

# ADAPTIVE FILTER ALGORITHMS AND MISALIGNMENT CRITERIA FOR BLIND BINAURAL CHANNEL IDENTIFICATION IN HEARING-AIDS

Gerald Enzner, Ivo Merks, Tao Zhang

Starkey Laboratories Inc., 6600 Washington Avenue South, Eden Prairie, MN 55344, USA

## ABSTRACT

Blind channel identification (BCI) is known from communications as a bitrate saving alternative to the more conventional pilot-based identification of source-receiver transfer functions. In multi-microphone signal processing, however, the source signal is per se not available for acoustic channel identification. The emerging binaural signal processing discipline is a good example where BCI may thus be considered necessary for tasks such as acoustic localization or equalization. This paper evaluates current algorithms from BCI in order to make a pair of hearing-aids aware of time-varying head-related transfer functions without knowledge of the source signal. Thereby, unrestricted bitrate is assumed to share both ear signals in a central processor. Using simulations of binaural signals, we explore four two-channel adaptive identification algorithms and three evaluation criteria as candidates for hearing-aids. Depending on the criteria, the study shows striking similarities for specific configurations of algorithm and data, but also reveals important differences.

## 1. INTRODUCTION

In binaural signal processing applications, such as digital hearing-aids, the impulse responses from a point sound source in space to the left and right ear describe the acoustic system. In dynamic environments with changing locations of the target sound source with respect to the human head, these head-related impulse responses (HRIRs) are usually not known to a binaural signal processing algorithm. Their availability would, however, foster the design of new binaural algorithms for acoustic equalization, dereverberation, noise reduction, source localization, or even head tracking. This motivates the online estimation of HRIRs from the left and right microphone signals of, e.g., a pair of hearing aids. Using only this information then requires a blind binaural channel identification approach.

BCI has been explored in digital communications with applications to multi-antenna and oversampled single-sensor processing. It was found that blind channel identifiability up to a gain factor requires the absence of observation noise and common zeros between the channels [1], [2]. BCI was later considered in audio and acoustics [3]. Especially the more recent adaptive approaches based on recursive error signal minimization can be applicable for online acoustic impulse response inference [4], [5], [6]. In addition, the online principal component analysis [7], as used in frequency-domain adaptive beamforming [8], or its translation to a time-domain implementation [9], may provide inherent channel identification.

In the acoustics domain, however, the identifiability conditions are hardly ever met. The dynamic range of, e.g., speech, causes a large range of signal-to-noise ratios. Furthermore, acoustic channels contain random or systematic common zeros (the latter in case of system order overmodeling). Algorithms to overcome the problems related to common zeros and noise have been published, e.g., [10], [11]. More recently, sophisticated evaluation criteria have been developed [12], [13] to absorb the truly ill-conditioned part of the identification problem, so that the better-conditioned and more interesting part can be examined in more depth.

This paper studies the applicability of BCI in binaural hearing-aids. A set of candidate adaptive algorithms for two-channel HRIR identification is outlined with unified notation. In doing so, we aim to cover the range from quasi-supervised adaptive filtering to truly blind adaptive filter design. We characterize relationships between the algorithms and rank their suitability for the problem at hand. Two different single-number metrics for performance evaluation are then presented. By relaxing the usual expectation of strict channel identification up to a gain, we aim to bridge between theory and practice of BCI. Using the relaxed metrics for evaluation across the diverse set of algorithms, we demonstrate that in fact all candidates can achieve similar and surprisingly good performance in the presence of independent and uncorrelated observation noise. In addition to single-number metrics, spectral analysis of blind channel estimates is recommended and presented. We finally test the algorithms and metrics under more realistic acoustic conditions with colored observation noise and point out limitations of current BCI technology.

The paper is organized as follows. Sec. 2 introduces the binaural signal model and Sec. 3 reviews the candidate algorithms for BCI. Sec. 4 then presents misalignment criteria with inherent discussion of results for the independent and uncorrelated observation noise case. For this particular scenario, a proof of the observed equivalence of both single-number metrics is included. Sec. 5 separately discusses the colored noise case and Sec. 6 draws conclusions.

## 2. HEAD-RELATED SIGNAL MODEL

Fig. 1 explains our notation. An unknown signal  $s(k)$  at discrete time  $k$ , due to a sound source at angle  $\varphi$ , is convolved with left and right HRIRs,  $h_{l,k}$  and  $h_{r,k}$ , to yield the binaural signals  $x_l(k)$  and  $x_r(k)$  at the ears. After addition of acoustic observation noises  $n_l(k)$  and  $n_r(k)$ , the signals  $y_l(k)$  and  $y_r(k)$  are available for binaural signal processing with adaptive digital filters  $\hat{h}_{l,k}$  and  $\hat{h}_{r,k}$ . If the adaptive filters match the HRIRs, i.e.,  $\hat{h}_{l/r,k} = h_{l/r,k}$ , the depicted cross-correlation (CR) processing would obviously yield an error signal  $e(k) = 0$  in the absence of observation noise [1], [3], [5].

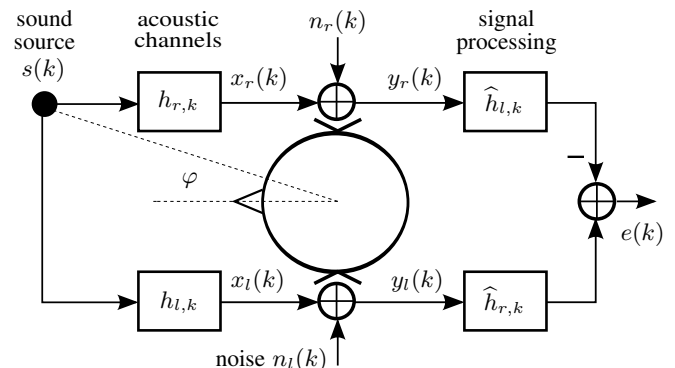


Fig. 1. Binaural reception and adaptive signal processing model.

### 3. TWO-CHANNEL ADAPTIVE ALGORITHMS

Each of the following subsections contains a candidate algorithm for blind estimation of the left and right HRIR, now as vectors

$$\mathbf{h}_{l/r} = [h_{l/r,0} \ h_{l/r,1} \ \cdots \ h_{l/r,L-1}]^T, \quad (1)$$

based on the most recent observations contained in vectors

$$\mathbf{y}_{l/r}(k) = [y_{l/r}(k) \ y_{l/r}(k-1) \ \cdots \ y_{l/r}(k-L+1)]^T. \quad (2)$$

The estimated binaural channel  $\widehat{\mathbf{h}}_{l/r}(k)$  at time  $k$  is defined in accordance with  $\mathbf{h}_{l/r}$ . The linear filtering of Fig. 1 can then be expressed as inner vector products resulting in the error signal

$$e(k) = \widehat{\mathbf{h}}_r^T(k) \mathbf{y}_l(k) - \widehat{\mathbf{h}}_l^T(k) \mathbf{y}_r(k). \quad (3)$$

This error signal will be a common factor of all CR-type algorithms below, but it will not be employed in a PCA-type algorithm. A further common factor of all algorithms will be the unit-norm constraint on the estimated channels, i.e.,

$$\|\widehat{\mathbf{h}}\|_2^2 \triangleq \widehat{\mathbf{h}}_l^T(k) \widehat{\mathbf{h}}_l(k) + \widehat{\mathbf{h}}_r^T(k) \widehat{\mathbf{h}}_r(k) = 1. \quad (4)$$

This constraint will be enforced via normalization after each iteration of the adaptive algorithm loop to ensure that the blind channel estimate is neither falling nor rising to extremes.

#### 3.1. Stereo Least-Mean-Square (LMS) Algorithm

Considering the known inputs to the adaptive filters, we shall first harvest from supervised adaptive filtering [14] in order to achieve blind binaural channel identification. Let us assume low-noise observations  $y_{l/r}(k)$  and exploit the aforementioned implication of nulling the cross-relation error  $e(k)$  by a good channel estimate. Using the known ‘‘supervised’’ LMS algorithm [14], we can then adjust the two-channel adaptive filter  $\widehat{h}_{l/r,k}$  to drive the adaptive filter output, i.e., the signal  $e(k)$  in Fig. 1, to equality with the desired zero. Formally speaking, we minimize the square error  $e^2(k)$  with respect to  $\widehat{h}_{l,k}$  and  $\widehat{h}_{r,k}$  to obtain, via gradient descent, two recursive update rules for acoustic channel estimation:

$$\widehat{\mathbf{h}}_l(k+1) = \widehat{\mathbf{h}}_l(k) + \mu e(k) \mathbf{y}_r(k) \quad (5)$$

$$\widehat{\mathbf{h}}_r(k+1) = \widehat{\mathbf{h}}_r(k) - \mu e(k) \mathbf{y}_l(k). \quad (6)$$

In further analogy with [14], we use a normalized stepsize factor

$$\mu = \mu_0 (\mathbf{y}_l^T(k) \mathbf{y}_l(k) + \mathbf{y}_r^T(k) \mathbf{y}_r(k))^{-1} \quad (7)$$

to control the speed of adaptation with choices of  $0 < \mu_0 < 1$ .

#### 3.2. Adaptive Eigenvalue Decomposition Algorithm (AEDA)

In [4], the author relies on two-channel blind acoustic system identification in order to perform time-delay estimation. The BCI task, subject to unit-norm constraint, is recognized as a minimum eigenvector estimation problem and solved by iterative minimization of the Rayleigh quotient  $R = e^2(k) / \|\widehat{\mathbf{h}}\|_2^2$ . When the unit-norm constraint is maintained in the iterative loop, as described in the preamble of this section, the following version of AEDA can be invoked:

$$\widehat{\mathbf{h}}_l(k+1) = \widehat{\mathbf{h}}_l(k) + \mu e(k) (\mathbf{y}_r(k) + e(k) \widehat{\mathbf{h}}_l(k)) \quad (8)$$

$$\widehat{\mathbf{h}}_r(k+1) = \widehat{\mathbf{h}}_r(k) - \mu e(k) (\mathbf{y}_l(k) - e(k) \widehat{\mathbf{h}}_r(k)). \quad (9)$$

Because of additional terms next to  $\mathbf{y}_{l/r}(k)$  in the updates rules, AEDA obviously represents an extension of the stereo LMS algorithm. The additional terms may, however, not be very intuitive. As the results in Sec. 4 will not indicate large impact of the extension, we leave extensive interpretation aside. In order to take the input signal level into account, we will rely on the same normalized stepsize factor  $\mu$  as shown by (7) for the stereo LMS algorithm.

#### 3.3. Blind Multichannel LMS (MCLMS) Algorithm

The authors of [5] propose the MCLMS algorithm as a generalization of AEDA to efficiently solve  $P$ -channel ( $P \geq 2$ ) blind identification problems by recursive update. They showed for this case that blind identification again can be seen as a minimum eigenvector estimation problem. Generalizing AEDA, the cost function of the adaptive MCLMS algorithm utilizes all pairwise cross-relations between available channel-pairs. In contrast to AEDA, MCLMS relies on fully populated matrix algebra to describe the update rules for all channels. Apart from solving the comprehensive  $P$ -channel identification, an implementation in terms of matrix algebra would unfortunately cause a too high computational load on almost any platform used in online adaptive audio signal processing.

Naturally, MCLMS can also be invoked for two-channel problems. Considering time-varying and data-dependent instantaneous-correlation matrices  $\widetilde{\mathbf{R}}_{y_i y_j}(k) = \mathbf{y}_i(k) \mathbf{y}_j^T(k)$ ,  $i, j \in \{l, r\}$ , and again assuming that unit-norm is enforced explicitly in the iterative loop, we can adopt the MCLMS for our  $P = 2$  application:

$$\begin{aligned} \widehat{\mathbf{h}}_l(k+1) &= \widehat{\mathbf{h}}_l(k) + \\ &+ \mu (\widetilde{\mathbf{R}}_{y_r y_r}(k) \widehat{\mathbf{h}}_l(k) - \widetilde{\mathbf{R}}_{y_r y_l}(k) \widehat{\mathbf{h}}_r(k) - e^2(k) \widehat{\mathbf{h}}_l(k)) \end{aligned} \quad (10)$$

$$\begin{aligned} \widehat{\mathbf{h}}_r(k+1) &= \widehat{\mathbf{h}}_r(k) + \\ &+ \mu (-\widetilde{\mathbf{R}}_{y_l y_r}(k) \widehat{\mathbf{h}}_l(k) + \widetilde{\mathbf{R}}_{y_l y_l}(k) \widehat{\mathbf{h}}_r(k) - e^2(k) \widehat{\mathbf{h}}_r(k)). \end{aligned} \quad (11)$$

Upon substituting the definition of  $\widetilde{\mathbf{R}}_{y_i y_j}(k)$  everywhere in the update rules, we immediately notice for instance in (11) that

$$-\widetilde{\mathbf{R}}_{y_l y_r}(k) \widehat{\mathbf{h}}_l(k) + \widetilde{\mathbf{R}}_{y_l y_l}(k) \widehat{\mathbf{h}}_r(k) = \mathbf{y}_l(k) e(k). \quad (12)$$

Hence, for  $P = 2$ , the MCLMS algorithm is equivalent to AEDA. However, AEDA offers great advantage in that it requires inner-vector-product arithmetic only and is therefore our choice of implementation in later parts of this paper.

#### 3.4. Adaptive Principal Component Analysis (APCA)

This learning algorithm can be derived by maximization of the output power of a filter-and-sum array [8], again subject to the unit-norm constraint on the filters. It has been originally known as Oja’s rule in principal component analysis for neural network processing [7]. Here, we rely on a time-domain interpretation consisting of iterative channel identification and equalization [9].

Denoting by  $\mathbf{h}_{l/r}^{\leftarrow}(k)$  the time-reversed acoustic channels, a two-channel matched-filter array (i.e., channel equalizer) is first employed to yield an estimate  $\widehat{s}(k)$  of the source signal,

$$\widehat{s}(k) = \widehat{\mathbf{h}}_l^{\leftarrow T}(k) \mathbf{y}_l(k) + \widehat{\mathbf{h}}_r^{\leftarrow T}(k) \mathbf{y}_r(k). \quad (13)$$

On this basis, just considering  $\widehat{s}(k)$  in place of  $s(k)$  in Fig. 1, we can formulate quasi-supervised channel update rules to be executed independently for each channel,

$$\widehat{\mathbf{h}}_{l/r}(k+1) = \widehat{\mathbf{h}}_{l/r}(k) + \mu e_{l/r}(k) \widehat{s}(k), \quad (14)$$

based on the individual error signals for left and right,

$$e_{l/r}(k) = y_{l/r}(k-L+1) - \widehat{\mathbf{h}}_{l/r}^T(k) \widehat{s}(k), \quad (15)$$

and the assembly of the most recent equalizer output samples,

$$\widehat{s}(k) = [\widehat{s}(k) \ \widehat{s}(k-1) \ \cdots \ \widehat{s}(k-L+1)]^T. \quad (16)$$

Here, we again use (7) to adjust the stepsize  $\mu$ . The delayed version of the two input signals, i.e.,  $y_{l/r}(k-L+1)$ , has to be used in (15) in order to compensate for  $L$  samples of delay between  $s(k)$  and  $\widehat{s}(k)$  according to the matched-filtering in (13).

#### 4. MISALIGNMENT CRITERIA

This section presents two types of single-number metrics for performance analysis across the two-channel adaptive algorithms. The first type indirectly evaluates the channel estimate via the CR-error, while the second is an explicit impulse-response distance. Because of their coherent results, we further establish the relationship of both metrics and inspect estimated head-related transfer functions (HRTFs).

##### 4.1. Cross-Relation-Error Attenuation (CREA)

A native measure to assess CR-type algorithms is the square-error attenuation with respect to the total input signal variance, i.e.,

$$\text{CREA} = \frac{E\{e^2(k)\}}{E\{y_l^2(k)\} + E\{y_r^2(k)\}}. \quad (17)$$

Here, the expectation operator  $E\{\cdot\}$  can be evaluated by short-time averaging of the respective signal energies, e.g., via recursive or non-recursive smoothing with appropriate time-constants. Apart from proving correct implementation of minimum-eigenvector algorithms, the CREA can also be evaluated for the PCA-type algorithm by mimicking the CR-error in (3) using the observed signals  $y_{l/r}(k)$  with the estimated channels  $\hat{h}_{l/r,k}$  from Sec. 3.4.

In case of very good channel identification, the CREA will saturate according to the signal-to-noise ratio, because of CREA calculation from noisy observations  $y_{l/r}(k)$ . If, under lab conditions, the noise-free binaural signals  $x_{l/r}(k)$  are available, we may further calculate a noise-free CR-error

$$\tilde{e}(k) = \hat{\mathbf{h}}_r^T(k)\mathbf{x}_l(k) - \hat{\mathbf{h}}_l^T(k)\mathbf{x}_r(k) \quad (18)$$

and evaluate the corresponding noise-free error signal attenuation

$$\widetilde{\text{CREA}} = \frac{E\{\tilde{e}^2(k)\}}{E\{x_l^2(k)\} + E\{x_r^2(k)\}}. \quad (19)$$

While the noisy CREA is biased according to the observation noise level, the  $\widetilde{\text{CREA}}$  provides us with an unbiased view on the binaural target-signal cancellation (i.e., blocking) ability and, indirectly, on the channel estimation performance of the algorithms.

Fig. 2 depicts learning curves for the described algorithms. The binaural signals  $x_{l/r}(k)$  were generated by convolving white Gaussian noise with anechoic HRIRs measured on a KEMAR dummy-head (<http://sound.media.mit.edu/resources/KEMAR.html>) and re-sampled to 16 kHz for our simulation. The addition of independent white Gaussian observation noises yields  $y_{l/r}(k)$  at the signal-to-noise ratio  $\text{SNR} = (\sigma_{x_l}^2 + \sigma_{x_r}^2)/(\sigma_{n_l}^2 + \sigma_{n_r}^2)$ , where  $\sigma_{n_l}^2 = \sigma_{n_r}^2$ . The point source starts at  $\varphi = 45^\circ$  (right-front) and its location abruptly changes in the middle of the simulation to  $\varphi = 5^\circ$  (near-front) in order to study reconvergence behavior. In this experimental configuration, it turns out that the performances of the different algorithms can be hardly distinguished from each other. All candidates exponentially converge with similar rate to about 16 dB below the noise floor in the signals. This performance is plausible, by analogy with supervised adaptive filtering, when considering the stepsize factor  $\mu_0 = 0.1$  and adaptive filter length  $L = 128$  used here. In the middle of the simulation, all solutions reconverge at similar rate. It is noteworthy that the blind identification is quite insensitive to the diverse SNR levels at left and right ears for  $\varphi = 45^\circ$  and  $\varphi = 5^\circ$ .

##### 4.2. Normalized Filter-Projection Misalignment (NFPM)

For more explicit impulse response evaluation, we build on the concept of *projection misalignment* which was originally conceived to

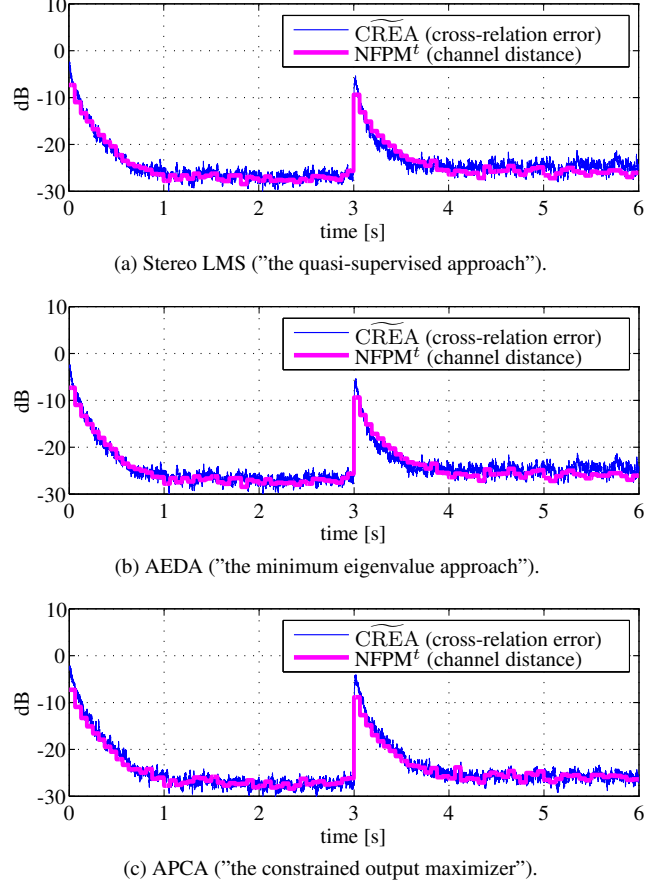


Fig. 2. Learning curves of different algorithms at SNR = 10 dB.

excuse the unavoidable gain error in blind identification [15]. Here, we rely on the *normalized filter-projection misalignment* [13] as a recent generalization with improved usability in the presence of common zeros in the channels. In contrast to original projection misalignment, NFPM essentially absorbs an unavoidable common filter error [9, 12, 13] from the estimated impulse responses. As a result, it indicates whether relative impulse response characteristics are estimated well. This quality may prove to be helpful and in some cases even sufficient to solve binaural signal processing tasks such as time-delay estimation or acoustic signal enhancement.

With reference to [13], we rely on transposed usage of estimated and true channels. We express individual impulse response errors  $\epsilon_{l/r}(k)$  in terms of length  $L + 2D$  zero-padded channels  $\hat{\mathbf{h}}_{l/r}^z = [0 \cdots \hat{\mathbf{h}}_{l/r}^T \cdots 0]^T$  and size  $(L+2D) \times (2D+1)$  convolution matrices

$$\hat{\mathbf{H}}_{l/r} = \begin{bmatrix} \hat{h}_{l/r,0} & 0 & \cdots & \cdots & \cdots & 0 \\ \hat{h}_{l/r,1} & \hat{h}_{l/r,0} & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \hat{h}_{l/r,L-1} & \hat{h}_{l/r,L-2} \\ 0 & \cdots & \cdots & \cdots & 0 & \hat{h}_{l/r,L-1} \end{bmatrix} \quad (20)$$

formed from the estimated channels, i.e.,

$$\epsilon_{l/r}(k) = \mathbf{h}_{l/r}^z - \hat{\mathbf{H}}_{l/r}(k) \mathbf{f}(k). \quad (21)$$

The common filter  $\mathbf{f}$  and NFPM are then defined via least-squares as

$$\text{NFPM}^t(k) = \min_{\mathbf{f}(k)} (\|\epsilon_l(k)\|_2^2 + \|\epsilon_r(k)\|_2^2) / \|\mathbf{h}\|_2^2. \quad (22)$$

NFPM can be applied universally to all channel estimators described in this paper. Fig. 2 depicts its evolution as a function of time along with the cross-correlation error. For the same data as before, it can be seen that NFPM almost perfectly coincides with CREA, as mean-square error signal and impulse-response misalignment do in supervised adaptive filtering when broadband input is used [14].

We finally confirm that an evaluation of the algorithms in terms of the just gain-absorbing normalized projection misalignment [15] does not provide insight into the algorithms for the data at hand. It resides only slightly below 0 dB and is thus not depicted here.

### 4.3. Relationship between CREA and NFPM

We shall not leave the observed and possibly intuitive coincidence of  $\widehat{\text{CREA}}$  and NFPM to the experiment alone. Hence, in this section, we formally confirm their relationship.

Define estimated binaural signals  $\widehat{\mathbf{x}}_{l/r}(k) = \mathbf{s}^T(k)\widehat{\mathbf{H}}_{l/r}(k)$  via length  $(2L+2D-1)$  source vectors  $\mathbf{s}(k) = [\dots s(k) s(k-1) \dots]^T$  and size  $(2L+2D-1) \times (L+2D)$  convolution matrices formed from estimated channels as before. Further invoking commutativity of the linear filters in Fig. 1, the noise free cross-correlation error as shown by (18) equivalently reads

$$\begin{aligned} \tilde{e}(k) &= \widehat{\mathbf{x}}_r^T(k)\mathbf{h}_l^z(k) - \widehat{\mathbf{x}}_l^T(k)\mathbf{h}_r^z(k) \\ &= -\mathbf{s}^T(k)\widehat{\mathbf{H}}^o(k)\mathbf{h}^z(k), \end{aligned} \quad (23)$$

where  $\widehat{\mathbf{H}}^o = [-\widehat{\mathbf{H}}_r, \widehat{\mathbf{H}}_l]$  and  $\mathbf{h}^z = [\mathbf{h}_l^z, \mathbf{h}_r^z]^T$ . Also, we can rewrite (21) more compactly as  $\mathbf{h}^z = \boldsymbol{\epsilon} + \widehat{\mathbf{H}}\mathbf{f}$ , where  $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_l^T, \boldsymbol{\epsilon}_r^T]^T$  and  $\widehat{\mathbf{H}} = [\widehat{\mathbf{H}}_l^T, \widehat{\mathbf{H}}_r^T]^T$ . Then substituting (21) into (23), the cross-correlation error is expressed directly in terms of  $\boldsymbol{\epsilon}(k)$ , i.e.,

$$\begin{aligned} \tilde{e}(k) &= -\mathbf{s}^T(k)\widehat{\mathbf{H}}^o(k)(\boldsymbol{\epsilon}(k) + \widehat{\mathbf{H}}(k)\mathbf{f}(k)) \\ &= -\mathbf{s}^T(k)\widehat{\mathbf{H}}^o(k)\boldsymbol{\epsilon}(k), \end{aligned} \quad (24)$$

by exploiting the orthogonality of  $\widehat{\mathbf{H}}^o$  and  $\widehat{\mathbf{H}}$ , i.e.,  $\widehat{\mathbf{H}}^o\widehat{\mathbf{H}} = \mathbf{0}$ .

On this basis, we proceed to evaluate the mean-square error as

$$\begin{aligned} E\{\tilde{e}^2(k)\} &= E\{\boldsymbol{\epsilon}^T(k)\widehat{\mathbf{H}}^{oT}(k)\mathbf{s}(k)\mathbf{s}^T(k)\widehat{\mathbf{H}}^o(k)\boldsymbol{\epsilon}(k)\} \\ &= \sigma_s^2 ( \|\widehat{\mathbf{h}}_r(k)\|_2 \|\boldsymbol{\epsilon}_l(k)\|_2 + \|\widehat{\mathbf{h}}_l(k)\|_2 \|\boldsymbol{\epsilon}_r(k)\|_2 )^2 \\ &= \sigma_s^2 ( \|\boldsymbol{\epsilon}_l(k)\|_2^2 + \|\boldsymbol{\epsilon}_r(k)\|_2^2 ) \end{aligned} \quad (26)$$

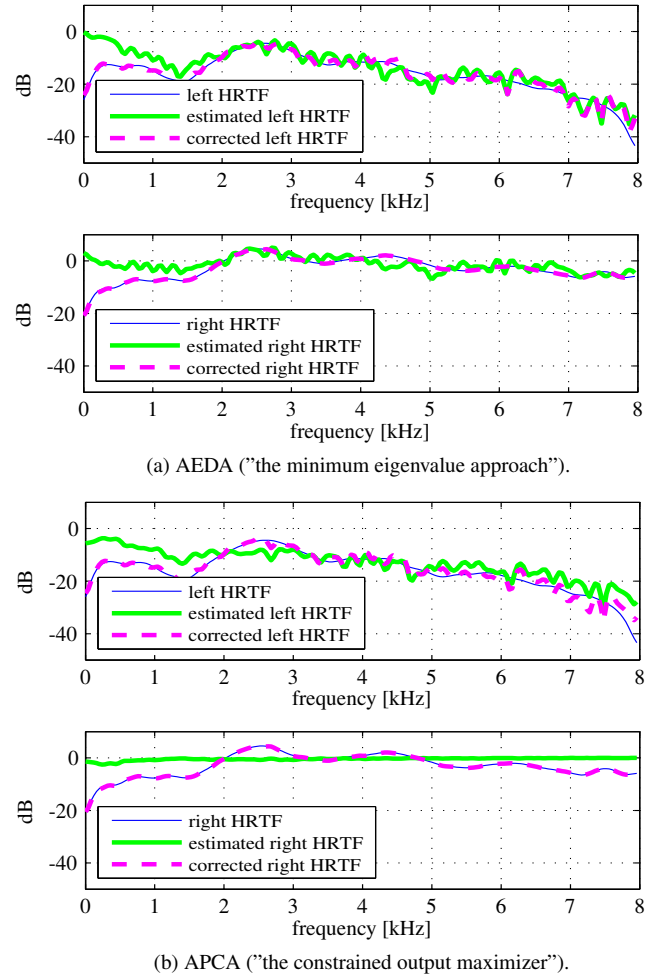
assuming white noise input  $s(k)$  with variance  $\sigma_s^2$ . The amplitude addition in (26) is applied because of strong correlation between filtered-impulse-response errors  $\boldsymbol{\epsilon}_l$  and  $\boldsymbol{\epsilon}_r$  after two-channel NFPM regression in (22). The step towards (27) firstly exploits our observation of transposed proportionality of channel error and channel estimate after the NFPM regression, i.e.,  $\|\boldsymbol{\epsilon}_l(k)\|_2 = \alpha\|\widehat{\mathbf{h}}_r(k)\|_2$  and  $\|\boldsymbol{\epsilon}_r(k)\|_2 = \alpha\|\widehat{\mathbf{h}}_l(k)\|_2$ , and secondly the unit-norm constraint according to (4) is taken into account. The equivalence of  $\widehat{\text{CREA}}$  in (19) and NFPM in (22) is eventually seen by recognizing the normalization of  $E\{\tilde{e}^2(k)\}$  to  $E\{x_l^2(k)\} + E\{x_r^2(k)\} = \sigma_s^2 \|\mathbf{h}\|_2^2$ .

### 4.4. Spectral Analysis of Blind Channel Estimates

Depending on the target application, e.g., acoustic equalization or signal enhancement, single-number metrics alone may not be regarded as sufficient criteria to judge the applicability of BCI. Moreover, we want to gain further insight into performance differences of the blind estimators under investigation. For more in-depth considerations, we therefore delve into spectral characteristics of true and estimated channels,  $\mathbf{h}_{l/r}$  and  $\widehat{\mathbf{h}}_{l/r}(\infty)$ , respectively. A further quantity of interest, the product  $\widehat{\mathbf{H}}_{l/r}\mathbf{f}$  as defined via the NFPM calculation in (21) and (22), will be termed the common-filter-corrected

channel estimate. For the sake of brevity, we limit our following presentation to AEDA and APCA, because we observed that the stereo LMS algorithm again behaves very similar to AEDA.

Fig. 3 shows results for experimental conditions as before, but only for right-front source location. AEDA tends to estimate magnitude spectral characteristics more accurately than APCA, but the low-frequency range is not well represented by both estimators. APCA tends to adapt a flat transfer function for the ipsilateral (right) HRIR channel. This behavior is reasonable when considering that the algorithm, with the help of a unit-filter, simply selects the dominant ear signal as the principal component. The contralateral impulse response is adjusted accordingly to represent relative channel characteristics well. The latter is expected from the NFPM results and confirmed here by the observation that the common filter  $\mathbf{f}$  obviously corrects the channel estimates into the actual channels.



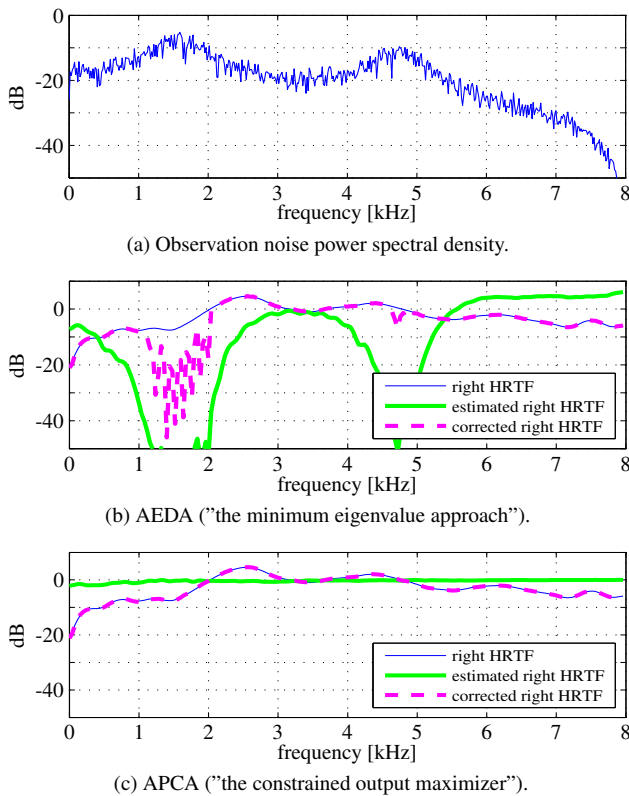
**Fig. 3.** Spectral analysis of the blindly estimated binaural channels in white observation noise. SNR = 10 dB.  $\varphi = 45^\circ$ .  $L = 128$ .

## 5. COLORED OBSERVATION NOISE CONDITIONS

More realistic conditions are of great interest in audio and acoustic signal processing. We thus expose the selected algorithms, AEDA and APCA, to a colored observation noise simulation, while using the same source and binaural signals as before. As seen from Fig. 4a, we introduce acoustic resonances as they may occur in ambient noise. However, we do not introduce binaural coherence to avoid possible ambiguities in the interpretation of the results.

Fig. 4b (right ear only) shows that the performance of AEDA is strongly affected by the coloration of the noise. The energy of the adaptive filters is obviously pushed towards the low-energy bands of the observation noise. The algorithm in fact utilizes the degrees of freedom related to the common-filter error to achieve strong error signal attenuation, while maintaining the unit-norm constraint at the same time. Not even the common-filter correction can perfectly restore the original channel. It further turns out that the assumptions stated after (27) are violated and, hence,  $CREA \ll NFPM$ .

APCA in Fig. 4c is found to be more robust in the sense that the estimate of the dominant channel flattens out as seen for white observation noise. The left HRTF (not depicted for brevity) is again adjusted accordingly so that the common-filter correction yields a very nice match with the true channel. In terms of the single-number criteria, we obtain  $CREA \approx NFPM \approx 28$  dB and even the convergence rate is comparable to the white-noise case in Fig. 2.



**Fig. 4.** Spectral analysis of blind channel estimates in the presence of colored observation noise.  $SNR = 10$  dB,  $\varphi = 45^\circ$ ,  $L = 128$ .

## 6. CONCLUSIONS

BCI has been a delicate approach in acoustic signal processing, since the strict identifiability conditions are not met. This paper has put recent advances in algorithm design and evaluation metrics into perspective in order to better judge the applicability of BCI for binaural channel identification. We presented "minimum-eigenvector estimation" and "principal component analysis" as two major algorithm classes. Under controlled conditions with white observation noise, the results for both classes indicate usability of BCI. Especially the normalized filter-projection misalignment, which absorbs the truly ill-conditioned part of the BCI problem, coherently lines up the good algorithm performances. Under the more realistic condition of colored observation noise, minimum-eigenvector analysis turns out to be very sensitive in terms of channel misidentification. Here, the

principal component analysis indicates more inherent robustness, but still it detects only the relative channel information very well.

## 7. ACKNOWLEDGMENTS

The authors thank D. Schmid and S. Vishnubhotla for comments and discussion. The first author thanks Starkey for providing an excellent working environment during a research visit in fall 2011.

## 8. REFERENCES

- [1] 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.
- [2] 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.
- [3] S. Gannot and M. Moonen, "Subspace methods for multicrophone speech dereverberation," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 11, pp. 1074–1090, 2003.
- [4] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *J. Acoust. Soc. Am.*, vol. 107, no. 1, pp. 384–391, Jan. 2000.
- [5] Y. Huang and J. Benesty, "Adaptive multi-channel least mean square and Newton algorithms for blind channel identification," *Signal Proc.*, vol. 82, no. 8, pp. 1127–1138, Aug. 2002.
- [6] R. Ahmad, A. W. H. Khong, M. K. Hasan, and P. A. Naylor, "An extended normalized multichannel FLMS algorithm for blind channel identification," in *Proc. European Signal Process. Conf.*, Florence, Italy, Sept. 2006.
- [7] A. Hyvärinen, J. Karhunen, and E. Oja, *Principal Component Analysis*, John Wiley & Sons, New York, USA, 2001.
- [8] E. Warsitz and R. Haeb-Umbach, "Acoustic filter-and-sum beamforming by adaptive principal component analysis," in *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Mar. 2005, pp. iv/797 – iv/800.
- [9] D. Schmid and G. Enzner, "Robust subsystems for iterative multichannel blind system identification and equalization," in *Proc. IEEE Pacific Rim Conf. on Commun., Comput. and Signal Process.*, Victoria, Can., Aug. 2009, pp. 889–893.
- [10] N. D. Gaubitch, M. K. Hasan, and P. A. Naylor, "Noise robust adaptive blind channel identification using spectral constraints," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process.*, Toulouse, France, May 2006, pp. 93–96.
- [11] N. D. Gaubitch, J. Benesty, and P. A. Naylor, "Adaptive common root estimation and the common zeros problem in blind channel identification," in *Proc. European Signal Process. Conf.*, Antalya, Turkey, Sept. 2005.
- [12] M. R. P. Thomas, N. D. Gaubitch, E. A. P. Habets, and P. A. Naylor, "Supervised identification and removal of common filter components in adaptive blind SIMO system identification," in *Proc. Int. Workshop on Acoust. Echo and Noise Control*, Tel Aviv, Israel, Sept. 2010.
- [13] D. Schmid and G. Enzner, "Cross-relation-based blind SIMO identifiability in the presence of near-common zeros and noise," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 60–72, Jan. 2012.
- [14] S. S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, USA, 4th edition, 2002.
- [15] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Process. Lett.*, vol. 5, no. 7, pp. 174–176, July 1998.