Sub-band Acoustic Echo Cancellation

Sven Nordholm Jörgen Nordberg Sven Nordebo

August 19, 1997

Sammanfattning

Eko-utsläckning är en viktig del i dagens moderna kommunikationssystem. Användningen av ”hands-free” utrustning i bilar, datorapplikationer, videokonferanser e.t.c, har skapat ett ökande behov av högkvalitativ akustisk eko-utsläckning. Inom dessa användnings områden är typiska längder på den akustiska kanalens impulssvar ca. 500-1500 filterkoeficienter, vid en samplingsfrekvens på 12 khz. Genom att låta eko-utsläckaren jobba med filter-strukturer som delar in signalerna i olika frekvensband fås en minskad beräkningskomplexitet och en snabbare konvergens för eko-utsläckaren. I denna studie visas att med den här typen av filter, kan en ekoundertryckning på 30 dB uppnås och en förbättrad konvergenshastighet jämfört med en traditionell ekoutsläckare. I den här rapporten är tyngdpunkten lagd på konstruktionen av filter-banken som delar up signalerna i olik delband och införandet av en enkel taldetektor som på ett märkbart sätt ökar ekoutsläckarens effektivitet.

Abstract

Echo suppression is a vital part of every communications system. The use of hands-free communication in cars, computer applications and video conferencing have created further demands for high-quality acoustic echo cancellation. In these applications the acoustic channel has, typically, a long impulse response in the order of 100ms. Typical lengths of adaptive FIR-filters are 500-1500 taps, assuming a 12 kHz sampling frequency. In order to reduce the computational load and also to improve the convergence rate, sub-band processing schemes have been suggested. This paper presents a study of a delayless sub-band adaptive filter. The study shows a possible echo suppression of about 30 dB and also an improved convergence rate when compared to a fullband LMS-filter. The main issues discussed are filter bank design and a simple speech detector that gives a drastic performance improvement.

Chapter 1

Introduction

In modern hands-free communication systems such as hands-free car phones, loudspeaker phones and video conference systems, it is necessary to perform an acoustic echo cancellation of the far-end speaker[1, 2, 3]. In order to track variations in the acoustic channel, the echo cancellation is made adaptive. The filter length of the acoustic canceller is typically 500-1500 FIR taps for normal sampling frequencies. Long filters imply a large computational burden and slow convergence rate. The slow convergence rate is especially obvious in signals with a large spectral dynamic range such as speech signals. A sub-band echo canceller [4, 5] gives several advantages when compared to a full-band echo canceller such as:

1. The computational burden is essentially reduced by the number of sub-bands due to decimation.

2. A faster convergence since the spectral dynamic range in each sub-band will be less.

3. The signal controlled adaptation can be performed in each sub-band individually, hence enhanced performance.

4. A well separated structure for parallel implementation.

This paper presents an improved version of a delayless sub-band adaptive filter (DSAF), presented by Morgan and Thi [5]. This adaptive filter structure employs the benefits of adaptive sub-band filtering, but does not suffer from the inherent delay usually found in sub-band schemes. This is due to the fact that the FIR filtering is performed without delay directly on the full-band signal. The following improvements are presented in this paper:

1. Improved filter bank design which makes it possible to improve the convergence rate.

2. A signal detection scheme operating in each sub-band, thereby improving the convergence rate for signals with a highly varying spectral content.

3. A detailed analysis of the polyphase decomposition.

The outline of the paper is as follows:

- In chapter 2 the sub-band adaptive filter is presented

- Chapter 3 shows how different prototype filters in the filter bank affect the performance of the echo-canceller.
- Chapter 4 presents improvements to the standard scheme. These improvements give better performance and convergence rates when using speech signals for identification of the acoustic channel.
- In chapter 5 the results derived from real signals (gathered in a car environment) are presented.
- Chapter 6 concludes the paper and suggests further improvements.

Chapter 2

Sub-band Adaptive Filters

An acoustic echo canceller, see Fig. 2.1, identifies the channel between the loudspeaker and the hands-free microphone. This identified impulse response is then employed to achieve an inhibition of the echo. One of the fundamental characteristics of this channel is the bulk delay. A typical distance between loudspeaker and microphone is 1 m. This corresponds to a 3 ms delay and with 8-12 kHz sample frequency this corresponds to about 20-30 samples. However, a 50 taps long FIR filter will only characterize the direct wave and give a suppression of about 5-10 dB. In order to achieve the suppression goal which is to suppress 30-40 dB, filter lengths of 500-1000 FIR taps become necessary. The filter should also be able to track variations in the acoustic environment. An appealing approach is to use a multirate technique since this technique reduces the computational burden and also gives a faster convergence rate. The latter is due to the reduction of spectral dynamic range in each sub-band. A major drawback is the delay that is introduced by the filter bank. This can, however, be circumvented by using a modified structure for the sub-band adaptive filter [5].

2.1 The Delayless Sub-band Adaptive Filter

The delayless attribute of this technique comes from the fact that the new adaptive weights are computed in sub-bands and then transformed to an equivalent full-band filter with means of an inverse FFT, see Fig. 2.1. The filter works in real time on the loudspeaker signal. The coefficients are calculated separately in each band. They can be calculated either by employing the error signal $e(k)$ (closed loop case) or the microphone input signal $d(k)$ (open loop case). If the signal $d(k)$ is used, a local error signal in each band is created and the calculations do not need to be performed in real time. This approach will, however, give less suppression since the algorithm is working blind with respect to the real error signal. The full-band signal is divided into several sub-band signals by using a polyphase FFT technique [6].

2.2 Polyphase FFT Filter Banks

A set of M filters is said to be a uniform DFT filter bank if they are related as

$$H_l(z) = H_0(zW^l) = \sum_{n=-\infty}^{\infty} h_0(n)(zW^l)^{-n}, \tag{2.1}$$

where $W = e^{-j2\pi/M}$ and $l \in [0, M-1]$. The polyphase decomposition can be used to implement such a filter bank in a very efficient manner [6]. The number of filters in

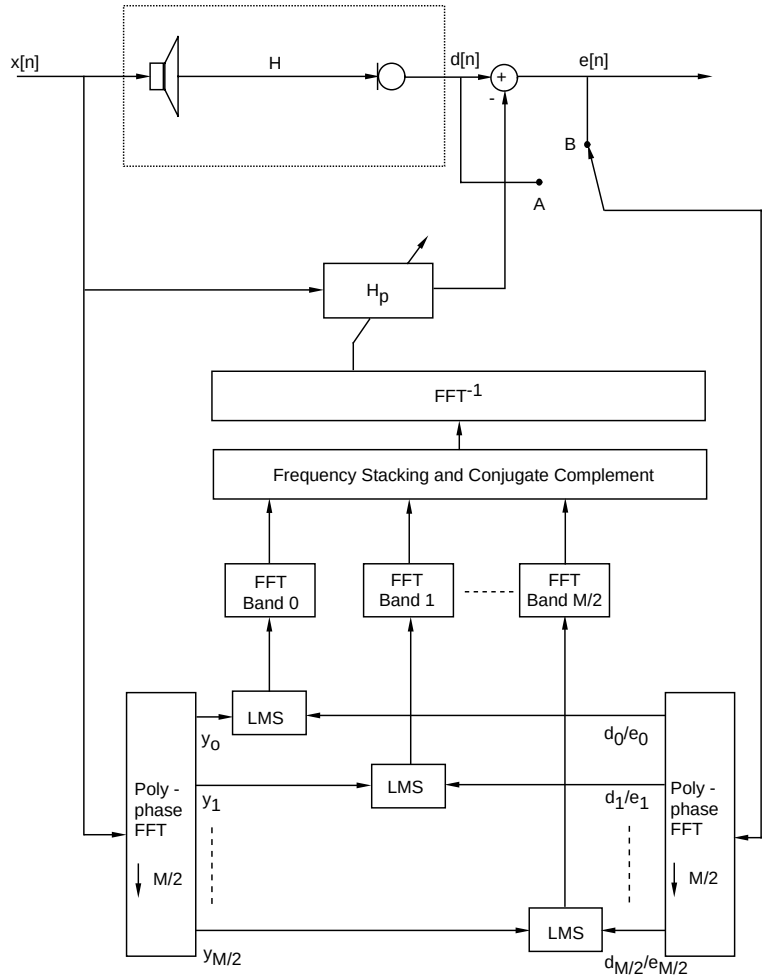

Figure 2.1: Delayless sub-band acoustic echo canceller; position A open loop configuration and position B closed loop configuration

the filter bank is M, thus the passband frequency of the prototype filter is set to $\frac{1}{2M}$. Since only full-band filters with real coefficients are considered, it is enough to calculate $\frac{M}{2}+1$ complex sub-band signals and, in order to avoid aliasing, the signals in the filter bank are decimated by a factor of only $\frac{M}{2}$. The polyphase decomposition of the DFT filter bank is performed accordingly. The resulting filter after decimation will have even sub-bands centered at dc, while odd sub-bands are centered at $\frac{1}{2}$, see Fig. 2.2. The prototype filter $H_0(z)$ is polyphase decomposed as

$$H_0(z) = \sum_{m=-\infty}^{\infty} h_0(n)z^{-n} = \sum_{m=0}^{M/2-1} z^{-m} \sum_{n=-\infty}^{\infty} h_0(n\frac{M}{2}+m)z^{-nM/2}. \tag{2.2}$$

An arbitrary filter in the filter bank Eq. (2.1) and (2.2) yields,

$$H_l(z) = \sum_{m=-\infty}^{\infty} h_0(n)(W^l z)^{-n} = \sum_{m=0}^{M/2-1} (W^l z)^{-m} \sum_{n=-\infty}^{\infty} h_0(n\frac{M}{2}+m)(W^l z)^{-nM/2}. \tag{2.3}$$

where

$$W^{-lnM/2} = (e^{j\pi l})^n = \begin{cases} (-1)^n & \text{l odd} \\ 1 & \text{l even} \end{cases} \tag{2.4}$$

Eq. (2.4) indicates that odd and even sub-bands must be treated separately, see Fig. 2.3.

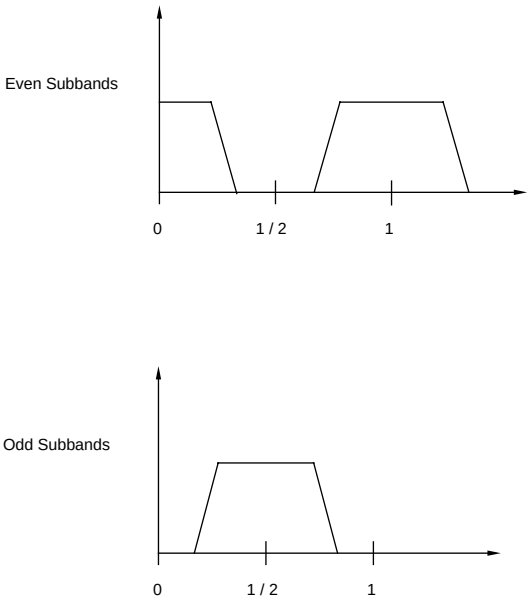

Figure 2.2: Filter-bank response for odd and even sub-bands after decimation.

For odd l Eq. (2.3) yields,

$$H_l(z) = \sum_{m=0}^{M/2-1} (W^l z)^{-m} \sum_{n=-\infty}^{\infty} h_0(n\frac{M}{2} + m)(-1)^n z^{-nM/2} \tag{2.5}$$

defining $E'_m(z)$ as

$$E'_m(z) = \sum_{n=-\infty}^{\infty} h_0(n\frac{M}{2} + m)(-1)^n z^{-n}, \tag{2.6}$$

then Eq. (2.5) can be rewritten as

$$H_l(z) = \sum_{m=0}^{M/2-1} (W^l z)^{-m} E'_m(z^{M/2}). \tag{2.7}$$

For even l Eq. (2.3) yields

$$H_l(z) = \sum_{m=0}^{M/2-1} (W^l z)^{-m} \sum_{n=-\infty}^{\infty} h_0(n\frac{M}{2} + m) z^{-nM/2} \tag{2.8}$$

defining $E_m(z)$ as

$$E_m(z) = \sum_{n=-\infty}^{\infty} h_0(n\frac{M}{2} + m) z^{-n} \tag{2.9}$$

then Eq. (2.8) can be rewritten as

$$H_l(z) = \sum_{m=0}^{M/2-1} (W^l z)^{-m} E_m(z^{M/2}) \tag{2.10}$$

This means that the polyphase filter bank is divided into two filter structures: one for even sub-bands and one for odd sub-bands, see Fig. 2.3

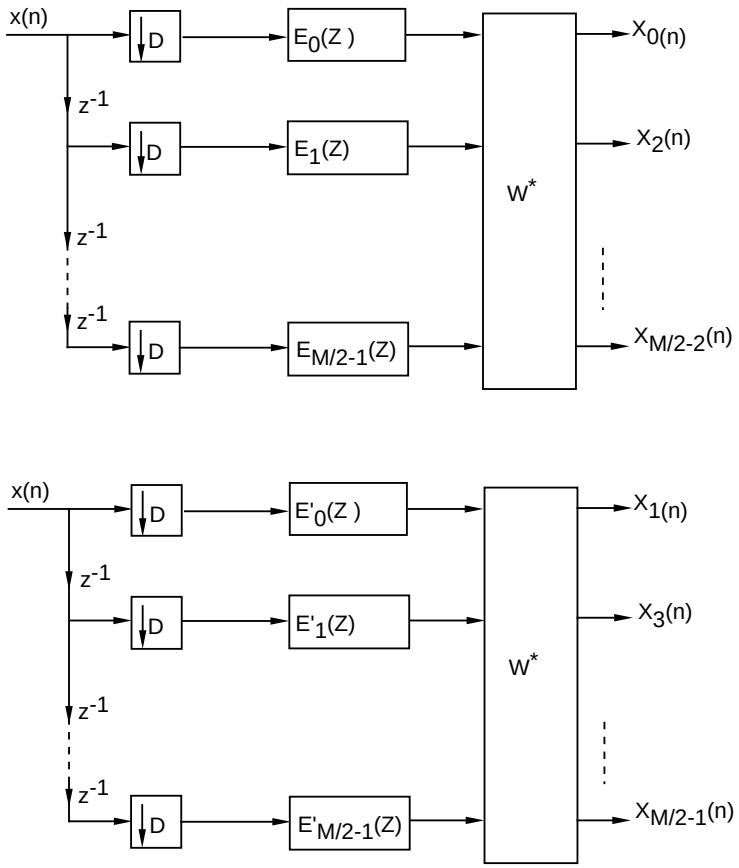

Figure 2.3: A filter bank design with polyphase FFT where even and odd sub-bands are calculated separately.

2.3 Transformation from sub-band coefficients to full-band coefficients:

If the full-band filter has N taps, the filter length in each sub-band will be $\frac{N}{D}$, $D = \frac{M}{2}$. A $\frac{N}{D}$ point FFT will be calculated on the adaptive weights in each sub-band. These are subsequently stacked to form a $[0...(\frac{N}{2}-1)]$ element array. The array is then completed by setting element $N/2$ to zero and using the complex conjugate of elements $[1...(\frac{N}{2}-1)]$ in reverse order. Finally, the N element array is transformed by a N point inverse FFT to obtain the full-band filter weights. The method is best described by an example, in this example $N = 512$ and $M = 32$. The correspondence between the FFT bins for the sub-band filter and the full-band filter are given according to table 2.4.

Subband FFT Bin Number	Wideband FFT Bin Number						
	Subband 0	Subband 1	Subband 2	. . .	Subband 14	Subband 15	Subbband 16
0	0	-	32		224	-	-
1	1	-	33		225	-	-
2	2	-	34		226	-	-
3	3	-	35		227	-	-
4	4	-	36		228	-	-
5	5	-	37		229	-	-
6	6	-	38		230	-	-
7	7	-	39		231	-	-
8	-	8	-		-	232	-
9	-	9	-		-	233	-
10	-	10	-		-	234	-
11	-	11	-		-	235	-
12	-	12	-		-	236	-
13	-	13	-		-	237	-
14	-	14	-		-	238	-
15	-	15	-		-	239	-
16	-	16	-		-	240	-
17	-	17	-		-	241	-
18	-	18	-		-	242	-
19	-	19	-		-	243	-
20	-	20	-		-	244	-
21	-	21	-		-	245	-
22	-	22	-		-	246	-
23	-	23	-		-	247	-
24	-	-	24		216	-	248
25	-	-	25		217	-	249
26	-	-	26		218	-	250
27	-	-	27		219	-	251
28	-	-	28		220	-	252
29	-	-	29		221	-	253
30	-	-	30		222	-	254
31	-	-	31		223	-	255

Figure 2.4: Frequency mapping from sub-band FFT bin numbers to wideband FFT bin numbers for a 32-sub-band polyphase FFT implementation with 512-point impulse responses and 32 taps per sub-band.

Chapter 3

The Design of Prototype Filter for Polyphase DFT Filter Bank

In this chapter a method to design general prototype filters is outlined. This method allows filter design parameters, such as magnitude response, length, phase linearity and group delay of the filter to be included in the design. These parameters will affect the convergence rate of the echo canceller as well as the suppression if not properly chosen.

3.1 Parameters and Convergence

It is well known that the convergence rate of the LMS algorithm is heavily dependent on the eigenvalue spread of the covariance matrix, such eigenvalues being related to the spectral dynamics of the input signal[8]. The DSAF employs a M-band filter bank, but the signals are only decimated with a factor $\frac{M}{2}$ which results in that the corresponding filter function in each band being as given in Fig. 2.2. The consequence of this is that the adaptive filter in each band has only an excitation signal over half the band. This will yield a slow convergence if one is interested in the adaptive filter parameters outside that band [12]. In this case, it is only necessary to obtain the adaptive filter parameters over half the band (only half of the DFT bins are used in the conversion to the full band filter). It is, however, essential that the prototype filter has low ripple in the passband.

A parameter of essential interest is the filter bank delay, as any delay will affect the speed with which the adaptive filters respond to any sudden change in the acoustic channel. This is especially important for a closed loop type of implementation (the method using the error signal $\varepsilon(k)$). A linear phase FIR filter has a group delay of $T_d = \frac{N}{2}$. A prototype filter can be designed with less group delay compared to a linear phase FIR filter, whilst retaining a similar magnitude response. This type of filter creates a system which can detect changes in the environment faster. An essential part is that the filter has a good phase linearity in the passband otherwise, the adaptive filters must compensate for the phase distortion (this argument is posed for the closed loop case). This will slow down the convergence rate.

3.2 Filter Design Procedure

In this chapter a filter design procedure will be presented which allows the designer to design filters with an arbitrary group delay using the ordinary filter types LP, BP, BS and HP. The specification in this context is such that the linear phase requirement is set

only in the pass-band . Other filter types can also be designed but these will, however, require a change in the desired filter specification, Eq. (3.4). The design filter is the result of a minimisation of the mean square error between the specified desired filter and the design filter. The result will be a compromise between the two design parameters, magnitude response and group delay. The designer can choose to emphasize either of the design parameters. This is done by employing two different weighting matrices (one for magnitude and one for group delay).

3.2.1 Mathematical Outline

The frequency function of a N-tap (causal) FIR filter is given by

$$H(\omega) = \sum_{n=0}^{N-1} h(n) e^{-j\omega n} \tag{3.1}$$

where the impulse response is assumed real. The Eq. (3.1) can also be written as

$$H(\omega) = \phi^H(\omega)\mathbf{h} \tag{3.2}$$

where

$$\phi(\omega) = \begin{pmatrix} 1 \\ e^{j\omega} \\ \vdots \\ e^{j\omega(N-1)} \end{pmatrix} \text{ and } \quad \mathbf{h} = \begin{pmatrix} h(0) \\ h(1) \\ \vdots \\ h(N-1) \end{pmatrix} \tag{3.3}$$

H_{di} is the desired complex filter specification, Eq. (3.4), where T_d is the group delay

$$H_d(\omega) = e^{-j\omega T_d} \tag{3.4}$$

Let ω_i be discrete frequency points between $[0, \pi]$ for $i \in [1, \ldots, I]$. If $\phi_i = \phi(\omega_i)$ and $\mathbf{\Phi} = [\phi_1 \ldots \phi_I]$ then Eq. (3.2) can be rewritten as

$$\mathbf{H} = \mathbf{\Phi}^H \mathbf{h} \tag{3.5}$$

Defining $\mathbf{H}_d, \mathbf{W}_m$ (the magnitude weighting matrix) and $\mathbf{W}_g$ (the group delay weighting matrix) accordingly to Eq. (3.6)

$$\mathbf{H}_d = \begin{pmatrix} H_{d1} \\ \vdots \\ H_{dI} \end{pmatrix} \quad \mathbf{W}_m = \begin{pmatrix} v_{m1}^2 & & \\ & \ddots & \\ & & v_{mI}^2 \end{pmatrix} \quad \mathbf{W}_g = \begin{pmatrix} v_{g1}^2 & & \\ & \ddots & \\ & & v_{gI}^2 \end{pmatrix} \tag{3.6}$$

where v_{mi} and v_{gi} are positive frequency weights. The derivation is split into two parts; a magnitude and a group delay solution. These two will later be combined to create a total solution.
The mean square solution to the magnitude specification is the impulse response $\mathbf{h}$, which minimises the cost function, Eq. (3.7).

$$J = \sum_{i=1}^{I} v_{mi}^2 |\mathbf{H}_{di} - \mathbf{H}(\omega_i)|^2 = (\mathbf{H}_d - \mathbf{H})^H \mathbf{W}_m (\mathbf{H}_d - \mathbf{H}) \tag{3.7}$$

Inserting Eq. (3.2) into Eq. (3.7), it can be rewritten as,

$$J = Re\{(\phi^H \mathbf{h} - \mathbf{H}_d)^H \mathbf{W}_m (\phi^H \mathbf{h} - \mathbf{H}_d)\}. \tag{3.8}$$

Eq. (3.8) yields

$$J = \mathbf{h}^T\mathbf{R}\mathbf{h} - 2\mathbf{h}^T\mathbf{P} + \mathbf{H}_d^H\mathbf{W}_m\mathbf{H}_d, \tag{3.9}$$

where

$$\mathbf{R} = Re\{\mathbf{\Phi}\mathbf{W}_m\mathbf{\Phi}^H\} \text{ and } \quad \mathbf{P} = Re\{\mathbf{\Phi}\mathbf{W}_m\mathbf{H}_d\}. \tag{3.10}$$

By inserting Eq. 3.10 into Eq. 3.8 and using completion of squares that equation,

$$J = (\mathbf{h} - \mathbf{R}^{-1}\mathbf{P})^T\mathbf{R}(\mathbf{h} - \mathbf{R}^{-1}\mathbf{P}) - \mathbf{P}\mathbf{R}^{-1}\mathbf{P} + \mathbf{H}_d^H\mathbf{W}_m\mathbf{H}_d. \tag{3.11}$$

The fact that $\mathbf{R}$ is a positive definite matrix yields the solution,

$$\mathbf{h}_* = \mathbf{R}^{-1}\mathbf{P} \tag{3.12}$$

where $\mathbf{h}_*$ is the impulse response that minimized the cost function J.
An approximate expression for the group delay error in the passband was derived in [7],

$$e_\tau(\omega) \approx \sum_{n=0}^{N-1}(n - T_d)h(n)cos(\omega(n - T_d)), \tag{3.13}$$

where τ_d is the desired group delay. Eq. 3.13 can be rewritten as

$$e_\tau(\omega) = \psi^T(\omega)\mathbf{h}, \tag{3.14}$$

where

$$\psi(\omega) = \begin{pmatrix} (T_d)\cos(\omega(T_d)) \\ (1 - T_d)\cos(\omega(1 - T_d)) \\ \vdots \\ (N - 1 - T_d)\cos(\omega(N - 1 - T_d)) \end{pmatrix} \tag{3.15}$$

Let ω_k be discrete frequency points between $[0, \pi]$ for $k \in [1, \ldots, K]$. If $\psi_k = \psi(\omega_k)$ and $\mathbf{\Psi} = [\psi_1 \ldots \psi_K]$, a cost function that can be used to minimise the group delay error can be expressed as follows

$$J_{T_d} = \mathbf{h}^T\mathbf{S}\mathbf{h}, \tag{3.16}$$

where

$$\mathbf{S} = \mathbf{\Psi}\mathbf{W}_g\mathbf{\Psi}^T. \tag{3.17}$$

Eq. (3.9) and Eq. (3.17) are combined to form a common cost function. By optimising this cost function $\mathbf{h}_{opt}$ is obtained. This solution satisfies both the magnitude and the group delay cost functions,

$$J_{tot} = (\mathbf{h} - \mathbf{R}^{-1}\mathbf{P})^T\mathbf{R}(\mathbf{h} - \mathbf{R}^{-1}\mathbf{P}) - \mathbf{P}\mathbf{R}^{-1}\mathbf{P} + \mathbf{H}_d^H\mathbf{W}_m\mathbf{H}_d + \mathbf{h}^T\mathbf{S}\mathbf{h}. \tag{3.18}$$

Using completion of squares in Eq. (3.18) yields,

$$J_{tot} = (\mathbf{h}-(\mathbf{R}+\mathbf{S})^{-1}\mathbf{P})^T(\mathbf{R}+\mathbf{S})(\mathbf{h}-(\mathbf{R}+\mathbf{S})^{-1}\mathbf{P})-\mathbf{P}(\mathbf{R}+\mathbf{S})^{-1}\mathbf{P}+\mathbf{H}_d^H\mathbf{W}_m\mathbf{H}_d \tag{3.19}$$

and the optimal solution is given as

$$\mathbf{h}_{opt} = (\mathbf{R} + \mathbf{S})^{-1}\mathbf{P} \tag{3.20}$$

Utilizing this solution a filter with a specified group delay can be obtained. The filter designer only needs to determine a magnitude response ($\mathbf{H}_d$), a filter length (I) and a group delay (T_d) and then calculate $\mathbf{h}_{opt}$ from Eq. (3.20).

Chapter 4

Echo Cancellation Employing Speech Signals

The echo cancellation problem is an identification problem and as such it is necessary to have a persistingly exiting [9] input signal. A straight forward approach is to use self generated noise as the input signal. This will give a very efficient scheme, but it is desired to be able to identify directly on the speech signal. The speech signal is a non-stationary signal which can be characterized as short time stationary. The typical length of a short time frame is 20 ms. The speech signal[11] is a combination of unvoiced noise like sounds(broad band), for example the S sound; voiced quasi-periodic sounds(narrow band), for example the vowel I sound, and long periods of silence.

The acoustic echo canceller must be adaptive. Therefore, an echo canceller that can adapt effectively to the speech signal itself must be developed. When the input signal is a speech signal there will be frequency bands that are not excited, thus using a NLMS algorithm directly will give a slow convergence. Some sort of speech detector needs to be developed.

4.1 Speech Detector

We have found that a speech detector that works as a simple signal energy detector in each sub-band with a threshold will work effectively. This simple detector controls the adaptive algorithm individually in each sub-band, see Fig. 4.1, and will enhance the convergence rate. The weights will only be updated in frequency bands where the signal energy is above the threshold, whilst in the others the old weights will remain. The approach is especially good when the signal energy is concentrated in one or a few frequency bands or is of highly varying spectral content such as a speech signal. When some of the sub-bands contain little signal energy, their contribution to the over all solution will be of negative nature because the NLMS algorithm has too noisy gradient estimate. Noisy gradient estimates or poor signal excitation in certain frequency bands are especially crucial in fixed point implementations [10]. This type of adaptive system will therefore utilize the speech signal much better than a full band echo suppressor or a system with a full band speech detector. The used signal detector is very rudimentary but effective, it works in each sub-band individually and takes the absolute value of the "x" signal used in the NLMS and averages it. The obtained value is compared to a threshold. If the value is below a threshold the adaptation in that sub-band is switched off, see Fig. 4.2. This simple signal detector worked very well and made it possible to achieve good suppression results (over 20 dB) on speech signals.

In the start up phase it is advantageous to use a broad band self generated noise sequence to get a good initial channel estimate.

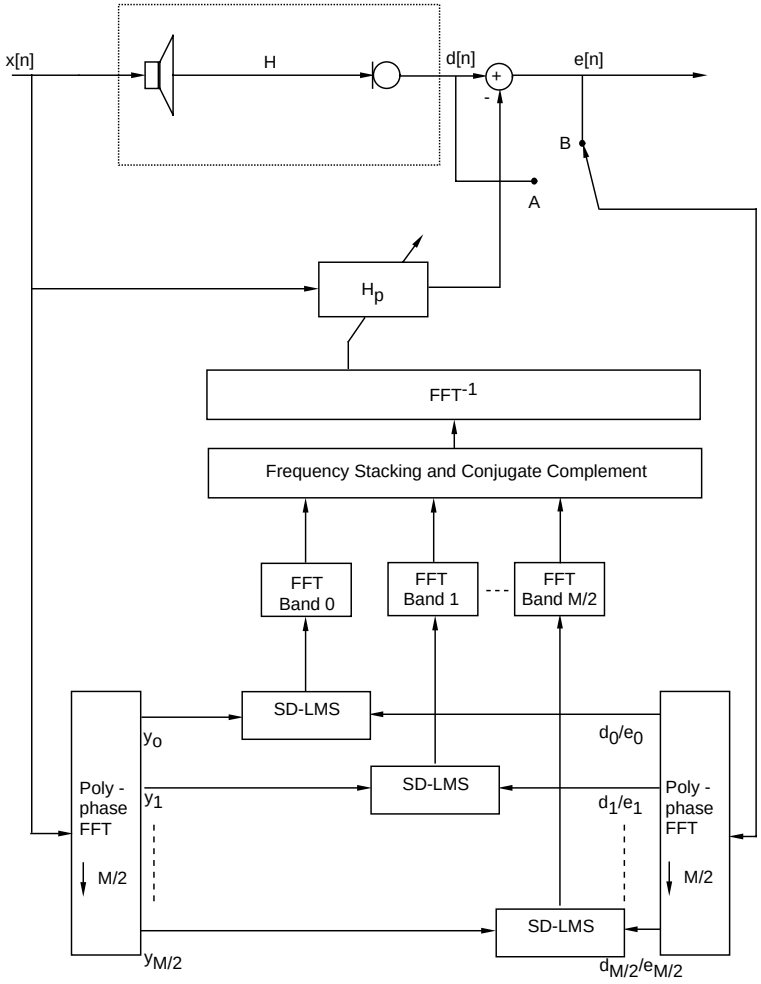

Figure 4.1: Delayless sub-band acoustic echo canceller with a signal detector in each sub-band.

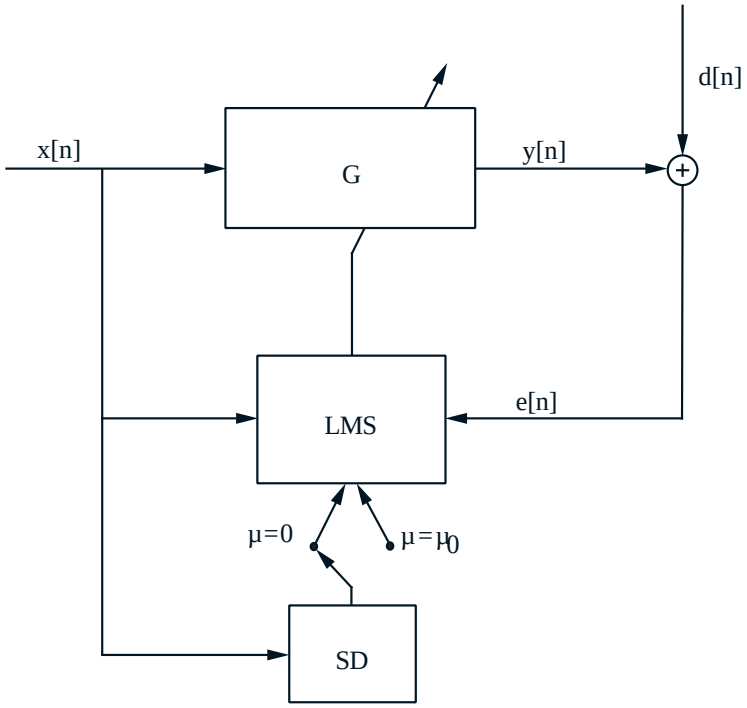

Figure 4.2: The SDLMS configuration for sub-band j in the open loop case

Chapter 5

Simulation Examples

In this section is results from the study of the acoustic echo canceller are presented. In this study the acoustic signals gathered in a car as well as computer simulated signals have been employed. The adaptations have been evaluated by using bandlimited flat noise and speech signals. The simulations are performed using a 512 tap full-band filter and a 16 sub-band echo canceller.

The SuPpression Ratio, SPR is defined as:

$$SPR = 10 \log \left(\frac{\sum_{n=1}^{N} (x[n])^2}{\sum_{n=1}^{N} (e[n])^2} \right). \tag{5.1}$$

$N = 400$ corresponding to 33 ms. We have assumed no near-end speech.

5.1 Fullband v.s Sub-band

The convergence rate for the DSAF(open loop type scheme) shows an improvement when compared to the full-band filter, see Fig. 5.1. The filter weight error norm $||\mathbf{h} - \mathbf{w}||$ is used, where $\mathbf{h}$ is the true impulse response of a known channel and $\mathbf{w}$ is the adaptive FIR filter weights. In this comparison, noise with a frequency content of between $[80, 5000]$ Hz been used as input signal, which means that all bands with the exception of the uppermost sub-band have been excited . The open loop type of DSAF has a larger distance to the true weights than the full band scheme. This is due to the fact that it is blind to the actual error. The error signals are created in each sub-band and they will overlap spectrally between the bands and thus they will not give the same solution. This can be improved by using different choices of filter banks. In this paper we have not made such an evaluation. In a practical situation the echo suppression will be the same but the convergence rate will be faster. An other advantage with the DSAF is that the number of floating point operations are reduced by approximately 30% in the studied case.

5.2 Closed Loop v.s Open Loop

In this section suppression results for the open loop DSAF and the closed loop DSAF are compared. The first results presented are those which reflect the difference between the open loop and closed loop DSAF, see Fig. 5.2. In order to clearly show these differences,

a measured room acoustic impulse response was used in combination with self generated noise. This means that the input signal and output signals are perfectly known and the output signal is a linearly filtered version of the input signal. We therefore have a coherence factor of 1 for every frequency. Ideally the output signal can be totally canceled. The closed loop algorithm resulted in a 70 dB suppression and the open loop in a 40 dB suppression. For a real situation, i.e., signals which have been gathered in a car before the loudspeaker and after the hands-free microphone, there will be less difference between the two methods, see Fig 5.3. In this situation there will be noise in the microphone and nonlinearities in the amplifiers and loudspeakers that will limit the possible suppression level. This is advantageous to the open loop DSAF. Both methods achieve a 30 dB echo suppression. The open loop method will converge faster than the closed loop method, which will give slightly better suppression once it has converged.

5.3 Closed Loop DSAF Study

In the sequel, the experiments concentrate on the closed loop DSAF configuration. This type of scheme has shown better suppression. The convergence rate is, however, initially a bit slower than for the open loop. This is due to the extra delay in the filter bank that is apparent in the error signal(for the open loop situation the error is formed locally in each sub-band). In this section some results from experiments using alternative prototype filters in the filter bank are presented. Results are also presented for the modified scheme which operates directly on the speech signal.

5.3.1 The Filter Bank's Importance in the Overall Performance of the DSAF Scheme

The main objectives in the prototype filter design are:

- a low passband ripple
- a short group delay
- a stopband suppression that is sufficient to avoid aliasing
- a linear phase in passband

Fig. 5.5 - 5.8 show the importance of choosing a prototype filter with an appropriate group delay. Fig. 5.5 shows the difference in convergence rates for prototype filters with the same amplitude function, see Fig. 5.4, but different group delays (self generated noise is filtered through a measured channel). Fig. 5.6 shows the same as in Fig. 5.5, but for real signals. When a prototype filter with a reduced group delay is used, the canceller responds more quickly to sudden changes in the channel H. This fact is shown in Fig. 5.7 -5.8, both for self generated signals and real signals.
The importance of having a linear phase in the passband in the prototype filter is shown in Fig. 5.10 - 5.11. The nonlinear prototype filter used in this comparison between the linear and nonlinear phase is a minimum phase version of the ordinary linear prototype filter. A minimum phase filter is constructed in the following way:

1. Design a FIR filter.
2. Determine the zero-plot in the z-domain.

3. Take those zero-passages ,Z_k, that are outside the unit circle and mirror them inside by the following transformation, $Z'_k = \frac{1}{Z_k}$.

The new nonlinear filter will have the same magnitude response as the linear FIR-filter, see Fig. 5.9, but it will have all the "zeros" inside the unit circle.

5.3.2 Results from employing the DSAF scheme directly on speech signals

In order to obtain a good suppression result (20 - 35 dB) using speech signals, the echo canceller has been fitted with a signal detector in each sub-band. Fig. 5.12 shows that when the signal detector is included, the filter coefficients converges faster and the resulting echo suppression will be increased. When the sub-detector is switched off because of a low energy level, the adaption is frozen and the last set of established coefficients is used. The next time there is a signal in that/those sub-bands, a good start value is at hand, see Fig. 5.13. This learning feature makes it possible to use a self-generated training sequence to create good start estimate of the channel. Figure 5.14 describes the resulting suppression when the echo canceller has been trained with a sequence of white noise, and thus given good values to work from.

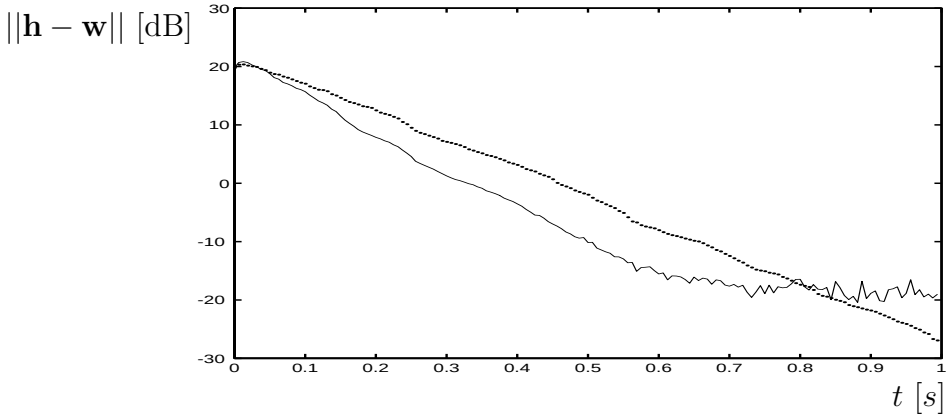

Figure 5.1: Learning curves for a full-band normalized LMS(dashed line) and a delayless adaptive acoustic echo-canceller(solid line).

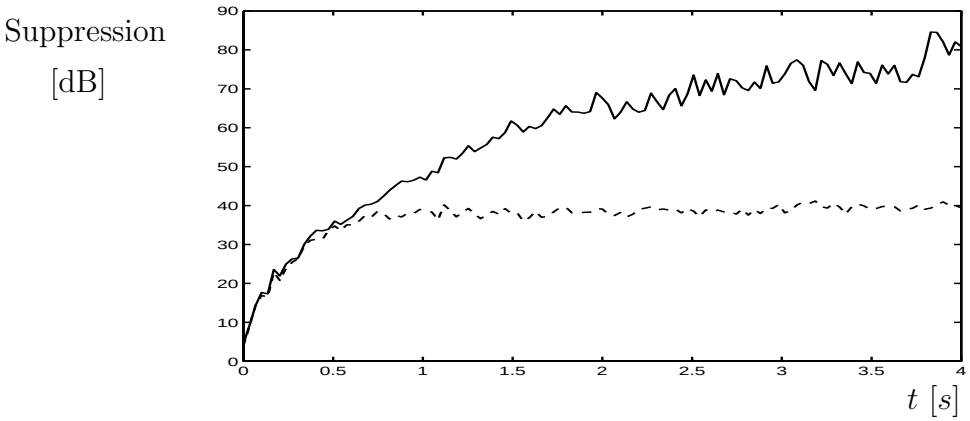

Figure 5.2: Computer simulations using a measured room acoustic impulse response resulted in the following suppression results: the dashed line shows the open loop case and solid line shows the closed loop case.

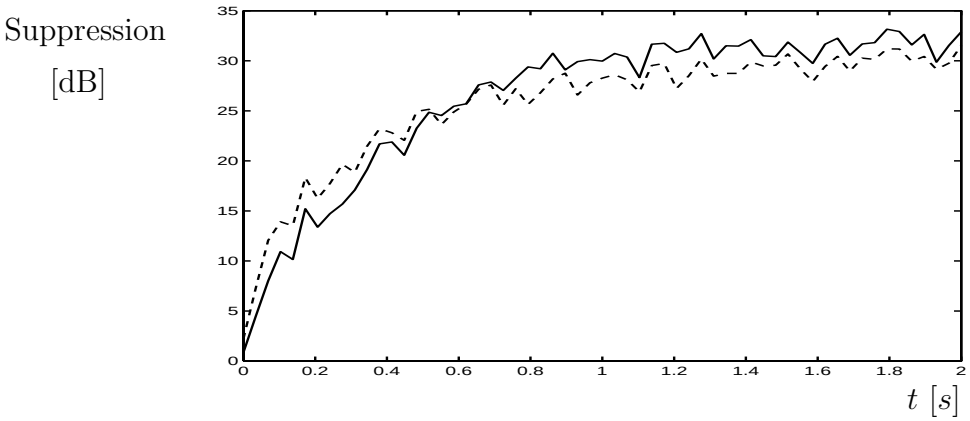

Figure 5.3: The suppression results in a simulation with "real signals": the dashed line shows the open loop case and solid line shows the closed loop case.

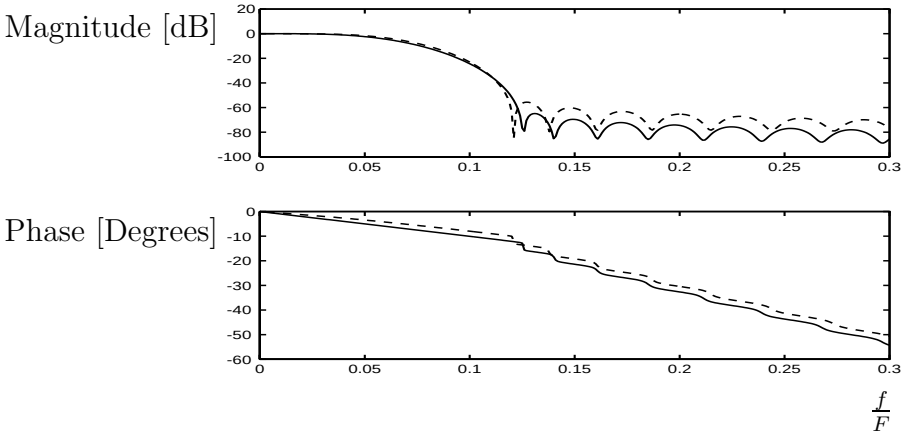

Figure 5.4: The magnitude and phase response for two 64 tap prototype filters, one with a group delay at 32 samples (solid line) and the other one with a groupdelay at 20 samples (dashed line).

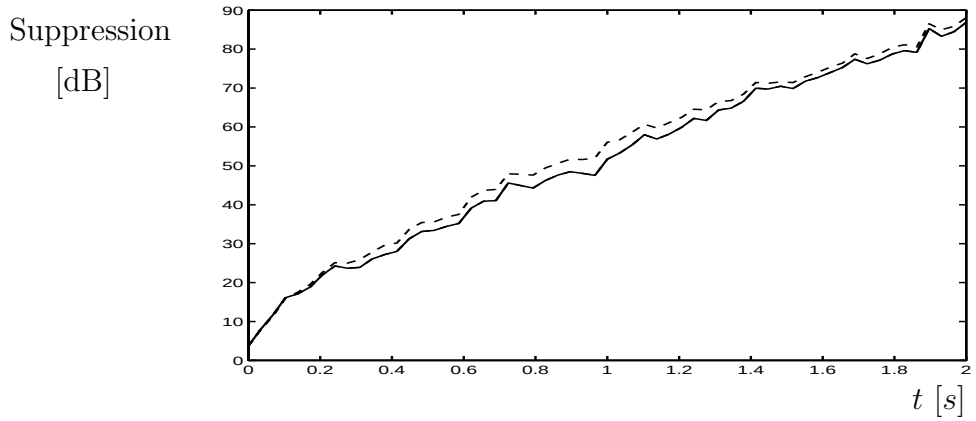

Figure 5.5: Computer simulations with a measured room acoustic impulse response resulted in the following suppression results for the closed loop case: the dashed line shows the result with a group delay at 20 samples and solid line shows the result with a group delay at 32 samples. In the simulations a 64 tap prototype filter was used.

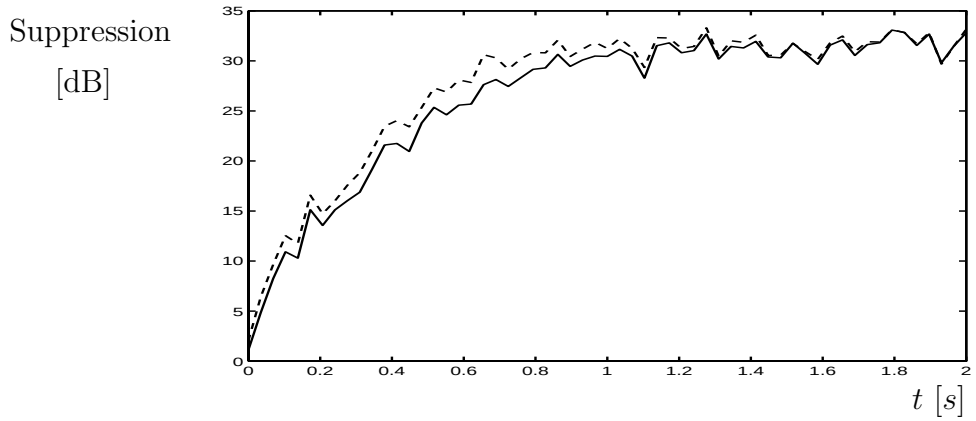

Figure 5.6: Computer simulations on real signals resulted in the following suppression results for the closed loop case: the dashed line shows the result with a group delay at 20 samples and solid line shows the result with a group delay at 32 samples. In the simulations a 64 tap prototype filter was used.

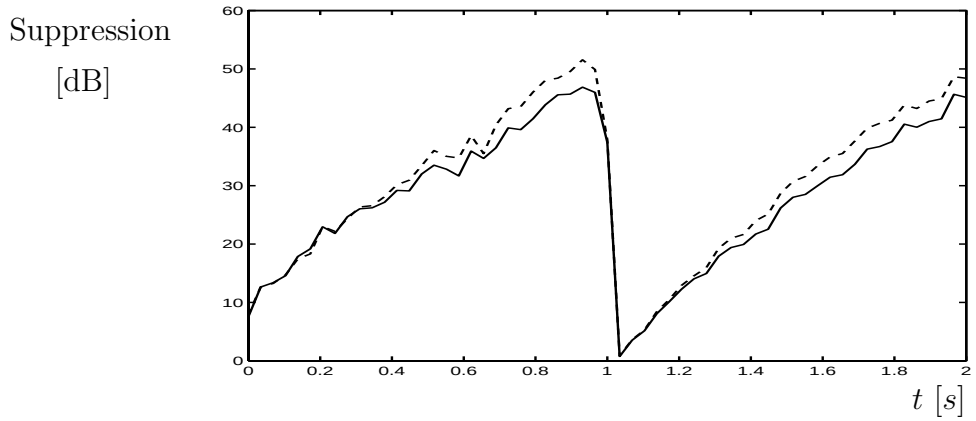

Figure 5.7: Computer simulations with a measured room acoustic impulse response when a sudden change occurs in the channel, resulted in the following suppression results for the closed loop case: the dashed line shows the result with a group delay at 20 samples and solid line shows the result with a group delay at 32 samples.In the simulations a 64 tap prototype filter and a self generated bandlimited flat noise were used.

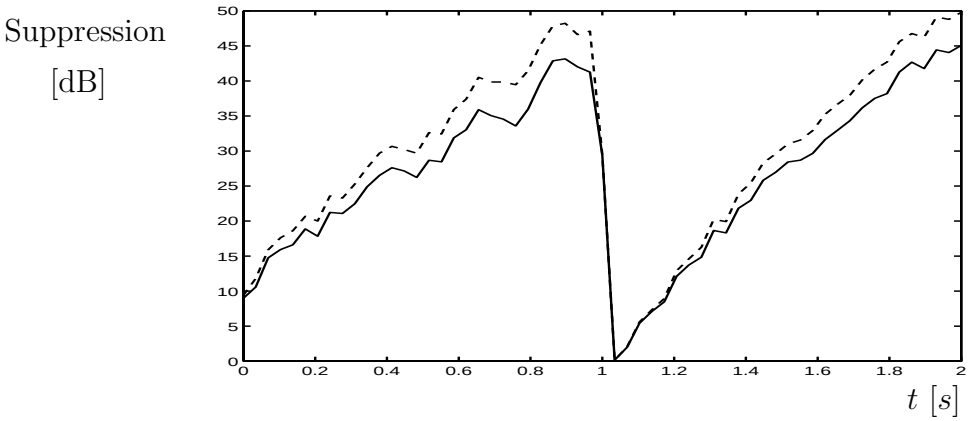

Figure 5.8: Computer simulations with a measured room acoustic impulse response when a sudden change occurs in the channel, resulted in the following suppression results for the closed loop case: the dashed line shows the result with a group delay at 20 samples and the solid line shows the result with a group delay at 32 samples. In the simulations a 64 tap prototype filter and acoustic bandlimited flat noise were used.

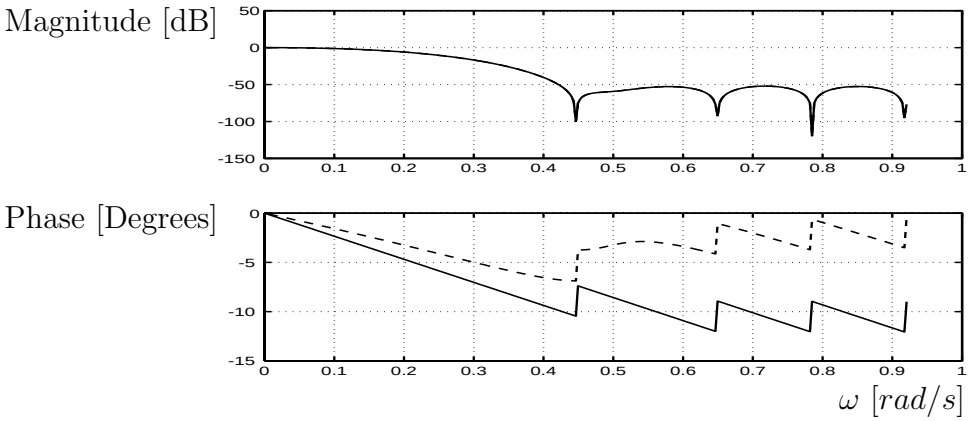

Figure 5.9: The magnitude and phase response for two 48 tap prototype filters; one with a nonlinear (min-phase) phase (dashed line) and the other one with a linear phase (solid line).

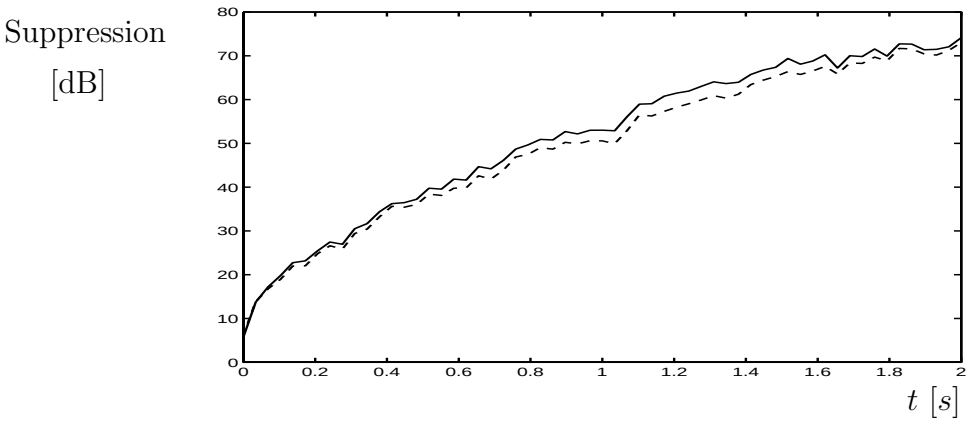

Figure 5.10: Computer simulations with a Chebychev filter as a channel yield the following suppression results for the closed loop case: the dashed line shows the result with a nonlinear phase prototype filter and the solid line shows the result using a linear phase prototype filter. In the simulations a 48 tap prototype filter and self generated bandlimited flat noise were used.

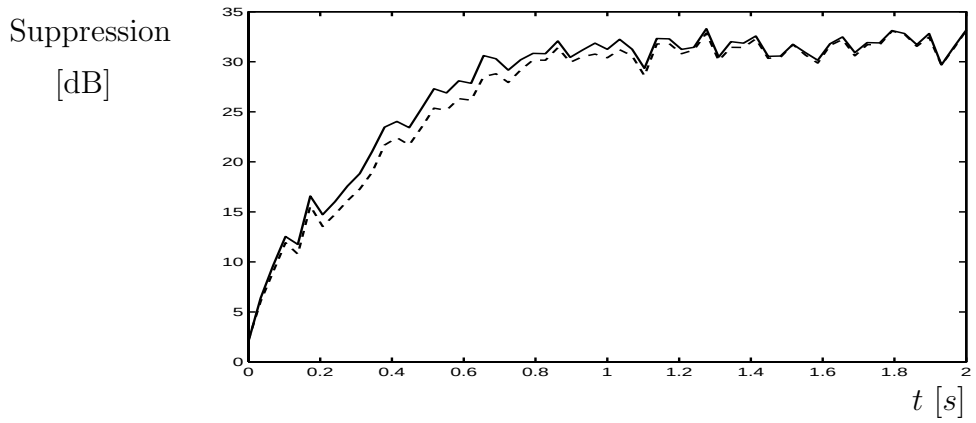

Figure 5.11: Computer simulations on real signals resulted in the following suppression results for the closed loop case: the dashed line shows the result with a nonlinear phase prototype filter, and the solid line shows the result with a linear phase prototype filter. In the simulations a 64 tap prototype filter was used.

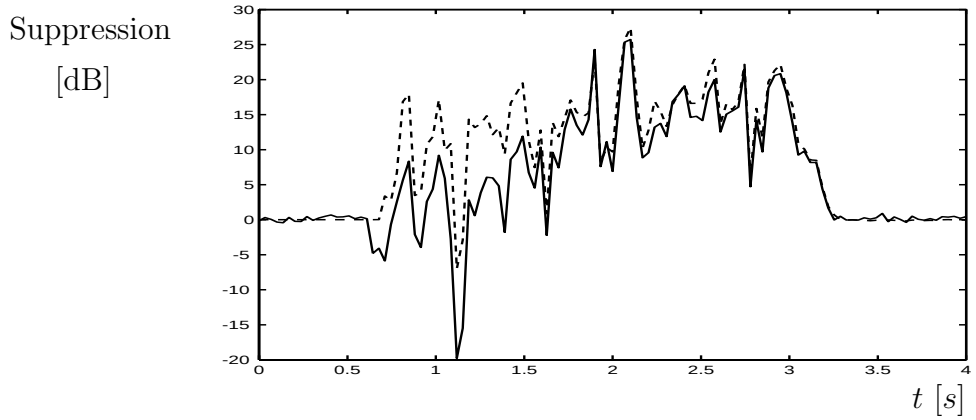

Figure 5.12: Computer simulations on real signals resulted in the following suppression results for the closed loop case: the dashed line shows the result when signal detectors were used, and the solid line shows the result without signal detectors.

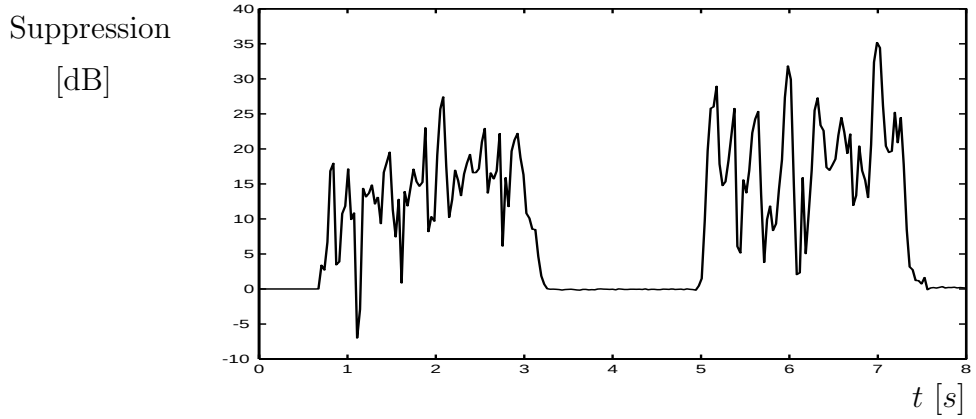

Figure 5.13: Computer simulations on real signals resulted in the following suppression results for the closed loop case. The result depicts the learning feature that signal detector introduced. The echo canceller has much better performance on the second speech sequence than the first (these are two different sequences) .

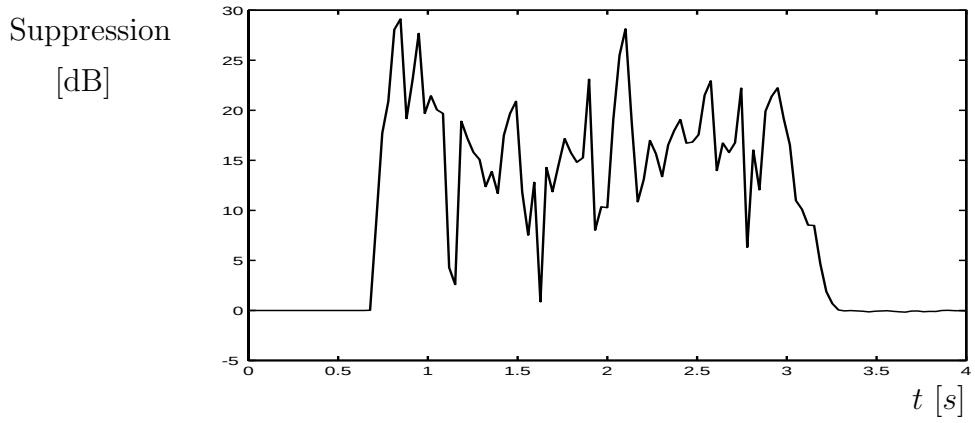

Figure 5.14: Computer simulations on real signals resulted in the following suppression results for the closed loop case. In this simulation a calibrated echo canceller was used, which should be compared with the result in Fig. 5.12. Bandlimited flat noise was used to calibrate the echo canceller.

Chapter 6

Summary and Conclusions

In this paper a sub-band adaptive filter scheme has been studied. It is called the Delayless Sub-band Adaptive Filter (DSAF) scheme, its main feature is that the filtering is performed directly on the full band signal, thereby avoiding delays associated with sub-band filtering and the adaptive process is separated from the filtering operation. The scheme has been used as an adaptive echo canceller and showed good results. The scheme has faster convergence and demands fewer operations compared to a full band scheme. The original scheme[5] has been improved to get faster convergence and also enhanced for adaptation directly on speech signals. The scheme shows good suppression results, up to 20-35 dB suppression on a speech signal. A method for designing filters with very general specifications has been presented. Filters designed with this design method have improved the convergence rate for the acoustic echo canceller. The biggest improvement has been achieved through using a simple speech-detector in each sub-band.

Bibliography

[1] B. Widrow, S. D. Stearns
Adaptive Signal Processing
Prentice Hall, 1985

[2] M. M. Sondhi, W. Kellermann
"Adaptive Echo Cancellation for Speech Signals"
in *Advances in Speech Signal Processing*, New York: Marcel Decker, 1992 , ch 11

[3] D. R. Morgan
"Slow Asymptotic Convergence of LMS Acoustic Echo Cancelers"
IEEE Trans. on Speech and Audio Processing, vol. 3, no. 2., pp. 126-136, March 1995

[4] Y. Ono, H. Kiya
"Performance Analysis of Sub-band Adaptive Systems using an Equivalent Model"
IEEE Proc ICASSP'94(Adelade, Australia), part III, pp. 53-56, 1994

[5] D. R. Morgan, J. C. Thi
"A Delayless Sub-band Adaptive Filter Architecture"
IEEE Trans. on Signal Processing, vol. 43, no. 8., pp. 1819-1830, Aug 1995

[6] P.P. Vaidyanathan
Multirate Systems and Filter Banks
Prentice Hall, 1993

[7] D. R. Xiangkun Chen, Thomas W. Parks
"Design of FIR Filters in the Complex Domain"
in *IEEE Trans. on Acoustic,Speech and Signal Processing*, vol. ASSP-35, no. 2, Feb 1987

[8] R.M. Gray
"On the Asymptotic Eigenvalue Distribution of Toeplitz Matrices"
in *IEEE Trans. on Information Theory*, vol. IT-16, p.p. 725-730, 1972

[9] T. Söderström, P. Stoica
System Identification
Prentice Hall International, 1989

[10] R. D. Gitlin, H. C. Meadors, S. B. Weinstein
"The tap-leakage algorithm: An algorithm for stable operation of digitally implemented, fractionally spaced adaptive equalizer"
in *bell Syst. Tech. J.*, vol. 61, no. 8, oct. 1982

[11] J. R. Deller, J. G. Proakis, J. H. L. Hansen
Discrete-Time Processing of Speech Signal
Macmillan, 1993

[12] I. Claesson, S. Nordholm, P. Eriksson
"Noise Cancelleing Convergence Rates for the LMS Algorithm"
in *Mechanical Systems and Signal Processing*, vol. 5, p.p. 375-388, 1991