FIFTH INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
DECEMBER 15-18, 1997
ADELAIDE, SOUTH AUSTRALIA

# DELAYLESS SUBBAND ECHO CANCELLATION

Sven Nordholm Jörgen Nordberg
Dept. of Signal Processing, University of Karlskrona/Ronneby, Sweden

August 6, 1997

**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.

## 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, the computational burden is essentially reduced by the number of sub-bands due to decimation, a faster convergence since the spectral dynamic range in each sub-band will be less, the signal energy adaptation control can be performed in each sub-band individually(thus enhancing the performance) and a well separated structure suitable 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, improved filter bank design which makes it possible to enhance the convergence rate and a signal detection scheme operating in each sub-band, thereby improving the convergence rate for signals with a highly varying spectral content.

# 2 Sub-band Adaptive Filters

An acoustic echo canceller, see Fig. 1, identifies the channel between the loudspeaker and the hands-free microphone and electrically subtracts the echo. One of the fundamental characteristics of this channel is the the long reverberation time. Thus filter lengths of 500-1000 FIR taps become necessary in order to get a sufficient suppression, about 30-40 dB. 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], a 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. 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.

The full-band filter with real coefficients has $N$ taps, thus 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. A detailed description of the procedure can be found in [5].

# 3 The Design of Prototype Filter for Polyphase DFT Filter Bank

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. A parameter of essential interest is the group delay, as the delay will affect the speed with which the

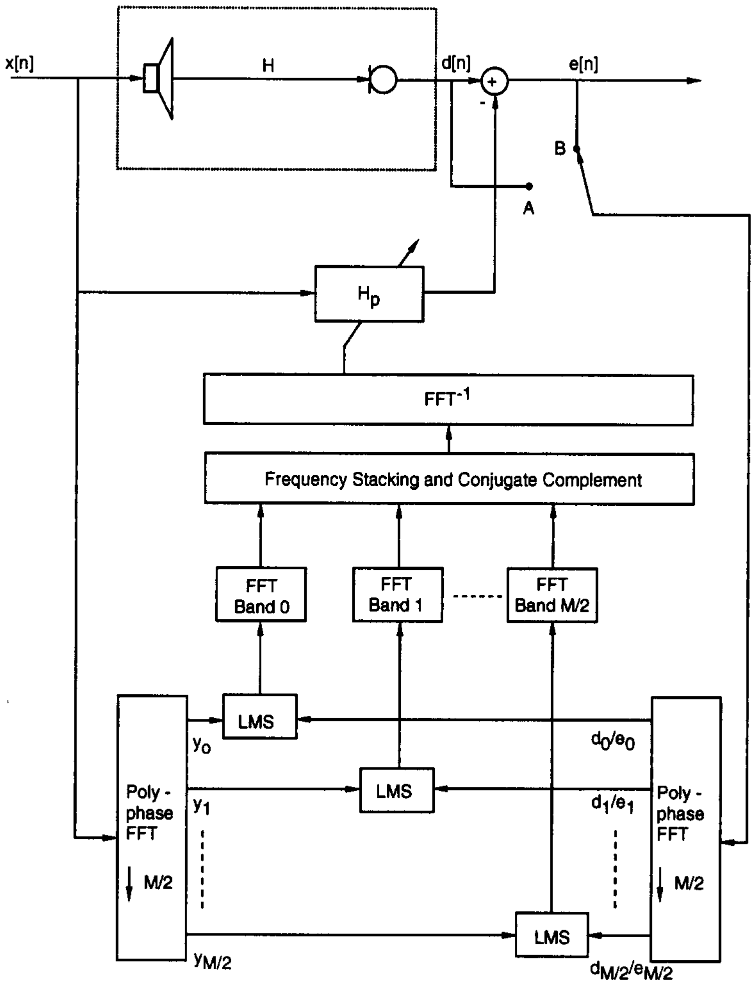


Figure 1: Delayless sub-band acoustic echo canceller; position A open loop configuration and position B closed loop configuration

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 $\varepsilon(k)$). The specification in this context is such that the linear phase requirement is set only in the pass-band . 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).

## 3.1 Mathematical Outline

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

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

where the impulse response is assumed real. This equation can alternatively be written on matrix form

$$H(\omega) = \phi^H(\omega)\mathbf{h} \tag{2}$$

If $\phi_i = \phi(\omega_i)$ and $\mathbf{\Phi} = [\phi_1 \dots \phi_I]$ then Eq. (2) can be rewritten as

$$\mathbf{H} = \mathbf{\Phi}^H\mathbf{h} \tag{3}$$

$\mathbf{H}$, is a vector containing frequency functions over a grid.

The mean square solution to the magnitude specification is the impulse response $\mathbf{h}$, which minimises the cost function,

$$J = \sum_{i=1}^{I} v_{mi}^2 |\mathbf{H}_{di} - \mathbf{H}(\omega_i)|^2 = (\mathbf{H}_d - \mathbf{H})^H \mathbf{W}_m (\mathbf{H}_d - \mathbf{H}). \tag{4}$$

where $\mathbf{H}(\omega_i)$ is a vector containing the frequency function in $I$ frequency points and $\mathbf{H}_{di}$ is the desired complex filter specification. The last expression is obtained by gathering the vectors into matrix form, where $\mathbf{W}_m$ is a weighting matrix.

This equation can be reformulated yielding a solution:

$$\mathbf{h}_* = \mathbf{R}^{-1}\mathbf{P} \tag{5}$$

where $\mathbf{h}_*$ is the impulse response that minimized the cost function J, where

$$\mathbf{R} = Re\{\mathbf{\Phi}\mathbf{W}_m\mathbf{\Phi}^H\} \text{ and } \mathbf{P} = Re\{\mathbf{\Phi}\mathbf{W}_m\mathbf{H}_d\}. \tag{6}$$

An approximate expression for the group delay error in the passband was derived in [6],

$$e_\tau(\omega) \approx \sum_{n=0}^{N-1} (n - T_d)h(n)cos(\omega(n - T_d)), \tag{7}$$

where $\tau_d$ is the desired group delay. Rewritten on matrix form

$$e_\tau(\omega) = \psi^T(\omega)\mathbf{h}. \tag{8}$$

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

$$J_{T_d} = \mathbf{h}^T\mathbf{S}\mathbf{h}, \tag{9}$$

where

$$\mathbf{S} = \mathbf{\Psi}\mathbf{W}_g\mathbf{\Psi}^T. \tag{10}$$

Eq. (4) and Eq. (10) are combined to form a common cost function, using Eq. (6). By optimising this cost function $\mathbf{h}_{opt}$ is obtained. This solution satisfies both the magnitude and the group delay cost functions,

$$J_{tot} = (\mathbf{h} - \mathbf{R}^{-1}\mathbf{P})^T\mathbf{R}(\mathbf{h} - \mathbf{R}^{-1}\mathbf{P}) - \mathbf{P}\mathbf{R}^{-1}\mathbf{P} + \mathbf{H}_d^H\mathbf{W}_m\mathbf{H}_d + \mathbf{h}^T\mathbf{S}\mathbf{h}. \tag{11}$$

yielding an optimal solution given as

$$\mathbf{h}_{opt} = (\mathbf{R} + \mathbf{S})^{-1}\mathbf{P} \tag{12}$$

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. (12).

# 4 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/or speech signals. The simulations are performed using a 512 tap full-band filter and a 16 sub-band delayless echo canceller.

The SuPpression Ratio, $SPR$ is defined as:

$$SPR = 10 \log \left( \frac{\sum_{n=1}^{N} (x[n])^2}{\sum_{n=1}^{N} (e[n])^2} \right). \tag{13}$$

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

## 4.1 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. 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. 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 2. In this situation there will be noise in the microphone and nonlinearities in the amplifiers and loudspeakers that will limit the possible suppression level. The open loop method will converge faster than the closed loop method, which will give slightly better suppression once it has converged.

## 4.2 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.

### 4.2.1 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. 3 shows the importance of choosing a prototype filter with an appropriate group delay. Fig. 3 shows the difference in convergence rates for prototype filters with the same amplitude function but different group delays

The importance of having a linear phase in the passband in the prototype filter is shown in Fig. 4. 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.

### 4.2.2 Results from employing the DSAF scheme directly on speech signals

A straight forward approach is to use self generated noise as an input signal. This will give a very efficient scheme, but it is desired to employ the speech signal itself for the identification. The speech signal is a non-stationary signal which varies much in energy as well as frequency content.

In order to obtain a good suppression result(20 - 35 dB) using speech signals, the echo canceller has been extended with a signal detector in each sub-band. 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, 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. Fig. 5 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.

## 5 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 20-25 dB suppression adapting using 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 largest improvement has been achieved through using a simple speech-detector in each sub-band.

## References


[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] 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

[7] 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

[8] 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

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

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

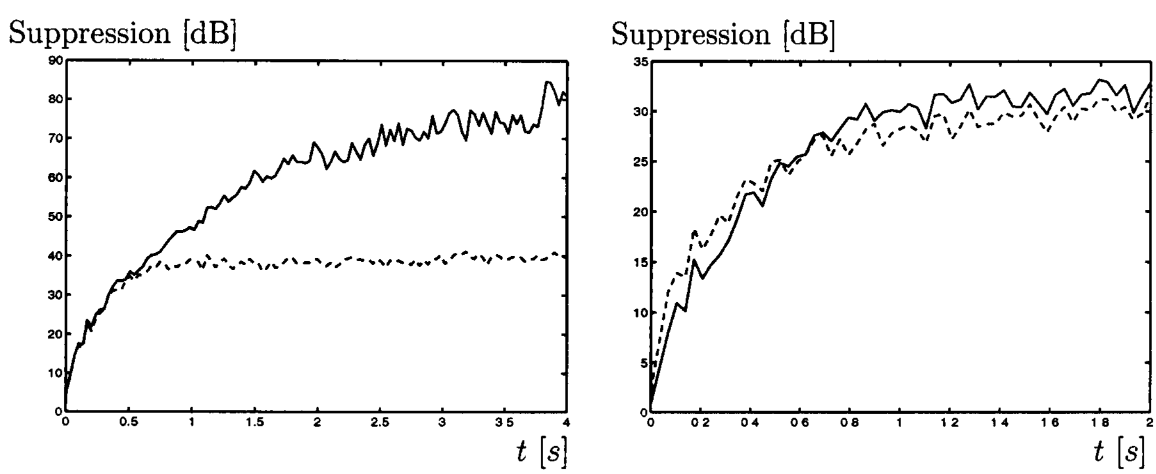


Figure 2: Computer simulations using a measured room acoustic impulse response(left) and "real signals"(right) resulted in the following suppression results: the dashed line shows the open loop case and solid line shows the closed loop case.

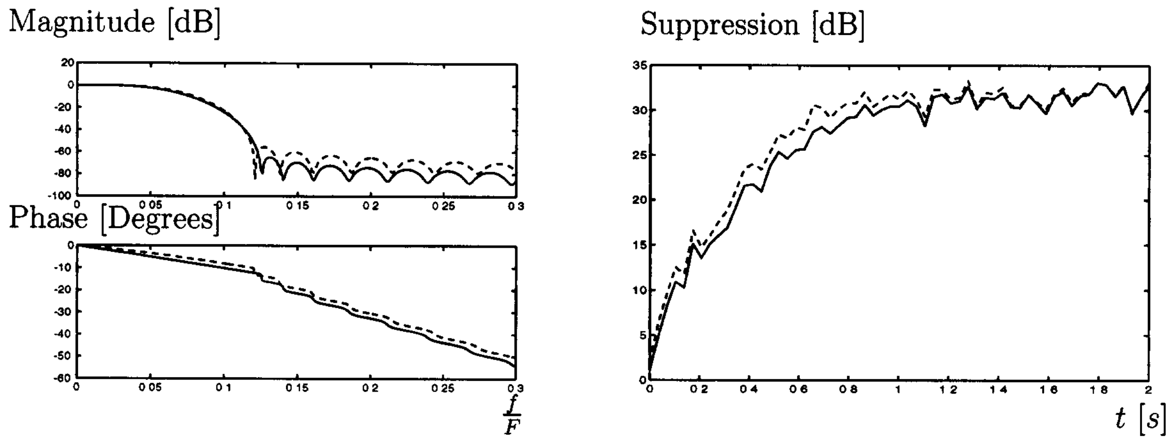


Figure 3: **Left**: Amplitude and phase response for two 64 tap prototype filters, 32 samples group delay(solid line) and 20 samples group delay(dashed line). **Right**: Suppression result for real broadband noise signals: 20 samples group delay(dashed line) and 32 samples group delay(solid line).

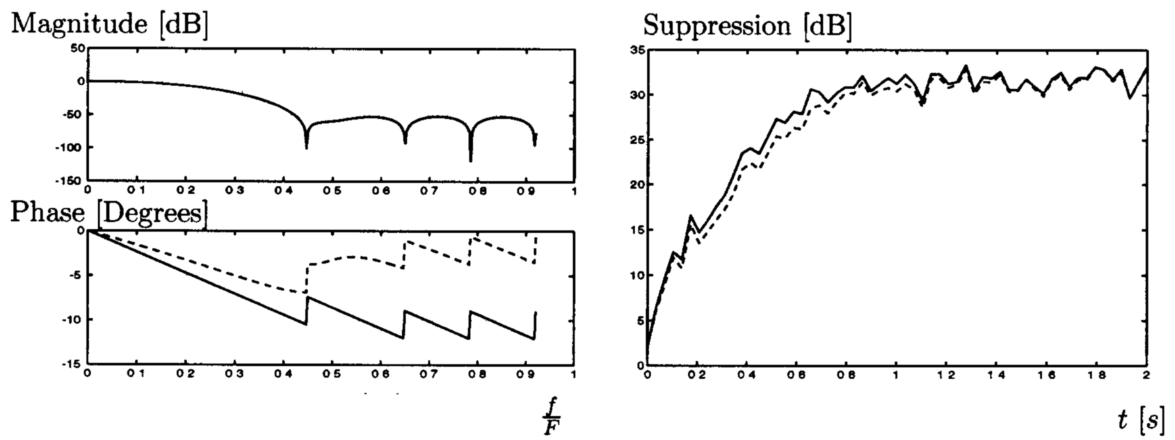


Figure 4: **Left**: Amplitude and phase response for two 48 tap prototype filters; min-phase(dashed line) and linear phase(solid line). **Right**: Suppression result for real broadband noise signals: min-phase phase(dashed line) and linear phase(solid line).

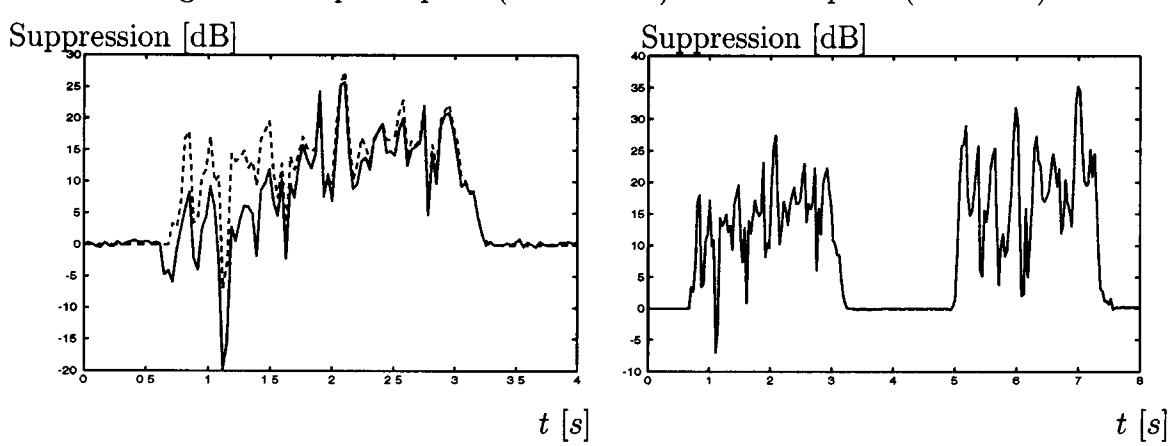


Figure 5: **Left**:Suppression result for real speech signals, with signal detector(dashed line) and without signal detector. **Right**: The echo canceller has much better performance on the second speech sequence than the first(these are two different sequences) due to the signal detector.