Computationally Efficient Adaptive System Identification in Subbands with Intersubband Tap Assignment for Undermodelled Problems

Stephan Weiß, Uwe Sörgel, and Robert W. Stewart

Signal Processing Division, Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow G1 1XW, Scotland, UK
weiss@spd.eee.strath.ac.uk

Abstract

This paper presents an efficient approach to adaptive system identification in undermodelled cases, ie. when the impulse response to be identified is too long to be represented by an adaptive FIR filter operating within the computational benchmarks of a DSP chip. Beyond computational savings gained from adaptive filtering in decimated subbands, the subband approach enables to use different filter lengths in the frequency bands through distribution of taps from an overall number of filter weights determined by the computational benchmark. Compared to a fullband filter of the same complexity, this method can greatly enhance the model representation of long impulse responses.

1 Introduction

When identifying very long impulse responses with an adaptive filter implemented on a DSP, as eg. found in acoustic echo control, often the benchmark limits the identification to an undermodelled problem, where IIR filters are not applicable to the nature of the problem [6]. Two possibilities to reduce the computational complexity are applying adaptive FIR filters to decimated subbands, as well as distributing the available number of computations intelligently between the subbands such that bands with longer decay or higher energy can be modelled closer than bands with few or no contribution to the overall system.

The adjustment of filter lengths differently according to a predefined scheme for a better exploitation of system resources has been reported eg. in [1], and successfully employed for commercial systems [2], based on a priori knowledge of system and signal characteristics involved in the identification process. Automatic tap-assignment between subbands was first introduced by [3], with possible refinements in [4] and [5]. However, [3],[4], and [5] work with undecimated subbands, which avoids aliasing problems in the subbands, but also counteracts the task of efficiency and low number of computations. To work with critical decimation, gap filter banks [6] or adaptive cross-terms between the subbands [7,8] have to be employed. As gap filter banks lose parts of the spectrum, we have decided on the later possiblity.

In the following, we will motivate our approach by comparing the model lengths of fullband and subband implementations based on equal complexity, and by showing how to reconstruct a fullband model from the adapted subband responses. The third section is dedicated to optimizing the tap profile, where adaptive tap-assigment strategies are derived. Their performance and characteristics will be demonstrated in section four.

2 Algorithm Complexity and Model Length

2.1 Analysis of Complexity

Fullband Adaptive Filter. From an input sequence $x(k)$, the adaptive filter forms an output $y(k)$, which subtracted from a desired signal $d(k)$ yields an error signal $e(k)$. When using the normalized least-mean square (NLMS) algorithm to minimize this error, the update equation for the adjustable filter weights, $\mathbf{w}$, given by [7]

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\tilde{\mu}}{\alpha + \mathbf{x}_k^T \mathbf{x}_k} e(k)\mathbf{x}_k, \tag{1}$$

where both weights and input signal are written in vector notation

$$\begin{aligned} \mathbf{w}_k &= [w_0(k), w_1(k), \cdots w_{N-1}(k)]^T & (2)\\ \mathbf{x}_k &= [x(k), x(k-1), \cdots, x(k-N+1)]^T & (3) \end{aligned}$$

with N being the filter length. Using the same notation, the error signal, $e(k)$, arises from the filter equations

$$\begin{aligned} y(k) &= \mathbf{x}_k^T \mathbf{w}_k & (4)\\ e(k) &= d(k) - y(k) \quad . & (5) \end{aligned}$$

The complexity for this fullband NLMS algorithms, C_{FB}, is given by

$$C_{FB} = 2N + 2, \qquad (6)$$

where complexity refers to the number of multiply-accumulate operations (MACs) per sampling period. Obviously, the finite filter length N may limit the model to a truncated estimate of the real system to be identified.

Subband Adaptive Filtering. For subband adaptive filtering, input and desired signals are split into frequency bands by an analysis filter bank, and adaptive filtering is then performed in the subbands. A global error signal can be reconstructed from the subband error signals by a synthesis bank.

Using pseudo-QMF filters for analysis and synthesis banks, the complexity of such a filter bank operation, C_B, can be kept low using an efficient polyphase implementation [11], yielding

$$C_b = 2(L_p + M^2) \cdot \frac{1}{M} \quad , \qquad (7)$$

where M is the number of subbands and L_p the length of a prototype lowpass filter, from which all filters in the bank are derived by a discrete cosine transform (DCT-IV). We derived this prototype filter, h_p from a halfband, near perfect reconstruction (PR) filter, h_h, of length L_h by iteration

$$h_p^i(k) = h_p^{i-1}(2k) * h_h(k), \qquad (8)$$

$$\text{for} \quad i = 1(1)\log_2 M, \quad h_p^0 = h_h, \qquad (9)$$

such that M is restricted to powers of 2. This methods yields

$$L_p = (2M-1)L_h - 2M + 2 \qquad (10)$$

as length for the pseudo-QMF prototype filter.

For adaptive filtering in critically decimated subbands, the adaptive filter matrix $\mathbf{W}$ is at least tridiagonal due to aliasing between the bands, making cross terms between bands necessary. [4, 5]. However, these cross terms can be restricted to adjacent subbands, if sufficiently good analysis filters are employed, and be split into a fixed and an adaptive part [5]. The complexity for the fixed part becomes

$$C_f = \frac{2M-2}{M^2}(2L_p - 1), \qquad (11)$$

and main and adaptive cross terms have the complexity

$$C_a = \frac{10M-4}{3M^2} \cdot N + \frac{6M-4}{M} \qquad (12)$$

where N is the total number of taps in the main adaptive filters. Thus, a total complexity for subband adaptive filtering is given by

$$C_{SB} = 3C_b + C_f + C_a. \qquad (13)$$

2.2 System Benchmark and Time Representation of Model

Considering an algorithm implementation, its complexity is limited by a benchmark B imposed by the resources, eg. the number of operations that can be performed on a DSP chip. On the basis of equal complexity, $C_{FB} = C_{SB} \leq B$, the number of taps in both fullband and subband implementation of an adaptive filter can be calculated, from which the represented time of the achievable model can be derived.

For the fullband implementation of the NLMS, the number of taps is given by $N_{FB} = (B-2)/2$, derived from (6) for $C_{FB} = B$. Obviously, the time representation of the achieved model is $L_{FB} = N_{FB}$ taps.

Looking at the subband approach, by inserting $C_{SB} = B$ and (12) into (13), solving for the number of taps N_{SB} yields

$$N_{SB} = \frac{3M\left(M(B - 3C_b - C_f) - 6M + 4\right)}{10M - 4}. \qquad (14)$$

Whereas the overall number of taps in the main adaptive filters is fixed, the time representation L_{SB} is flexible and depends on the tap profile, ie. how the taps are distributed over the subbands. The time representation is however limited by two extreme cases,

$$N_{SB} \leq L_{SB} \leq M N_{SB} \qquad (15)$$

where the lower bound corresponds to a uniform tap profile, ie. all subband main adaptive filters have an equal number of taps N_{SB}/M, with a time representation increase by M due to the reduced sampling rate, and the upper bound to a case, where all taps are concentrated in one single subband. Fig. 1 gives the possible range of time representation for subband adaptive filters, relative to a fullband filter of equal complexity, for different number of subbands and system benchmarks. Note that generally a higher number of subbands allows longer time representations by the model, unless a limit is reached, where most of the available computations have to be spent on the filter banks.

2.3 Reconstruction of Fullband Model

If we employ filters having PR or near PR property, we can reconstruct an overall, fullband system response in good approximation by sending an impulse through the input analysis bank, the adapted matrix $\mathbf{W}$, and reconstruct the output by the synthesis bank. We show this for the two channel QMF case $M = 2$, following the notation in [5]; thus, the reconstruction procedure can be expressed as

$$\tilde{W}(z) = [1 \;\; 1]\, \mathbf{H}_m(z) \mathbf{W}(z^2) \begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} \qquad (16)$$

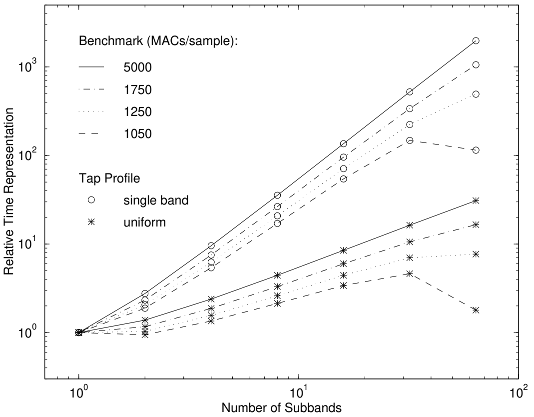

Figure 1: Upper and lower time representation bounds of a subband adaptive filter relative to a fullband implementation of the same complexity for different system benchmarks.

where $G_0(z)$ and $G_1(z)$ are the synthesis filters, $\mathbf{H}_m(z)$ the modulation expansion matrix of the analysis filters, and $\mathbf{W}(z^2)$ the optimum weight matrix [5]

$$\mathbf{W}(z^2) = \mathbf{H}_m(z)\mathbf{S}_m(z)\mathbf{H}_m^{-1}(z). \tag{17}$$

Choosing symmetrical QMF filters, with an analysis lowpass filter $H(z)$, we can solve (17) as

$$\tilde{W}(z) = \underbrace{\left(H^2(z) - H^2(-z)\right)}_{T(z)} S(z), \tag{18}$$

where $S(z)$ is the unknown system and $T(z)$ the distortion transfer function [11], covering amplitude and phase distortions in the filter banks. Thus, the overall system response can be reconstructed off-line with a lower error bound given by the distortions in the filter banks.

3 Tap Profile Adaptation

3.1 Optimum Tap Profile

Let us assume a sufficiently fine subband decomposition such that the characteristics of the signals and the unknown system are approximately constant in the mth subband. Further, let input signal and observation noise in each subband, x_m and n_m, be uncorrelated and wide sense stationary. Thus, the covariance matrix of x_m becomes $\mathbf{R}_{x_m x_m} = \mathcal{E}\{\mathbf{x}_m\mathbf{x}_m^T\} = \mathbf{I}\sigma_{x_m}^2$ and $\mathcal{E}\{n_m^2\} = \sigma_{n_m}^2$, where $\mathcal{E}\{\cdot\}$ denotes expectations.

Minimum Mean Squared Error. The minimum mean-squared error (MMSE) in the mth band is given by

$$\begin{aligned} \mathcal{E}\{e_m^2\}\,\big|_{\min} &= \mathcal{E}\{(d_m - y_m)^2\} & (19)\\ &= \mathcal{E}\Big\{\big(\mathbf{x}_m^T(\mathbf{w}_{\mathrm{opt},m} - \mathbf{w}_m) + n_m\big)^2\Big\} & (20)\end{aligned}$$

with the response of the unknown system in the mth band, $\mathbf{w}_{\mathrm{opt},m}$, and the also optimal, but after N_m taps truncated model $\mathbf{w}$. Substituting this truncation residual $\mathbf{w}_{\mathrm{opt},m} - \mathbf{w}_m$, the MMSE can be written as

$$\mathcal{E}\{e_m^2\}\,\big|_{\min} = \sum_{k=N_m}^{\infty} w_{\mathrm{opt},m}^2(k) \cdot \sigma_{x_m}^2 + \sigma_{n_m}^2 \tag{21}$$

based on the previously stated assumptions. Thus, the MMSE depends on the truncation error, weighted by the variance of the input signal and biased by the noise variance σ_{n_m}.

Global MMSE. Assume an orthonormal filter bank, wide sense stationary processes, and steady-state adaptive filters. By taking expectations of Parseval's Theorem we can express the global MSE, with the error $e(k)$ reconstructed by the synthesis bank, as a sum of the M subband MSEs:

$$\mathcal{E}\{e^2\} = \frac{1}{M}\sum_{m=1}^{M} \mathcal{E}\{e_m^2\}\,. \tag{22}$$

Here, minimizing the global MSE obviously leads to the problem of minimizing the M subband MSEs subject to the contraints

$$\mathcal{E}\{e_i^2\} \overset{!}{=} \mathcal{E}\{e_j^2\} \quad (\overset{!}{=} \min) \quad \forall i,j \in \{1, M\}, \tag{23}$$

and $\sum_{m=1}^{M} N_m = \text{const.}$, which could be solved using standard optimization techniques, eg. Lagrange.

Here, we are looking for an adaptive solution over the global MSE by controlling the filter lengths. Thus, if the subband errors converge to the subband MMSEs, and we can adjust the tap profile such that these errors are all of the same size, the global error will converge to its minimum. However, this decision may be unwillingly, eg. in the presence of observation noise due to (21), or willingly biased, eg. by introducing a weighting in (22) to account for pseudo-acoustic considerations [3, 12].

3.2 Tap-Assignment Algorithms

For adapting the tap profile of subband adaptive filters, Ma *et al.* [8] and Sugiyama *et al.* [10] provide two approaches. Both have in common to modify the tap profile every P samples, such that every filter is shortened by R taps and the pool of MR freed taps is then

S_1	θ_1/π	0.1	0.3	0.5	0.6	0.7
	ρ_1	0.9985	0.996	0.95	0.85	0.85
S_2	θ_2/π	0.3	0.48	0.5	0.6	0.9
	ρ_2	0.1	0.1	0.995	0.9	0.99

Table 1: Poles of systems S_1 and S_2.

redistributed according to some criterion,

$$N_m(k+1) = N_m(k) - R + MR\frac{c_m(k)}{\sum_{m=1}^{M} c_m(k)} \qquad (24)$$

where the quantity c_m reflects the MSE in the mth band. The last term in (24) has to appropriately quantized.

In an error-based criterion of [8], c_m is an estimate of the error variance over the past P samples [8]

$$c_m(\frac{k}{P}) = \tilde{\sigma}^2_{e_m} = \frac{1}{P}\sum_{p=0}^{P-1} e^2_m(k-p), \qquad (25)$$

We have exploited (21) to create a criterion similar to [10], where we measure the input signal power and take the power of the last S taps of an adaptive filter to estimate the truncation error

$$c_m(\frac{k}{P}) = \frac{1}{PS}\sum_{p=0}^{P-1} x^2_m(k-p)\cdot\sum_{s=0}^{S-1} w^2_{m,N_m-s}. \qquad (26)$$

This criterion introduces twofold bias in the truncation error estimate, as bands with shorter decay will be favoured, and the input signal power will deviate the result, but it suppresses the influence of observation noise. The criterion in [10] is a computationally intensive average of the right hand side of (26) over the last P iterations, while the criteria presented with (25) and (26) introduce a very low additional complexity of 1 MAC and M/P divisions, and $(1 + SP/M)$ MACs and M/P divisions, respectively. The tap profile of the cross-terms can be adapted in a similar fashion, although a required minimum number of taps for criterion (26) allows only few changes and does not give much benefit.

4 Simulations and Results

Fullband and Subband Comparison at Equal Complexity. To demonstrate the considerable gain in time representation for the identified model using a subband approach, we will compare the performance of fullband and subband adaptive filters based on equal complexity of implementations. The system to be identified, S_1, is described in Tab. 1 and Figs. 2(a) and 3(a), has an impulse response of several 1000 taps and