k^{10}(m) \right)^T \boldsymbol{\mathcal{F}}_{2L}^H \boldsymbol{\mathcal{F}}_{2L} \widehat{\mathbf{h}}_k^{10}(m) \\ &= 2L \sum_{k=1}^{M} \widehat{\mathbf{h}}_k^T(m) \widehat{\mathbf{h}}_k(m) \\ &= 2L, \end{aligned} \tag{69}$$

which indicates that the constant $C = 2L$. This completes the proof. ■

The effect of the ℓ_p-norm ($1 < p < 2$) penalty on the optimization of the cost function $\widetilde{\mathcal{J}}(m)$ can be illustrated by Fig. 2(b). As can be seen, under the condition of $\gamma_i(m) = \sqrt{\frac{C}{2ML}}$, $i = 1, 2, \ldots, 2ML$, the frequency-domain adaptive filter indeed can obtain a dense solution. For the acoustic channels with moderate dense spectra, such as moderate reverberation channels, however, the ℓ_1-norm and ℓ_p-norm penalties cannot yield an optimal solution according to Fig. 2(a) and (b). In this case, the ℓ_p-norm penalty is easier to lead to an approximate solution than the ℓ_1-norm penalty. It can be further seen from (68) that as the value of the parameter $p\,(1 \le p < 2)$ is increased, the maximum value of the penalty function $\mathcal{J}_{\mathrm{FP}}(m)$ decreases [note that the constant $C = 2L$ according to (69)], which makes the minimum value of the total cost function $\widetilde{\mathcal{J}}(m)$ larger. So, the larger is the value of the parameter p, the worse is the convergence performance for this type of sparse acoustic channels. For realistic reverberant acoustic channels, however, the amplitude spectra of the impulse responses are not flat. The penalty function $\mathcal{J}_{\mathrm{FP}}(m)$ does not, in general, converge to the maximum value in (68) but to a suboptimal value during the course of optimization. Thus, the corresponding convergence rate slows down. Note that if $p = 2$, the developed frequency-domain adaptive filter can no longer identify the dense acoustic channels, which can be seen from Fig. 2(c).

IV. Simulations

This section investigates and compares the performances of the NMCFLMS, RNMCFLMS, and developed ℓ_p-RNMCFLMS algorithms in a simulated room with the dimension of 7.0 m × 6.0 m × 3.0 m. For convenience, positions in the room are designated by (x, y, z) coordinates with reference to the northwest corner of the floor. An equispaced linear array which consists of three omnidirectional microphones is employed in the measurement and the spacing between adjacent microphones is 0.5 m. The three microphones of the array are placed at (3.0, 0.5, 1.4), (3.5, 0.5, 1.4), and (4.0, 0.5, 1.4), respectively. A loudspeaker, which serves as a sound source, is placed at (0.5, 4.0, 1.6).

The impulse responses from the source to the three microphones are generated using the image model [24], where the sampling rate is set to 16 kHz. Then the obtained channel impulse responses are truncated to 1024 samples, which will be viewed as the true impulse responses in the blind multichannel identification. The microphones' outputs are obtained by convolving the source signal with the corresponding generated impulse responses and then adding zero-mean white Gaussian noise to the results to control the SNR. The sound source signal is a white Gaussian sequence, which is approximately 4.7-minute long.

Four reverberation conditions with different values of reverberation time T_{60} (defined as the time for the sound to die away to a level 60 dB below its original level) are considered to test the performance of the three frequency-domain adaptive filters, i.e., anechoic ($T_{60} = 0$ ms), lightly reverberant ($T_{60} = 50$ ms), moderately reverberant ($T_{60} = 300$ ms), and random ($T_{60} \to \infty$) channels. The amplitude spectra of the first-channel impulse response for the four acoustic environments are plotted in Fig. 3(a)–(d) for illustration. For the three algorithms, the step size μ_{F} is set to 0.1 for the anechoic and light reverberation environments, and 0.5 for the moderate reverberation condition and the case of $T_{60} \to \infty$. We use the normalized projection misalignment (NPM) [25] as the metric to evaluate the BMCI performance. The NPM at block m is defined as

$$\mathrm{NPM}(m) = 20 \log_{10} \left[\frac{\|\boldsymbol{\epsilon}(m)\|_{\ell_2}}{\|\mathbf{h}\|_{\ell_2}} \right] \mathrm{dB}, \tag{70}$$

where

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_1^T & \mathbf{h}_2^T & \cdots & \mathbf{h}_M^T \end{bmatrix}^T \tag{71}$$

consists of the true impulse responses,

$$\boldsymbol{\epsilon}(m) = \mathbf{h} - \frac{\mathbf{h}^T \widehat{\mathbf{h}}(m)}{\widehat{\mathbf{h}}^T(m) \widehat{\mathbf{h}}(m)} \widehat{\mathbf{h}}(m) \tag{72}$$

is the projection misalignment vector. All the results are averaged over 100 Monte Carlo runs.

Fig. 4 depicts the convergence behavior of the three studied algorithms for the identification of acoustic channels with $L = 1024$ in the four reverberation conditions at the SNR of 5 dB, where the acoustic system is excited by a white Gaussian sequence. It can be seen from Fig. 4 that the NMCFLMS algorithm diverges in all the four acoustic environments due to the presence of noise, which indicates that this algorithm is sensitive to noise. In the anechoic environment, one can see from Fig. 4(a) that the RNMCFLMS and ℓ_p-RNMCFLMS algorithms are convergent. The smaller is the value of the parameter p, the smaller is the NPM of the ℓ_p-RNMCFLMS algorithm, which demonstrates the analysis of Proposition 1. Meanwhile, the NPM for the RNMCFLMS algorithm is the least in the anechoic acoustic environment. The underlying reason for this is that the acoustic impulse responses in this environment are sparsest and the spectra of the channels are flat. In this scenario, the logarithmic penalty function is closer to the ℓ_0-norm, which is better to

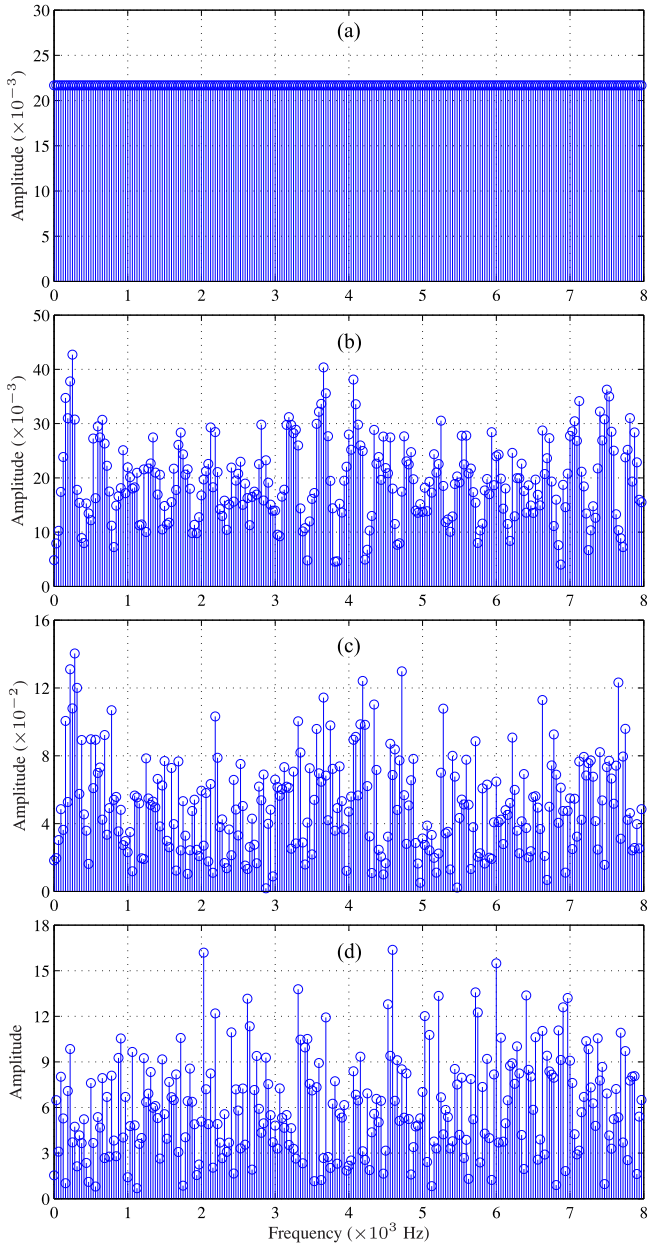

Fig. 3. Amplitude spectra of the impulse responses for the acoustic channel with different reverberation times. (a) Anechoic channel ($T_{60} = 0$ ms). (b) Light reverberation channel ($T_{60} = 50$ ms). (c) Moderate reverberation channel ($T_{60} = 300$ ms). (d) Random acoustic channel ($T_{60} \to \infty$).

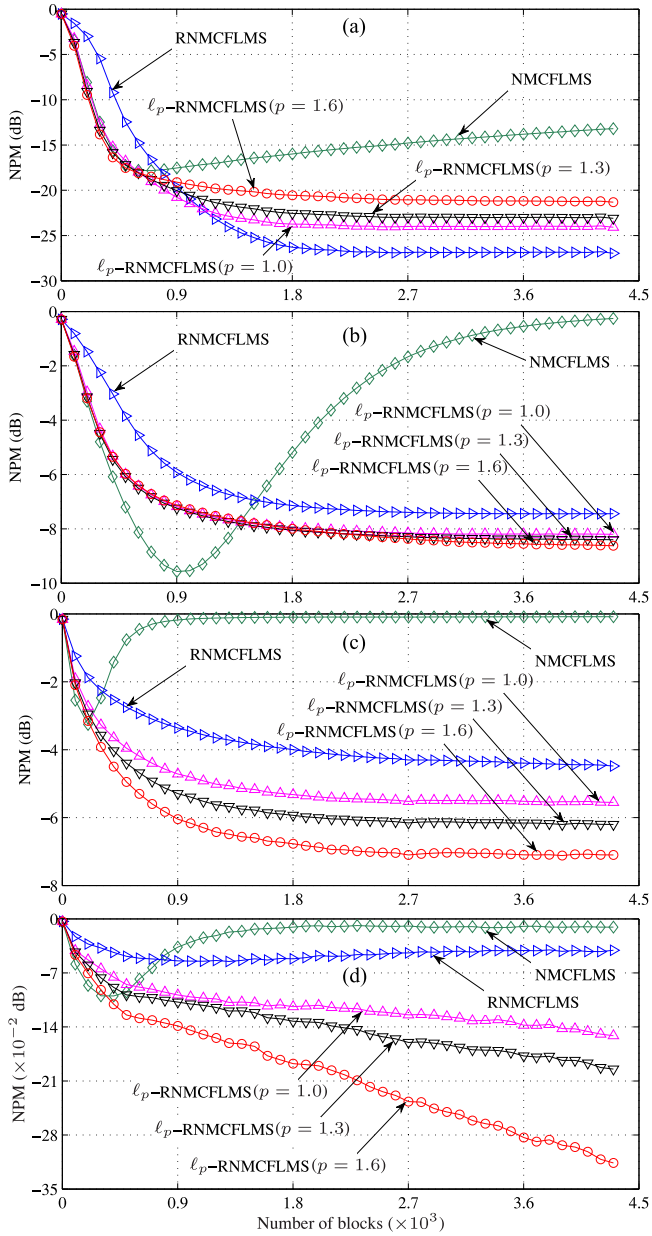

Fig. 4. Comparison of convergence among the NMCFLMS, RNMCFLMS (with logarithmic penalty), and ℓ_p-RNMCFLMS algorithms for the identification of a three-channel system excited by a white Gaussian sequence at the SNR of 5 dB. (a) Anechoic channel ($T_{60} = 0$ ms). (b) Light reverberation channel ($T_{60} = 50$ ms). (c) Moderate reverberation channel ($T_{60} = 300$ ms). (d) Random acoustic channel ($T_{60} \to \infty$).

measure the sparsity of this type of channel impulse responses. Therefore, the frequency-domain adaptive filter with the logarithmic penalty is better than that with the ℓ_p-norm ($1 \le p < 2$) penalty to blindly identify the impulse responses.

When there exists reverberation, the acoustic channel impulse responses become less sparse, which makes the logarithmic penalty function no longer optimal. It can be seen from Fig. 4(b)–(d) that the ℓ_p-RNMCFLMS algorithm obtains better performance than the RNMCFLMS algorithm, indicating that the penalty with ℓ_p-norm ($p = 1.0$, 1.3, 1.6) is better than the logarithmic penalty under reverberation environments. We can see from Fig. 4(b) that in a light reverberation environment, the adaptive filter with the three different ℓ_p-norm penalties ($p = 1.0$, 1.3, 1.6) has similar convergence behavior. As the reverberation of this room is increased, the penalty with greater parameter p results in a better convergence and a smaller NPM, as shown in Fig. 4(c) and (d). This demonstrates that the penalty with a greater value of p is favorable for the identification of the acoustic channels with heavier reverberation. We can also see from Fig. 4 that stronger reverberation makes the identification of acoustic channels more difficult, especially for the random channels with $T_{60} \to \infty$. Note that for the proposed ℓ_p-RNMCFLMS algorithm, the optimal value of p varies with

reverberation and how to find the optimal value given an acoustic environment is a problem worthy of further investigation.

In this section, the spacing between microphones is 50 cm. Note that the spacing between microphones may significantly affect the performance of BMCI. Generally, a smaller spacing between microphones increases the difficulty of blindly identifying the system since there is less diversity between different channels and common zeros are more likely to happen.

V. Experiments

In this section, we validate the effectiveness of the ℓ_p-RNMCFLMS algorithm in a real acoustic environment where the channel impulse responses were measured at the Bell Labs varechoic chamber [26]. The dimension of the chamber is 6.7 m $\times$ 6.1 m $\times$ 2.9 m. An equispaced linear array which is composed of three omnidirectional microphones is used in the measurement, which are placed at (2.437, 0.500, 1.400), (3.137, 0.500, 1.400), and (3.837, 0.500, 1.400), respectively. A loudspeaker is placed at (0.337, 3.938, 1.600), which plays back a sound signal to simulate a source. The transfer functions of the acoustic channels between the source and microphones were measured at a 48 kHz sampling rate when 89% panels on the walls were open and the corresponding reverberation time of the chamber is approximately 280 ms [26]. Then the obtained channel impulse responses are downsampled to a 16 kHz sampling rate and truncated to 1024 samples. The measured impulse responses will be treated as the true impulse responses in BMCI.

We consider two cases: 1) the sound source signal is a white Gaussian sequence and 2) the source signal is a recorded speech signal sampled at 16 kHz. The length of the two source signals is also approximately 4.7-minute long. The multichannel system outputs are generated by convolving the sound source signal with the corresponding measured channel impulse responses and zero-mean white Gaussian noise is then added to the results at a given level of SNR. The step size μ_{f} is set to 0.5. All the results are averaged over 100 estimates.

Fig. 5 plots the average NPM of the three algorithms versus SNR for the acoustic channel identification of the three-channel system with $L = 1024$, where the acoustic system is excited by a white Gaussian sequence. It can be seen that the NMCFLMS algorithm is most susceptible to noise. The introduction of the logarithmic penalty into NMCFLMS improves its robustness to noise due to the constraint on the spectral flatness. In comparison, the ℓ_p-RNMCFLMS algorithm with the p-norm ($1 \leq p < 2$) penalty obtains the best performance in most of the studied conditions when SNR is low. Again, this shows that the ℓ_p-norm ($1 \leq p < 2$) penalty is more appropriate than the logarithmic penalty to help robustness improvement of BMCI to noise. When SNR is high, however, the two algorithms with sparsity constraints do not produce better robustness than NMCFLMS. As SNR is increased, the performance of all the three algorithms tends to be similar. This is due to the following two aspects: 1) they all have the common part in the total cost function, i.e., $\mathcal{J}_{\mathrm{F}}(m)$, and 2) when SNR is high, the effect of the even spectrum of white noise on the modeling filter is so negligible

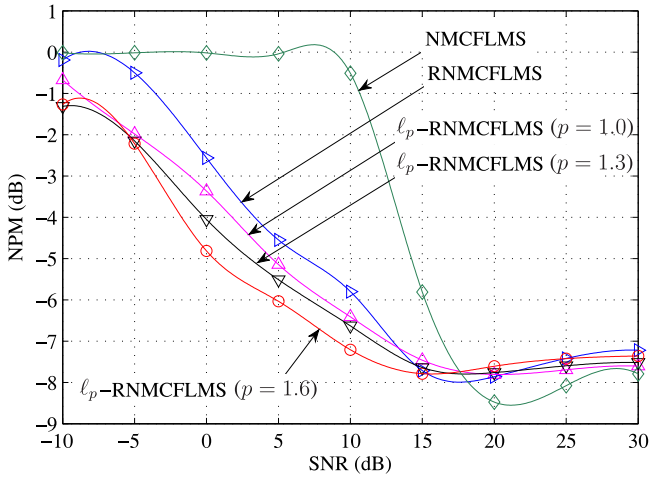

Fig. 5. Comparison of the NPM among the NMCFLMS, RNMCFLMS (with logarithmic penalty), and ℓ_p-RNMCFLMS algorithms versus SNR for the identification of a three-channel acoustic system excited by a white Gaussian sequence.

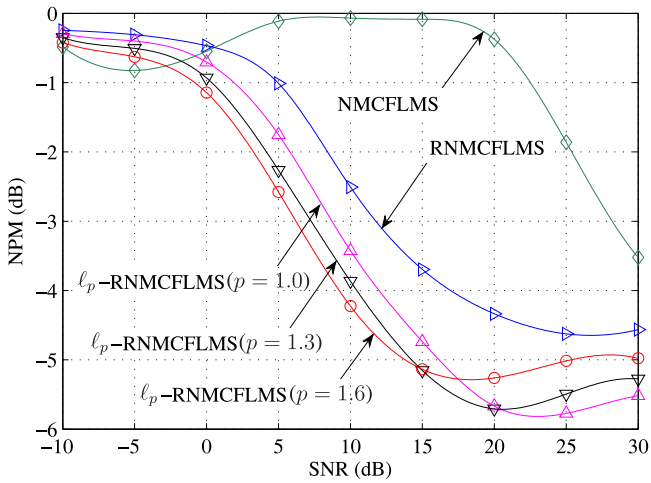

Fig. 6. Comparison of the NPM among the NMCFLMS, RNMCFLMS (with logarithmic penalty), and ℓ_p-RNMCFLMS algorithms versus SNR for the identification of a three-channel acoustic system excited by a speech signal.

that the *a priori* error $e_{ij}(n)$ mainly comes from a modeling error, which indicates that this filter is basically of sparsity during the course of adaptive optimization.

Fig. 6 illustrates the NPM of the three algorithms versus SNR for the acoustic channel identification with speech excitation. All the parameters are set to the same as in the previous experiment. Comparing the results to those of the previous experiment, one can see that the performance of all the frequency-domain adaptive filtering algorithms degradates when the excitation is speech. Again, the NMCFLMS algorithm is most sensitive to noise. The RNMCFLMS algorithm uses the logarithmic penalty to enhance its immunity to noise. In comparison with RNMCFLMS, the ℓ_p-RNMCFLMS algorithm exhibits better performance when SNR is low. One can notice from Fig. 6 that a large value of p can help improve the robustness of the

ℓ_p-RNMCFLMS algorithm with respect to noise, i.e., its performance is better for a large value of p when SNR is low. The underlying reason is as follows. As mentioned in [22], the ℓ_p-norm ($1 \leq p < 2$) is an approximation to the elastic net function, which is a convex combination of the lasso (corresponding to the ℓ_1-norm) and ridge (corresponding to the ℓ_2-norm) penalty. For a Laplace distributed acoustic channel, the ℓ_1-norm penalty is a good candidate, but for a Gaussian distributed acoustic channel, the ℓ_2-norm penalty provides a stable solution. When SNR is low (e.g., SNR $<$ 10 dB), the microphone signal is closer to Gaussian distribution than Laplace distribution. In this case, the ℓ_p-norm penalty with a relatively large value of p is preferable. In comparison, when SNR is high (e.g., SNR $\geq$ 20 dB), the probability distribution of the multichannel system output gets close to that of the speech source signal. A greater value of p generally makes the ℓ_p-RNMCFLMS algorithm more sensitive to adverse effect of speech nonstationarity and non-Gaussianity, which corroborates the observation from [27]. Therefore, considering performance consistency in both high and low SNR environments, we suggest to set the value of p not too large.

VI. Conclusion

In this paper, an ℓ_p-RNMCFLMS algorithm was developed to blindly identify the acoustic impulse responses of an acoustic SIMO system in noisy environments. This algorithm exploits the ℓ_p-norm with $1 \leq p < 2$ as a spectral penalty on the acoustic channel impulse responses to improve its robustness to noise. The property of this algorithm is analyzed both theoretically and experimentally and its performance depends on both the sparseness of the acoustic impulse responses as well as the nature of the source excitation. Basically, the algorithm emphasizes more on the sparseness of the acoustic impulse responses with a smaller value of p. If the excitation is white, a large value of p makes the algorithm more robust to noise. In practical acoustic environments with speech excitation, however, the performance varies with the value of p. Generally, a large value of p is preferred at low SNR conditions while a small value of p is preferred if SNR is high. In future work, we will investigate how different ℓ_p-norm penalties affect the performance of the adaptive filter for the identification of the acoustic channels in reverberant and non-Gaussian noise environments.

Acknowledgment

The authors would like to thank the associate editor and the three anonymous reviewers for their careful reading of the draft and providing many constructive comments and suggestions, which helped improve the clarity and quality of this paper.

References

[1] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.

[2] H. Luo and Y. Li, "The application of blind channel identification techniques to prestack seismic deconvolution," *Proc. IEEE*, vol. 86, no. 10, pp. 2082–2089, Oct. 1998.

[3] J. K. Tugnait, "A multidelay whitening approach to blind identification and equalization of SIMO channels," *IEEE Trans. Wireless Commun.* vol. 1, no. 3, pp. 456–467, Jul. 2002.

[4] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[5] C. G. Tsinos, A. S. Lalos, and K. Berberidis, "Sparse subspace tracking techniques for adaptive blind channel identification in OFDM systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2012, pp. 3185–3188.

[6] C. Yu, C. Zhang, and L. Xie, "A new deterministic identification approach to hammerstein systems," *IEEE Trans. Signal Process.*, vol. 62, pp. 131–140, Jan. 2014.

[7] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 340–349, Mar. 1994.

[8] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[9] C. Yu, L. Xie, and Y. Soh, "Blind channel and source estimation in networked systems," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4611–4626, Sep. 2014

[10] H-Z. Tan and T. Aboulnasr, "Tom-based blind identification of nonlinear volterra systems," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 1, pp. 300–310, Feb. 2006.

[11] D. Raphaeli, U. Suissa, and G. Kutz, "Blind channel identification from burst data using implicit matching of HOS," *Elsevier Signal Process.*, vol. 90, no. 3, pp. 795–808, Mar. 2010.

[12] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 661–672, Mar. 1996.

[13] F. Alberge, P. Duhamel, and M. Nikolova, "Adaptive solution for blind identification/equalization using deterministic maximum likelihood," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 923–936, Apr. 2002.

[14] C. Yu, C. Zhang, and L. Xie, "Blind system identification using precise and quantized observations," *Automatica*, vol. 49, no. 9, pp. 2822–2830, Sep. 2013.

[15] W. Qiu, S. K. Saleem, and M. Pham, "Blind identification of multichannel systems driven by impulsive signals," *Dig. Signal Process.*, vol. 20, pp. 736–742, May 2010.

[16] H. He, J. Lu, J. Chen, X. Qiu, and J. Benesty, "Robust blind identification of room acoustic channels in symmetric alpha-stable distributed noise environments," *J. Acoust. Soc. Amer.*, vol. 136, no. 8, pp. 693–704, Aug. 2014.

[17] Y. Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 11–24, Jan. 2003.

[18] M. A. Haque and M. K. Hasan, "Noise robust multichannel frequency-domain LMS algorithms for blind channel identification," *IEEE Signal Process. Lett.*, vol. 15, pp. 305–308, 2008.

[19] M. A. Haque, T. Islam, and Md. K. Hasan, "Robust speech dereverberation based on blind adaptive estimation of acoustic channels," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 4, pp. 775–787, May 2011.

[20] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.

[21] W. J. Fu, "Penalized regressions: The bridge versus the lasso," *J. Comput. Graph. Statist.*, vol. 7, no. 3, pp. 397–416, 1998.

[22] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. Roy. Statist. Soc. B*, vol. 67, pp. 301–320, 2005.

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[24] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, Apr. 1979.

[25] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Process. Lett.*, vol. 5, pp. 174–176, Jul. 1998.

[26] A. Härmä, "Acoustic measurement data from the varechoic chamber," Agere Systems, Allentown, PA, USA, Tech. Rep. 110101, Nov. 2001.

[27] S. Gazor and W. Zhang, "Speech probability distribution," *IEEE Signal Process. Lett.*, vol. 10, no. 7, pp. 204–207, Jul. 2003.

Hongsen He (M'14) was born in Sichuan, China. He received the B.E. degree in automation from Southwest University of Science and Technology (SWUST), Mianyang, China, in June 2000, and the Ph.D. degree in signal and information processing from Nanjing University (NJU), Nanjing, China, in June 2013.

He joined the School of Information Engineering, SWUST, as a member of faculty in July 2000. He is currently an Associate Professor, SWUST. From September 2007 to May 2013, he was a Research Assistant with the Institute of Acoustics, NJU, and engaged in research on adaptive signal processing and microphone array signal processing. His current research interests include signal processing and acoustic signal processing, including adaptive filtering, robust signal processing, sparse signal processing, multimicrophone signal processing, acoustic source localization, system identification, dereverberation, adaptive noise cancellation, and machine learning.

Dr. He is a Member for the IEEE and Acoustical Society of America, and is a Reviewer of several international journals and conferences. He was selected as one of the Distinguished Youth Academic and Technical Pacemakers of Sichuan Province in 2014 (Incubation).

Jingdong Chen (M'99–SM'09) received the Ph.D. degree in pattern recognition and intelligence control from the Chinese Academy of Sciences, Beijing, China, in 1998.

From 1998 to 1999, he was with ATR Interpreting Telecommunications Research Laboratories, Kyoto, Japan, where he conducted research on speech synthesis, speech analysis, as well as objective measurements for evaluating speech synthesis. Then, he joined the Griffith University, Brisbane, QLD, Australia, where he engaged in research on robust speech recognition and signal processing. From 2000 to 2001, he worked at ATR Spoken Language Translation Research Laboratories on robust speech recognition and speech enhancement. From 2001 to 2009, he was a Member of Technical Staff with Bell Laboratories, Murray Hill, NJ, USA, working on acoustic signal processing for telecommunications. He subsequently joined WeVoice Inc., Bridgewater, NJ, USA, serving as the Chief Scientist. He is currently a Professor with the Northwestern Polytechnical University, Xi'an, China. He co-authored the books *Fundamentals of Signal Enhancement and Array Signal Processing*, (Wiley, 2017), *Design of Circular Differential Microphone Arrays* (Springer, 2015), *Study and Design of Differential Microphone Arrays* (Springer, 2013), *Speech Enhancement in the STFT Domain* (Springer, 2011), *Optimal Time-Domain Noise Reduction Filters: A Theoretical Study* (Springer, 2011), *Speech Enhancement in the Karhunen-Loève Expansion Domain* (Morgan & Claypool, 2011), *Noise Reduction in Speech Processing* (Springer, 2009), *Microphone Array Signal Processing* (Springer, 2008), and *Acoustic MIMO Signal Processing* (Springer, 2006). His research interests include acoustic signal processing, adaptive signal processing, speech enhancement, adaptive noise/echo control, microphone array signal processing, signal separation, and speech communication.

Dr. Chen was an Associate Editor for the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING from 2008 to 2014. From 2007 to 2009, he was a Technical Committee (TC) member of the IEEE Signal Processing Society (SPS) TC on Audio and Electroacoustics. He is currently a Member for the IEEE SPS TC on Audio and Acoustic Signal Processing, and a Member of the editorial advisory board of the Open Signal Processing Journal. He was the General Cochair of IWAENC 2016, the Technical Program Chair of IEEE TENCON 2013, a Technical Program Cochair of IEEE WASPAA 2009, IEEE ChinaSIP 2014, IEEE ICSPCC 2014, and IEEE ICSPCC 2015, and helped organize many other conferences. He was the recipient of 2008 Best Paper Award from the IEEE Signal Processing Society (with Benesty, Huang, and Doclo), the Best Paper Award from the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics in 2011 (with Benesty), the Bell Labs Role Model Teamwork Award twice, respectively, in 2009 and 2007, the NASA Tech Brief Award twice, respectively, in 2010 and 2009, the Japan Trust International Research Grant from the Japan Key Technology Center in 1998, and the Young Author Best Paper Award from the 5th National Conference on Man-Machine Speech Communications in 1998. He is also a co-author of a paper for which C. Pan received the IEEE R10 (Asia-Pacific Region) Distinguished Student Paper Award (First Prize) in 2016. He was also a recipient of the Japan Trust International Research Grant from the Japan Key Technology Center in 1998 and the "Distinguished Young Scientists Fund" from the National Natural Science Foundation of China in 2014.

Jacob Benesty received the Master's degree in microwaves from Pierre & Marie Curie University, Paris, France, in 1987, and the Ph.D. degree in control and signal processing from Orsay University, Orsay, France, in April 1991.

During the Ph.D. (from November 1989 to April 1991), he was with adaptive filters and fast algorithms at the Centre National d'Etudes des Telecomunications (CNET), Paris, France. From January 1994 to July 1995, he was with Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a Member of the Technical Staff with Bell Laboratories, Murray Hill, NJ, USA. In May 2003, he joined the University of Quebec, INRS-EMT, Montreal, QC, Canada, as a Professor. He is currently a Visiting Professor with the Technion, Haifa, Israel, and an Adjunct Professor with Aalborg University, Aalborg, Denmark. He is the Inventor of many important technologies. In particular, he was the Lead Researcher with Bell Labs who conceived and designed the world-first real-time hands-free full-duplex stereophonic teleconferencing system. Also, he conceived and designed the world-first PC-based multi-party hands-free full-duplex stereo conferencing system over IP networks. He is the Editor of the book series *Springer Topics in Signal Processing*. He has co-authored and co-edited/co-authored numerous books in the area of acoustic signal processing. His research interests include signal processing, acoustic signal processing, and multimedia communications.

Dr. Benesty was the General Chair and the Technical Chair of many international conferences and a member of several IEEE technical committees. Four of his journal papers were awarded by the IEEE Signal processing Society and in 2010 he was the recipient of Gheorghe Cartianu Award from the Romanian Academy.

Tao Yang received the B.E. degree from Xi'an Jiaotong University, Xi'an, China, in 1993, the M.E. degree from South China University of Technology, Guangzhou, China, in 1996, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2003, respectively.

After receiving the M.E. degree in automatic control in 1996, he joined TCL Corporation, Guangdong, China, as a Telecommunication Software Engineer. In 2003, he was a Member of Faculty with Southwest University of Science and Technology, Mianyang, China, where he has been a Professor with the School of Information Engineering since 2011, and the Deputy Director of Robot Technology Used for Special Environment Key Laboratory of Sichuan Province since 2013. He was with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, as a Visiting Scholar engaged in a research project in collaborative signal processing for sensor networks in 2006. His current research interests include signal processing, modeling, simulation and control of electromechanical systems, including acoustic array signal processing for robot scene perception and machine fault diagnosis, and microelectromechanical systems.