Subband Scale Factor Ambiguity Correction Using Multiple Filterbanks

Boaz Castro[1], *Nikolay Gaubitch*[2], *Emanuël Habets*[2], *Sharon Gannot*[1], *Patrick Naylor*[2], and *Steven Grant*[3]

[1]School of Engineering
Bar-Ilan University, Ramat-Gan, Israel
bo0701@bezeqint.net, gannot@eng.biu.ac.il

[2]Department of Electrical and Electronic Engineering
Imperial College London, SW7 2AZ, UK
{ndg,e.habets,p.naylor}@imperial.ac.uk

[3]Missouri University of Science and Technology, USA
sgrant@mst.edu

***Abstract*—One of the problems with blind system identification in subbands is that the subband systems can only be identified correctly up to an arbitrary scale factor. This scale factor ambiguity is the same across all channels but can differ between the subbands and therefore, limits the usability of such estimates. In this contribution, a method that uses multiple filterbanks is proposed that utilizes overlapping passband regions between these filterbanks to find scalar correction factors that make the scale factor ambiguity uniform across all subbands. Simulation results are provided, showing that the proposed method accurately identifies and corrects for these scale factors at the cost of an increased computational burden.**

I. Introduction

Blind system identification (BSI) is of interest in several fields of engineering including communications, exploration seismology and speech and audio processing [1]. Within the area of speech and audio processing BSI is an important component for speech dereverberation [1]–[3]. The BSI problem can be stated as follows: consider a signal $s(n)$ which is produced in a multipath environment such as a reverberant room and transmitted to an array of M sensors at a distance from the source. The observed signal at the mth sensor is then given by

$$x_m(n) = \mathbf{h}_m^T \mathbf{s}(n) + \nu_m(n), \tag{1}$$

where $\mathbf{h}_m = [h_{m,0} \ h_{m,1} \ \ldots \ h_{m,L-1}]^T$ is the L-tap impulse response between the source and the mth sensor, $\mathbf{s}(n) = [s(n) \ s(n-1) \ \ldots \ s(n-L+1)]^T$ is the input signal vector, $\nu_m(n)$ is additive measurement noise and $[\cdot]^T$ denotes matrix transpose. The problem of BSI is to find the impulse responses $\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \ldots \ \mathbf{h}_M^T]^T$ using only the observations $x_m(n)$. Most existing BSI methods find the channel responses up to an arbitrary scale factor; the estimates, $\widehat{\mathbf{h}}$, are given by

$$\widehat{\mathbf{h}} = \alpha \mathbf{h}. \tag{2}$$

The accuracy of the estimates is degraded by observation noise and long impulse responses, commonly occurring in acoustic signal processing. Long impulse responses also result in large computational complexity and lower estimation performance due to increased occurrence of common or near-common zeros [4].

One way to improve the identification performance is to use a subband framework; this has shown benefits in terms of computational complexity and performance improvement in applications such as acoustic echo cancellation [5]. A study of a subband BSI system was presented in [2], highlighting the benefit of the shorter channels to be estimated compared with the full-band case. However, BSI in subbands has received much less attention than the non-blind case (for example, acoustic echo cancellation) and one reason for this is the scale factor ambiguity across the different subbands which limits the use of such approaches [2].

A method to correct the scale factor ambiguity in oversampled subband systems was proposed in [6], where the authors find the scale correction factors using a full-band cross-relation error operating on the observations signals together with the full-band impulse responses that are reconstructed from the subband estimates. Although it provides very accurate results if the subband channel estimates are accurate, it degrades rapidly as the level of estimation error increases.

In this paper, we propose an alternative method to correct this scale factor ambiguity. This method uses multiple filterbanks such that an overlapping passband region (OPR) between subbands and subband systems is created; this facilitates the calculation of scale correction factors. Since the proposed method operates solely on the estimated subband channels it is much less sensitive to estimation errors compared to that in [6]. The remainder of this paper is organized as follows. In Section II, the problem of scale correction factor estimation is formulated. The scale factor ambiguity correction (SFAC) method is then developed in Section III and simulation results evaluating its performance are presented in Section IV. Finally, conclusions from this work are drawn in Section V.

II. Problem Formulation

Consider a filterbank with K subbands decimated by a factor of N. Assuming that the identifiability conditions are satisfied[1], each subband estimate $\widehat{\mathbf{h}}_k$ (where k is the subband

[1]It is still an open question whether identifiability conditions can be met in subbands, since for non-critically decimated filterbank, the systems are not fully-excited.

index) will be determined up to a complex scale factor α_k such that

$$\widehat{\mathbf{h}}_k = \alpha_k \mathbf{h}_k, \quad k = 0, 1, \ldots, K-1. \tag{3}$$

It is evident from (3) that for a particular k the scale factor will be the same across the M channels but for any particular channel will be different across the K subbands. If the subband estimates are used to equalize the observed signal, the scaling discrepancy will propagate to the reconstructed full-band signal. Alternatively, if a full-band impulse response is reconstructed from the subband estimates, the reconstructed impulse response will be incorrect. The correction of this inter-band scale factor ambiguity does not require identification of the exact scale factors; it is sufficient to process the subband estimates such that the scale factors are the same across all K subbands. Consequently, we here define the (SFAC) in terms of a correction terms β_k that can take on any values which result in

$$\beta_0\alpha_0 = \beta_1\alpha_1 = \ldots = \beta_{K-1}\alpha_{K-1}.$$

The objective of SFAC is to find the coefficients β_k for $k \in \{0, 1, \ldots, K-1\}$. Note that the scale factors can be real or complex, depending on the choice of filterbank structure.

III. Scale Factor Ambiguity Correction Method

In this section, we present a novel SFAC method. Firstly, the principle of the proposed method is described. Secondly, the algorithm to find the SFAC coefficients is provided.

Here the generalized discrete Fourier transform (GDFT) filterbank with K subbands decimated by a factor of N is employed. The advantages of the GDFT filterbank include straightforward implementation of fractional oversampling $(N \leq K)$ and computationally efficient implementations [5]. Within the framework of the GDFT filterbank, the coefficients of the kth analysis filter, $\mathbf{u}_k = [u_{k,0}, u_{k,1} \ldots u_{k,,L_{\mathrm{pr}}-1}]^T$, are calculated from a single L_{pr}-tap prototype filter, q_i, with bandwidth $\frac{2\pi}{K}$ according to

$$u_{k,i} = q_i \mathrm{e}^{j\frac{2\pi}{K}(k+k_0)(i+i_0)}, \quad i = 0, 1, \ldots, L_{\mathrm{pr}} - 1, \tag{4}$$

where the frequency and time offset terms were set to $i_0 = 0$ and $k_0 = 1/2$ as in [7]. A corresponding set of synthesis filters, $v_{k,i}$ satisfying near perfect reconstruction is obtained from the time-reversed, conjugated version of the analysis filters [5], $v_{k,i} = u^*_{k,L_{\mathrm{pr}}-i-1}$. The oversampled subband structure allows aliasing between adjacent subbands to be suppressed to a very low level (around -90 dB in our implementation); this eliminates the need for cross-band filters and the full-band transfer function, $H_m(z)$, can be related to a set of subband filters $H_{m,k}(z),\ k = 0, 1, \ldots, K/2 - 1$ with only one filter per subband [7].

A. Principle

The proposed SFAC algorithm works with multiple overlapping filterbank systems. The incentive for using multiple filterbanks is to create an OPR between subband filters. It is argued that cross-relation based BSI algorithms (e.g. [2]), applied at different subbands, will yield estimated filters with almost identical frequency responses up to a scale factor in the OPRs. The goal of our method is therefore to resolve these gain factors by comparing multiple estimates at the same frequency due to BSI in overlapping subband filters. It should be stressed that the proposed method involves BSI procedure in multiple filterbanks, and hence impose a significant computational burden increase. The frequency response of the analysis filters and the OPRs are depicted in Fig. 1. Note, that only frequency bins for which the subband frequency response is approximately 0 dB are considered. We suggest to find the SFAC coefficient by minimizing the Euclidian distance between the frequency responses of the system in overlapping subbands from different filterbanks in the OPR.

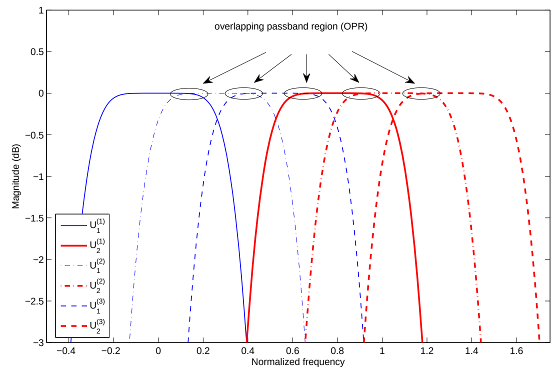


Fig. 1. Multiple filterbanks showing the OPRs.

It is important to note that the solution of estimating the subband systems is blind to the analysis filters since the roots of the latter filters constitute common zeros for all channels. The main objective of the current paper is the derivation of the SFAC algorithm. As in [7], we therefore replace the estimated filters in the subband domain by the LS equivalent subband filters obtained from the true full-band channels. To keep the exposition as realistic as possible, estimation errors are emulated by additive noise. A comprehensive method consisting of both estimation procedure and SFAC is a subject for future research.

The SFAC coefficients are calculated in the full-band domain, namely after upsampling the subband channels and filtering by the synthesis filters (but before subbands summation). The analysis and synthesis filters, due to their inherent ripple, insert small gain differences in the identified subband systems. These spurious gain differences should be compensated before calculating the SFAC coefficients[2]. Using the above considerations the kth subband at the pth filterbank system is given by:

$$u^{(p)}_{k,i} = q_i\, \mathrm{e}^{j\frac{2\pi}{K}(k+\frac{1}{2}+p/P)i};\ p = 0, \ldots, P-1. \tag{5}$$

It can be easily verified that these analysis filters have OPRs. Let $\mathbf{U}^{(p)}_k = [U^{(p)}_k(\omega_0), \ldots, U^{(p)}_k(\omega_{L_{\mathrm{DFT}}-1})]^T$ denote the discrete Fourier transform (DFT) of the impulse response of the

[2]Since the LS equivalent subband filters are based on the full-band system, they are all in-phase, rendering phase compensation unnecessary.

analysis filters $u_{k,i}^{(p)}$. Note that the DFT length L_{DFT} should be larger than (or equal to) the length of the resulting filters after upsampling and convolution. The same discussion also applies to the synthesis filters. The transfer functions of the synthesis filters $\mathbf{V}_k^{(p)}$ are defined in a similar fashion. All DFT operations in the sequel should be read as L_{DFT} points DFTs.

B. Algorithm

The algorithm consists of three steps. In *Step 1*, P different filterbanks are created. In *Step 2*, we estimate the subband systems and compute their DFTs in the full-band domain using the analysis and synthesis filters. In addition, we eliminate the influence of the analysis and synthesis filters. In *Step 3* the SFAC coefficient between adjacent subbands can be estimated using a weighted least square estimation (WLSE) approach. Here the weighting is used to specify the OPR in which the least square error is minimized.

Denote the estimated (or alternatively, the LS equivalent [7]) subband systems as $\widehat{\mathbf{h}}_{r,m}$, where $r = k \cdot P + p$. The full-band representation of these estimates is denoted by $\widetilde{\mathbf{h}}_{r,m}$ with the respective transfer function $\widetilde{\mathbf{H}}_{r,m}$, $m = 1, \ldots, M$. Denote also the transfer function of the ripple compensated filters as $\widetilde{\mathbf{H}}_{\text{c},r,m}$. Finally, a concatenation of all channel estimates is given by $\widetilde{\mathbf{H}}_r = [\widetilde{\mathbf{H}}_{r,1}^T, \widetilde{\mathbf{H}}_{r,2}^T, \cdots, \widetilde{\mathbf{H}}_{r,M}^T]^T$. $\widetilde{\mathbf{H}}_{\text{c},r}$ is defined in a similar fashion. In total we need to find $K/2$ SFAC coefficients. The $(r+1)$th SFAC coefficient is given by

$$\beta'_{r+1} = \text{WLSE}\left(\mathbf{w}_r, \widetilde{\mathbf{H}}_r, \widetilde{\mathbf{H}}_{\text{c},r+1}\right) = \underset{\beta}{\text{argmin}} \left\| \mathbf{w}_r \cdot \left(\widetilde{\mathbf{H}}_r - \beta\, \widetilde{\mathbf{H}}_{\text{c},r+1}\right) \right\|^2, \tag{6}$$

where here $\widetilde{\mathbf{H}}_{r,m}$ is after gain factor elimination by β'_r (calculated in the previous step) and $\mathbf{w}_r$ is the weight vector of length $M \times L_{\text{DFT}}$ with ones in the OPRs and zeros elsewhere. We define the OPR of the mth channel as the $2B+1$ bins wide region, where $B = \lfloor L_{\text{DFT}}/(6KP) \rfloor$. The symmetry axis is the intersection point of adjacent subband filters, satisfying

$$\eta_{r,m} = \text{round}\left\{ \frac{L_{\text{DFT}}}{2K} + \left(r + \frac{1}{2}\right) \frac{L_{\text{DFT}}}{KP} + (m-1)L_{\text{DFT}} \right\}. \tag{7}$$

To obtain larger overlap regions we can also use filterbanks with a decimation factor closer to the critical decimation factor. In this work, a decimation factor of half the critical decimation factor was used. It should be noted that the SFAC coefficient between the last and the first filters is not calculated here. However, when the correct SFAC coefficients have been identified this gain should be equal to 1. The proposed algorithm for estimating the SFAC coefficients is summarized in Algorithm 1. After estimating the SFAC coefficients, we can either equalize the reverberant signal in the subband domain, or extract the full-band filter for full-band equalization [6].

IV. Performance Evaluation

We now provide simulation results to demonstrate the performance of the proposed algorithm, comparing it to the method in [6] and to the case of no scale factor correction.

Step 1: Compute analysis and synthesis filters.
for $k = 0, 1, \ldots, K/2 - 1$ **do**
  for $p = 0, 1, \ldots, P - 1$ **do**

$$u_{k,i}^{(p)} = q_i\, \mathrm{e}^{j\frac{2\pi}{K}(k+\frac{1}{2}+p/P)i}$$
$$v_{k,i}^{(p)} = \left(u_{k,L_{\text{pr}}-i-1}^{(p)}\right)^*$$
$$\mathbf{U}_k^{(p)} = \text{DFT}\left(u_{k,i}^{(p)}\right) \text{ and } \mathbf{V}_k^{(p)} = \text{DFT}\left(v_{k,i}^{(p)}\right)$$

  end
end

Step 2: Estimate subband systems and compute frequency response in the full-band domain.
For $\mathbf{h}, \mathbf{U}, \mathbf{V}, \mathbf{H}$ and $\mathbf{H}_\text{c}$ use: $\{\circ\}_r \triangleq \{\circ\}_{\lfloor r/(K/2) \rfloor}^{(\text{mod}(r,P))}$.
begin
  for $k = 0, 1, \ldots, K/2 - 1$ **do**
    for $p = 0, 1, \ldots, P - 1$ **do**
      - Define $r = k \cdot P + p$
      - Estimate subband systems $\widehat{\mathbf{h}}_{r,m}$ and represent in the full-band domain:
      for $m = 1, 2, \ldots, M$ **do**

$$\widetilde{\mathbf{h}}_{r,m} = \left((\mathbf{u}_r)_{\downarrow N} * \widehat{\mathbf{h}}_{r,m}\right)_{\uparrow N} * \mathbf{v}_r$$

        - Compute the frequency response

$$\widetilde{\mathbf{H}}_{r,m} = \text{DFT}\left(\widetilde{\mathbf{h}}_{r,m}\right)$$

        - Ripple compensation:

$$\widetilde{\mathbf{H}}_{\text{c},r,m} = \widetilde{\mathbf{H}}_{r,m} \odot \mathbf{U}_r \oslash \mathbf{U}_{r+1} \odot \mathbf{V}_r \oslash \mathbf{V}_{r+1}$$

        where $\odot, \oslash$ denote element-wise multiplication and division, respectively.
      end

$$\widetilde{\mathbf{H}}_r = [\widetilde{\mathbf{H}}_{r,1}^T, \widetilde{\mathbf{H}}_{r,2}^T, \cdots, \widetilde{\mathbf{H}}_{r,M}^T]^T$$
$$\widetilde{\mathbf{H}}_{\text{c},r} = [\widetilde{\mathbf{H}}_{\text{c},r,1}^T, \widetilde{\mathbf{H}}_{\text{c},r,2}^T, \cdots, \widetilde{\mathbf{H}}_{\text{c},r,M}^T]^T$$

    end
  end
end

Step 3: Sequentially estimate the SFAC coefficients in the full-band domain.
- **Set** $\beta'_0 = 1$.
begin
  for $r = 0, 1, \ldots, P(K/2 - 1) - 1$ **do**
    - Compute SFAC coefficient for the estimated subband system $r + 1$:

$$\beta'_{r+1} = \text{WLSE}\left(\mathbf{w}_r, \widetilde{\mathbf{H}}_r, \widetilde{\mathbf{H}}_{\text{c},r+1}\right)$$

    - $\mathbf{w}_r$ is a vector with ones in the OPR of the M channels and zeros elsewhere.
    - Correct the gain in adjacent subband:

$$\widetilde{\mathbf{H}}_{r+1} = \beta'_{r+1} \widetilde{\mathbf{H}}_{r+1}$$

  end
end
- The SFAC coefficients are given by: $\beta_k = \beta'_{Pk}$ for $k = 0, 1, \ldots, K/2 - 1$.

Algorithm 1: Summary of the proposed SFAC algorithm.

The filterbank used for the following experiments comprises of $K = 8$ subbands with a decimation factor $N = 4$. An $L_{\text{pr}} = 64$-tap prototype filter was designed using the iterative least squares method [5], giving an estimated aliasing suppression of 92 dB. The source signal was white Gaussian noise with a system of $M = 2$ randomly generated channel responses of length $L = 1024$ whose tap values were drawn from a zero-mean Gaussian distribution and are multiplied by an exponentially decaying function. We then used the full-band channel responses and the method in [7] to find the equivalent subband filters; these were used as the subband channel estimates, $\widehat{\mathbf{h}}_k$. Scale coefficients α_k were generated randomly for each subband and white Gaussian noise was added to each subband to simulate different levels of estimation error. We used two different metrics in our evaluation. First, the normalized projection misalignment (NPM) was employed to measure the misalignment between two impulse responses (disregarding the full-band scale factor) and is defined as [8]

$$\text{NPM} = 20\log_{10}\left(\frac{1}{||\mathbf{h}||}\left\|\mathbf{h} - \frac{\mathbf{h}^T\widehat{\mathbf{h}}}{\widehat{\mathbf{h}}^T\widehat{\mathbf{h}}}\widehat{\mathbf{h}}\right\|\right)\text{dB}. \quad (8)$$

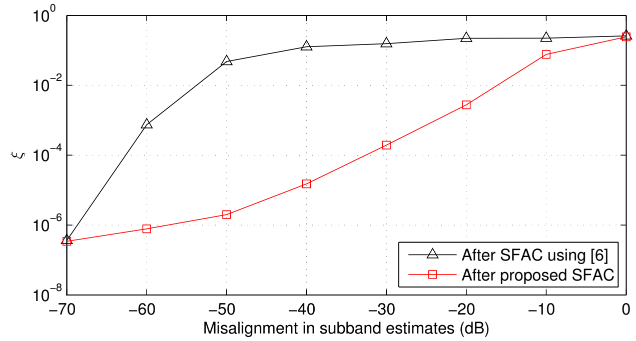
The NPM is applied to the full-band impulse responses which are reconstructed from the subband estimates using the approach in [7]. Second, we used the normalized sample variance of the corrected scale factors, defined as

$$\xi = \frac{\frac{1}{K/2}\sum_{k=0}^{K/2-1}(\alpha_k\beta_k - \mu)^2}{||\mathbf{A}\boldsymbol{\beta}||^2}, \quad (9)$$

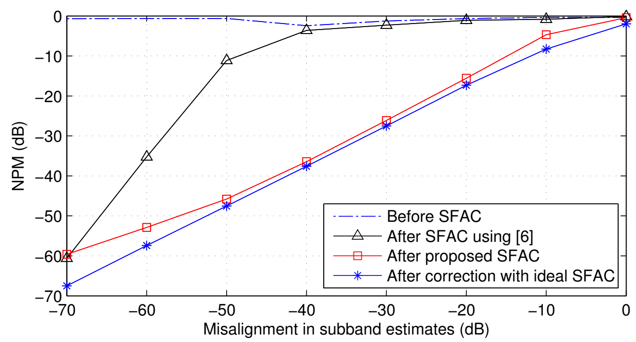
where $\mathbf{A} = \text{diag}\{\alpha_0\ \alpha_1\ \ldots\ \alpha_{K/2-1}\}$ is a diagonal matrix with the true scale factors, $\boldsymbol{\beta} = [\beta_0\ \ldots\ \beta_{K/2-1}]^T$ are the correction coefficients and $\mu = \frac{1}{K/2}\sum_{k=0}^{K/2-1}\alpha_k\beta_k$. If the parameters β_k correct the scale ambiguity such that the scale factor is uniform over all subbands, then $\xi = 0$. The results in terms of normalized variance for the proposed algorithm and for that of [6] are shown in Fig. 2(a). Figure 2(b) shows the results for the two algorithms in terms of NPM; this figure also includes the cases of no scale factor correction and the case of exact scale factor correction. The need for scale factor correction is confirmed, which shows that the reconstructed full-band impulse response does not match the actual system, even when the subband estimates are otherwise accurate. Furthermore, it can be seen that the proposed algorithm and that of [6] achieve the scale factor correction but the method in [6] rapidly deteriorates with increased misalignment. On the contrary, the proposed method is accurate in correcting for the scale factor ambiguity for a wide range of subband channel misalignment values, where its performance is close to that of the ideal scale factor correction.

V. Conclusion

We have presented a method for scale factor ambiguity correction in subband BSI. The method uses multiple filterbanks to create OPRs between different subband estimates and uses them to calculate a set of scale factor correction terms. Simulation results using a GDFT filterbank showed that the proposed method accurately corrects the scale factor differences between subbands for a wide range of subband channel estimation errors, with a performance close to the optimal case of ideal scale factor correction terms. The imposed computational burden is P times larger due to the use of multiple filterbanks.



(a) Normalized variance of the corrected scale factors.



(b) NPM of the full-band impulse responses.

Fig. 2. Metrics to evaluate the SFAC algorithm as a function of the misalignment.

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] S. Gannot and M. Moonen, "Subspace methods for multimicrophone speech dereverberation," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 11, pp. 1074–1090, 2003.

[3] Y. Huang, J. Benesty, and J. Chen, "A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 5, pp. 882–895, Sep. 2005.

[4] A. W. H. Khong, X. Lin, and P. A. Naylor, "Algorithms for identifying clusters of near-common zeros in multichannel blind system identification and equalization," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, USA, Apr. 2008, pp. 389–392.

[5] S. Weiss and R. W. Stewart, *On adaptive filtering in oversampled subbands*. Shaker Verlag, 1998.

[6] N. D. Gaubitch, X. Lin, and P. A. Naylor, "Scale factor ambiguity correction for subband blind multichannel identification," in *Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC)*, Seattle, Washington, USA, Sep. 2008.

[7] J. P. Reilly, M. Wilbur, M. Seibert, and N. Ahmadvand, "The complex subband decomposition and its application to the decimation of large adaptive filtering problems," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2730–2743, Nov. 2002.

[8] D. R. Morgan, J. Benesty, and M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Process. Lett.*, vol. 5, no. 7, pp. 174–176, Jul. 1998.