**Katholieke Universiteit Leuven**

Departement Elektrotechniek

ESAT-SISTA/TR 1996-65

# Frequency-domain adaptive echo cancellation as a special case of subband echo suppression [1]

Koen Eneman, Marc Moonen, Ian Proudler [2]

November 1996
Published in the Proceedings of ProRISC/IEEE Workshop on Circuits, Systems and Signal Processing,
Mierlo, the Netherlands, November 27-28 1996

[1] This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/eneman/reports/96-65.ps.gz*

[2] ESAT (SISTA) - Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven (Heverlee), Belgium, Tel. 32/16/321809, Fax 32/16/321970, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *koen.eneman@esat.kuleuven.ac.be*. Marc Moonen is a Research Associate with the F.W.O. Vlaanderen (Flemish Fund for Science and Research). This research was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven and has partly been made possible by the Concerted Research Action MIPS ('Model-based Information Processing Systems') of the Flemish Government and the Interuniversity Attraction Poles (IUAP-nr.17) initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture and was partially sponsored by Lernout & Hauspie Speech Products (Project 'Room Acoustic Echo Cancellation'). The scientific responsibility is assumed by its authors.

# Frequency-domain adaptive echo cancellation as a special case of subband echo suppression

*Koen Eneman*[1] *Marc Moonen*[1] *Ian Proudler*[2]

[1] ESAT - Katholieke Universiteit Leuven
Kardinaal Mercierlaan 94, 3001 Heverlee - Belgium
koen.eneman@esat.kuleuven.ac.be
marc.moonen@esat.kuleuven.ac.be
[2] Defence Research Agency, Room E506
St Andrews Road, Malvern Worcs., WR14 3PS, UK
proudler@signal.dra.hmg.gb

*Abstract*—**Modelling acoustic impulse responses with lengths up to 250 ms is required for high quality echo cancellation in strongly reverberating environments, leading to FIR adaptive filters with several thousands of taps. Classical LMS-based solutions then clearly fail as they exceed the computational capabilities of present-day DSPs. Cheaper alternative solutions have been proposed and are mainly based on either subband or frequency-domain techniques. Both the subband and the frequency-domain approach turn out to have their strong and weak points. Subband filter implementations have many interesting properties and apparently allow to achieve the lowest possible computational cost. However, their inherent delay and residual errors have made them unattractive for real-time applications up till now. On the other hand, it is known that frequency-domain adaptive filters do not suffer from these problems despite being (nearly) equivalent to subband adaptive filters with 'poor' filter banks. In this paper we explain the operation of the frequency-domain techniques in the 'subband jargon' and make an attempt at pointing out how to enhance subband performance. This latter is a subject of current research.**

## I. Introduction

For high quality echo cancellation long acoustic echoes need to be suppressed. Acoustic echo paths are characterised by FIR filters with lengths up to 250 ms. Filters clocked at a rate of, say, 10 kHz then require several thousands of filters taps to be identified. Classical LMS-based echo cancellers, **figure 1**, are unattractive for real-time processing as their computational requirements clearly exceed the capabilities of present-day DSPs. Moreover, speech signals have a coloured spectrum and it is well known that the performance of the LMS algorithm is suboptimal in that case, especially when extremely long FIR filters are being adapted. Therefore alternative solutions have been proposed and they are mainly based on either subband or frequency-domain techniques.

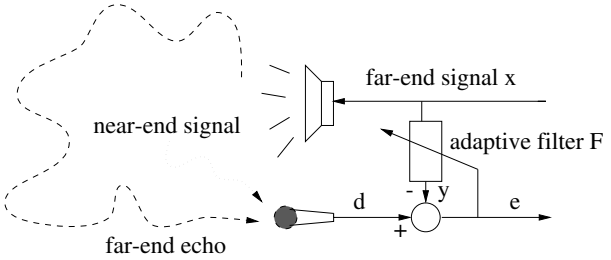



Fig. 1. classical approach

Such multirate adaptive echo cancellation schemes have been a topic of interest for many years now. Still, with the available techniques, it is difficult to meet all the specifications with respect to delay, echo suppression, distortion and computational requirements. Subband adaptive filtering and frequency-domain based techniques are mostly considered as being different approaches. In literature most attention is now paid to subband filtering, whereas frequency-domain schemes are already used in commercial long acoustic echo cancellers.

In this paper we will consider the frequency-domain approach as a special case of subband adaptive filtering having some desired properties and point out why frequency-domain techniques are better from certain perspectives, or -at least- are able to compete with

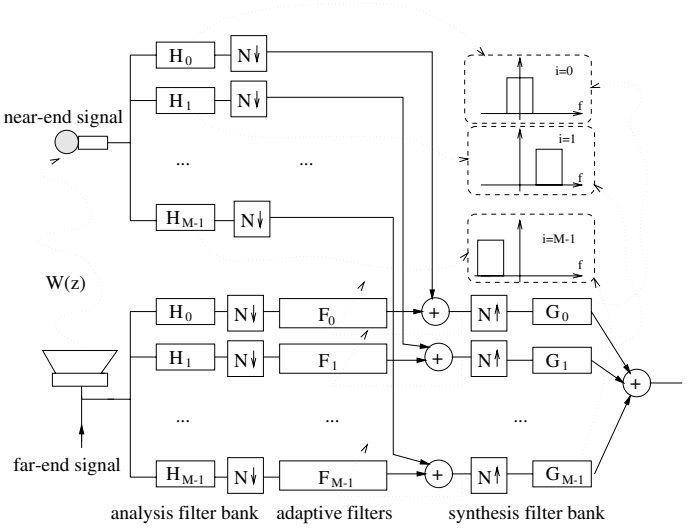



Fig. 2. subband adaptive echo canceller

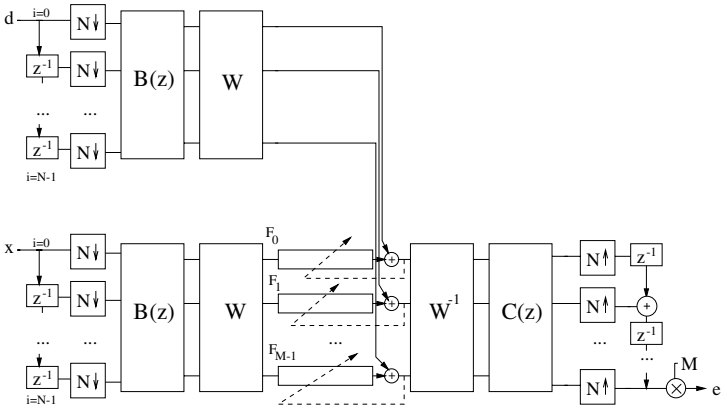



Fig. 3. DFT modulated adaptive echo canceller

subband schemes.

## II. Subband Adaptive Filtering

The general setup for a subband acoustic echo canceller is shown in **figure 2**. The loudspeaker and microphone signal are fed into identical M-band analysis filter banks. After subsampling with a factor N, (mostly LMS-based) adaptive filtering is done in each subband. The outputs of the subband adaptive filters are recombined in the synthesis filter bank leading to the final output. The ideal frequency amplitude characteristics of the analysis bank filters $H_i$ and synthesis bank filters $G_i$ are shown (ideal bandpass filters).
Splitting signals into subbands with subsequent subsampling has the advantage of giving better conditioned covariance matrices than obtained with fullband LMS, thus leading to *faster (initial) convergence and better tracking* if properly tuned. A second advantage of subband systems over classical fullband adaptation is the *lower achievable implementation cost* (due to downsampling).
A subband echo canceller requires a well designed filter bank. Filter banks introduce *aliasing errors* (which deteriorates the adaptive filtering operations), an inherent *delay* and perhaps unacceptable signal *distortion* (non-perfect reconstruction banks).

## III. DFT modulated Subband Adaptive Filters

Subband acoustic echo cancellers are mainly based on DFT modulated filter banks. In a DFT modulated filter bank, M subband filters are derived from a single prototype filter $h_0(k)$ : the subband filters are then frequency shifted versions of each other and the complete set of M filters covers the whole frequency spectrum.
A DFT modulated filter bank can be implemented efficiently thanks to the relation between the different subband filters and making use of polyphase decomposition and fast signal transforms. The filter bank basically consists of a tapped delay line of size N, a structured $M$x$N$-matrix $\mathbf{B}(z)$, containing polyphase components of the prototype filter $h_0$, and a $M$x$M$-DFT matrix $\mathbf{W}$ (**figure 3**). This filter bank representation and some filter bank design rules can be found in [1]. It is shown that element $(i,j)$ of $\mathbf{B}(z)$ is given by :

$$B_{ij}(z) = z^{-l} E_{(j+lN):K}(z^L) \qquad (1)$$

$$\text{with } i,j \geq 0, (j+lN) \bmod M = i \\ \text{and } (i-j) \bmod g = 0$$

$E_{k:K}(z)$ is the k-th element of the K-th order polyphase decomposition of the prototype filter $h_0$. $L = \frac{M}{g}$, $g$ is the greatest common divisor of $M$ and $N$, $K$ is the least common multiple of $M$ and $N$. The synthesis bank is constructed in a similar fashion (matrix $\mathbf{C}(z)$).
The perfect reconstruction condition, which (is certainly not the most general one, but is used in the sequel and which) ensures that the near-end source signal is not distorted by the analysis/synthesis system, is

$$\mathbf{C}(z)\mathbf{W}^{-1}\mathbf{W}\mathbf{B}(z) = \mathbf{C}(z)\mathbf{B}(z) = I \qquad (2)$$

Furthermore, the aim is to mimic the acoustic channel $W(z)$ by means of adaptive filters which corresponds to

$$\mathbf{C}(z)\mathbf{W}^{-1}\text{diag}\{F_i(z)\}\mathbf{W}\mathbf{B}(z) \approx$$

$$\underbrace{\begin{bmatrix} W_0 & W_1 & \dots & W_{N-1} \\ z^{-1}W_{N-1} & W_0 & \dots & W_{N-2} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix}}_{=\text{pseudo-circulant}} \quad (3)$$

where

$$\sum_{n=0}^{N-1} z^{-n} W_n(z^N) \quad (4)$$

is a polyphase decomposition of $W(z)$. As will be illustrated, this condition cannot always be satisfied.
Two examples are given : a critically downsampled subband system and a system which is 2-times oversampled.

### A. Critically Downsampled Subband Schemes

For critically downsampled subband systems $M = N$. It can be seen from Eq. 1 that the filter matrix $\mathbf{B}$(z) is a $M$x$M$-matrix and will take on the form :

$$\mathbf{B}(z) = \begin{bmatrix} E_0(z) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & E_{(M-1)}(z) \end{bmatrix} \quad (5)$$

From the P.R. condition it follows that $\mathbf{C}(z) = \mathbf{B}^{-1}(z)$.
Critically downsampled subband systems are attractive because *optimal computational savings* can be obtained when $N$ is as high as possible.
In the ideal case with infinite order analysis and synthesis filters the prototype filters are ideal lowpass and their impulse response is a sinc function. A causal implementation then introduces an *infinite delay* which obviously is intolerable. Hence, FIR filters have to be used with acceptable processing delay. Finite order filters, however, always have a non-negligible transition band and therefore critically downsampled subband systems will always introduce some *aliasing* in the subbands. Aliasing is detrimental for overall convergence. In [2] it is shown that critically downsampled subband systems lead to a residual error which is considerable unless cross filters are included between neighbouring subbands. Cross filters again increase the complexity, which is unwanted. Furthermore, cross filters fail to converge quickly to the optimum solution. This suggests the use of oversampled subband schemes where $M > N$.

### B. 2-times Oversampled Subband Systems

In this case the subband signals are subsampled 2-times less with respect to the critical scheme. Here $M = 2N$ and the $\mathbf{B}$(z)-matrix is a 2NxN matrix :

$$\mathbf{B}(z) = \begin{bmatrix} E_0(z^2) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & E_{N-1}(z^2) \\ z^{-1}E_N(z^2) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & z^{-1}E_{M-1}(z^2) \end{bmatrix} \quad (6)$$

From a computational point of view 2 times oversampled subband systems seem to be less interesting as -roughly speaking- their *implementation cost is 4 times higher* : the required subband adaptive filters are roughly twice as long and operate at twice the sampling rate. As will be shown later on, *frequency-domain adaptive filters can be cast in the subband approach as a special 2-times oversampled subband system.* Frequency-domain adaptive filters have desirable convergence properties, and hence provide probably the best solution in those cases where 2-times oversampling is affordable.
In conclusion, the above examples represent 2 extremes, the critically downsampled scheme interesting at first sight from a computational point of view but with poor performance, the 2-times oversampled system with the frequency-domain scheme in mind, more expensive but with desirable convergence properties. The ultimate goal may be to find acceptable performance with $\frac{M}{2} \leq N \leq M$.

## IV. Pros and Cons of Subband Echo Cancellers

We first return to the subband approach and consider the general case of oversampled subband systems for which $M > N$ : critically downsampled echo cancellers have already shown to be unattractive.
Splitting signals into subbands seems very promising since for coloured input spectra, fullband convergence is slow due to ill-conditioned covariance matrices. In the subband case, each subband signal will have a flatter spectrum after appropriate subsampling, leading to *improved convergence.* Instead of a single fullband $L_a$-taps FIR filter, M subband filters of, say, $\frac{L_a}{N}$ taps are used to model the acoustic path (see fig. 2). As the adaptive computations as well as the filter bank convolutions can be done at a reduced sampling rate, this subband approach is supposed to give a better performance at a *lower cost.*
It is clear that this picture is certainly too optimistic. Whereas critically downsampled subband schemes inherently suffer from subband aliasing, in oversampled subband adaptive systems reduced *aliasing distortion*

is exchanged for extra costs and slower steady-state convergence. Moreover *the assumption of having M subband filters with reduced length* $\frac{L_a}{N}$ *seems to be quite wrong.*

One can prove that in the case of an M-band, N-times downsampled ideally frequency selective filter bank the i-th subband adaptive filter should converge to $\hat{w}_i(k)$ :

$$\hat{w}_i(k) = \left( w(m) * e^{-j\frac{2\pi i m}{N}} \mathrm{sinc}(\frac{\pi m}{N}) \right)_{N\downarrow}, \qquad (7)$$

i.e., $\hat{w}_i(k)$ can be obtained by downsampling the convolution of the acoustic path $w(m)$ and a modulated double-sided sinc. This corresponds to an interpolation operation. The adaptive identification process therefore has to track more than $\frac{L_a}{N}$ samples and due to the spreading out in both directions of the time axis, an extra delay has to be inserted in the near-end signal path. This is illustrated in **figure 4**. Neglecting

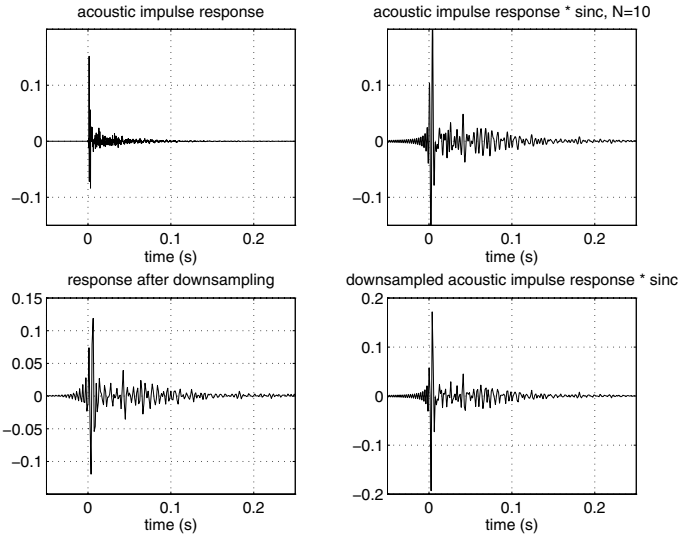



Fig. 4. The first (left, top) subplot shows an impulse response obtained from real, measured data. Loudspeaker and microphone are close to each other. In subplot 2 (right, top), this acoustic impulse response is convolved with a sinc. In subplot 4 (right, bottom), the downsampled version of subplot 2 is shown. This corresponds to what was obtained from real data (subplot 3).

the additional subband filter length due to these sinc-effects strongly limits the convergence of the adaptive filter and leads to a residual error. The ITU-T norms [6] suggest a 40 dB echo suppression such that even with a nonlinear postprocessor (e.g. a centre clipper) computational advantages of subband schemes shrink. These anti-causal sinc-effects are considerable in the case of a small loudspeaker-to-microphone distance. Furthermore, the *delay constraints* also make subband schemes unattractive. Selective filter banks are needed to avoid aliasing distortion within subbands. They introduce a substantial processing delay and thus put a constraint on the downsampling factor $N$. However, computational savings are more or less inversely proportional to $N$.

## V. Frequency Domain Adaptive Filters

As a cheaper alternative to LMS, the frequency-domain adaptive filter (FDAF) was introduced, which is a direct translation of Block LMS in frequency domain [5]. Correlation (weight updating) and convolution (filtering) operations are expensive but in the case of block processing, they may be implemented more efficiently in frequency domain. Instead of a linear convolution/correlation a circular operation is performed. This requires some 'restore' operations which can be of the overlap-save or overlap-add type. If only the convolution operation is corrected a so-called *unconstrained FDAF* is obtained requiring 3 FFTs. Two more FFTs are needed for the gradient estimate correction resulting in a *constrained FDAF.*

A major drawback concerning standard frequency-domain adaptive filters is the inherent delay. For a realistic cancellation setup with an adaptive filter length being 1600 taps @ 8 kHz, the delay is twice as long as the acoustic impulse response, i.e. 400 ms.

### A. Partitioned Block Frequency Domain Adaptive Filtering

By splitting the acoustic impulse response in equal parts, a kind of mixed time and frequency convolution canceller is obtained, called the *Partitioned Block Frequency-Domain Adaptive Filter* (PBFDAF)[3],[4]. Here block lengths can be adjusted, resulting in a cheap echo canceller with acceptable processing delay.

### B. The PBFDAF as a special case of Subband Adaptive Filtering

This PBFDAF scheme can be put into the oversampled subband framework proposed in [1]. The PBFDAF implements a simple filter bank with low frequency selectivity and 1-tap filter bank polyphase components. This easily admits re-transformation of subband errors into time domain where a repair operation restores aliasing errors without introducing extra delay. It is remarkable how an unselective filter bank can lead to satisfactory results.

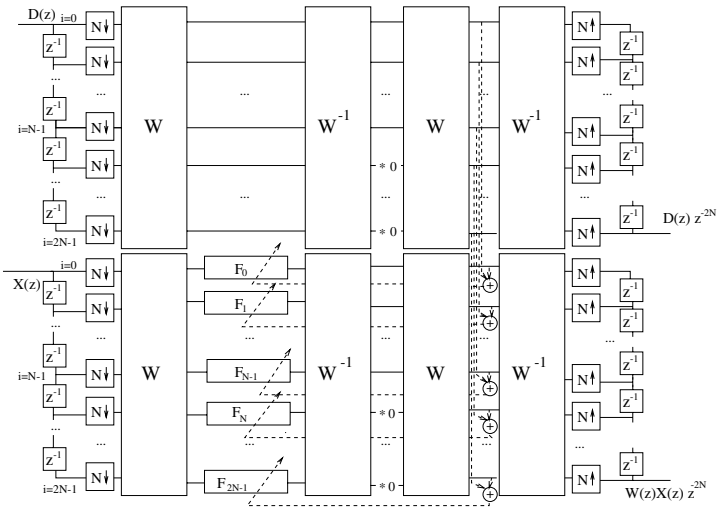



Fig. 5. PBFDAF as a special case of subband adaptive filtering

Call $X(z)$ the far-end signal and $D(z)$ the near-end signal, so

$$D(z) = S(z) + W(z)X(z) \tag{8}$$

where $S(z)$ is the contribution of a near-end source. The acoustic impulse response $W(z)$ can be split up in its N-th order polyphase components :

$$D(z) = S(z) + \sum_{n=0}^{N-1} z^{-n} W_n(z^N) X(z) \tag{9}$$

This leads to equation 10, which is rewritten as

$$\mathbf{D}(z) = \mathbf{S}(z) + \mathbf{M}(z)\mathbf{X}(z) \tag{11}$$

The transfer matrix $\mathbf{M}$(z) was made circulant so that it can be transformed into a diagonal matrix by means of DFT operations, i.e., $\mathbf{W}\mathbf{M}(z)\mathbf{W}^{-1} = \text{diag}\{\hat{W}_i(z^N)\}$. $\hat{W}_i(z)$ are related to the DFT coefficients of the first column of $\mathbf{M}$(z) and therefore they are of finite length. Instead of identifying a half-full matrix $\mathbf{M}$(z), a diagonal matrix can be tracked in frequency-domain. An adaptive identification process trying to match $W(z)$ in frequency domain based on the above formulas is depicted in **figure 5**.
Looking closer, fig. 5 can be cast in the oversampled subband framework of fig. 3, i.e. with size N (instead of size 2N) tapped delay lines, together with $\mathbf{B}(z)$ and $\mathbf{C}(z)$. The $\mathbf{B}(z)$-matrix for a 2-times oversampled DFT modulated *analysis filter bank* is given in eq. 6. The filter bank used here is a simple DFT

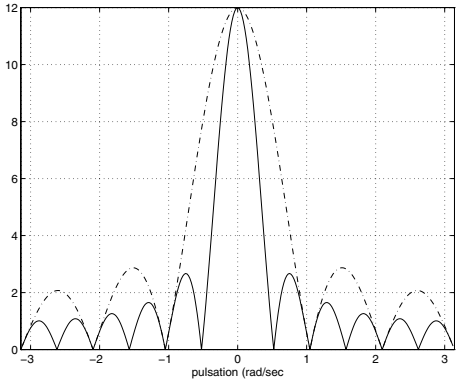



Fig. 6. analysis and synthesis bank prototype filters

filter bank for which $E_n(z) = 1$, i.e.

$$\mathbf{B}(z) = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ z^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z^{-1} \end{bmatrix} \tag{12}$$

The prototype frequency response has a sinc-like shape with a low frequency selectivity. The analysis prototype frequency amplitude response in shown in **figure 6** in full line for M=12 and N=6.
Also the *synthesis part* can be fit into the subband filter approach. The synthesis bank $\mathbf{C}(z)$-matrix is then given by :

$$\mathbf{C}(z) = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix} \tag{13}$$

The synthesis filters are DFT modulated versions of a prototype whose frequency response is twice as wide as the analysis equivalent. The synthesis prototype frequency response is shown in fig. 6 in the dashed dotted line.
It is easily verified that for this analysis/synthesis system, both the P.R. condition $\mathbf{C}(z)\mathbf{B}(z) = I$ and the additional condition $\mathbf{C}(z)\mathbf{W}^{-1}\text{diag}\{F_i(z)\}\mathbf{W}\mathbf{B}(z) \approx \begin{bmatrix} W_0 & \cdots & W_{N-1} \\ z^{-1}W_{N-1} & \cdots & W_1 \\ \vdots & \ddots & \vdots \end{bmatrix}$ is satisfied for $F_i = \hat{W}_i$.

## VI. Error Signal Correction

The implicit error 'restore' or projection operation in frequency-domain adaptive filters consists of a transformation to time domain, zeroing of certain components and transformation back to the frequency domain (see figure 5). No projection operations are

$$\begin{bmatrix} D(z) \\ z^{-1}D(z) \\ \cdots \\ \underline{z^{-(N-1)}D(z)} \\ * \\ * \\ \cdots \\ * \end{bmatrix} = \begin{bmatrix} S(z) \\ z^{-1}S(z) \\ \cdots \\ \underline{z^{-(N-1)}S(z)} \\ z^{-N}S(z) \\ z^{-(N+1)}S(z) \\ \cdots \\ z^{-(2N-1)}S(z) \end{bmatrix} +$$

$$\left[\begin{array}{cccccc} W_0(z^N) & \cdots & W_{N-1}(z^N) & 0 & \cdots & 0 \\ 0 & W_0(z^N) & \cdots & W_{N-1}(z^N) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & W_0(z^N) & \cdots & W_{N-1}(z^N) & 0 \\ \hline 0 & \cdots & 0 & W_0(z^N) & \cdots & W_{N-1}(z^N) \\ W_{N-1}(z^N) & \cdots & 0 & 0 & W_0(z^N) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ W_1(z^N) & \cdots & 0 & 0 & \cdots & W_0(z^N) \end{array}\right] \left[\begin{array}{c} X(z) \\ z^{-1}X(z) \\ \cdots \\ \underline{z^{-(N-1)}X(z)} \\ z^{-N}X(z) \\ z^{-(N+1)}X(z) \\ \cdots \\ z^{-(2N-1)}X(z) \end{array}\right] \tag{10}$$

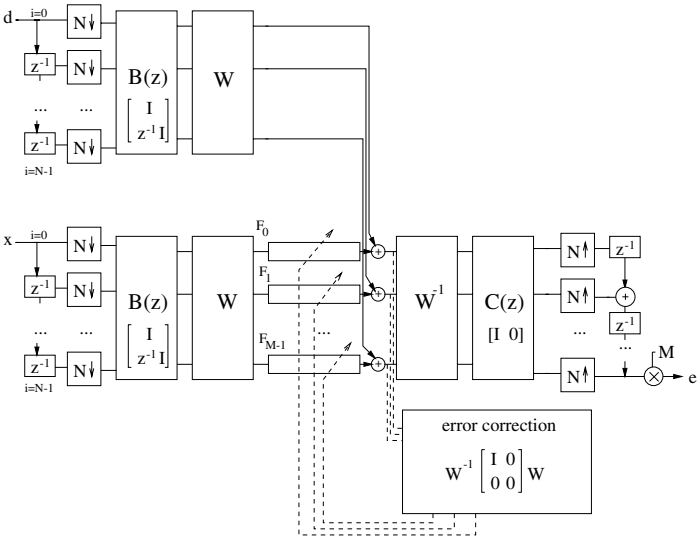



Fig. 7. Oversampled subband schemes with constraint adaptation

applied in a standard subband approach. The challenge is now to look for appropriate filter banks and an error signal correction to come up with an oversampled subband systems that inherits some of the properties of the PBFDAF but at a lower cost. A general scheme is an extension of fig. 3 and 5 and is depicted in **figure 7**. This is a topic of current research.

## VII. Conclusions

We showed that a PBFDAF implements a simple filter bank with low frequency selectivity based on 1-tap filter bank polyphase components. This easily admits re-transformation of subband errors into time domain where a repair operation restores aliasing errors without introducing extra delay. We indicated that exploiting these ideas for oversampled subband acoustic echo cancellers could alleviate the artefacts associated with subband acoustic echo echocancellers.

## VIII. Acknowledgements

Marc Moonen is a Research Associate with the F.W.O. Vlaanderen (Flemish Fund for Science and Research). This research was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven and has partly been made possible by the Concerted Research Action MIPS ('Model-based Information Processing Systems') of the Flemish Government and the Interuniversity Attraction Poles (IUAP-nr.17) initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture and was partially sponsored by Lernout & Hauspie Speech Products (Project 'Room Acoustic Echo Cancellation'). The scientific responsibility is assumed by its authors.

## References

[1] Z. Cvetković, Oversampled Modulated Filter Banks and Tight Gabor Frames in $l^2(\mathbb{Z})$, Proceedings of the 1995 IEEE Int. Conf. on Acoust., Speech and Signal Processing, Detroit, Michigan, USA, pp. 1456-1459, May 1995

[2] A. Gilloire and M. Vetterli, Adaptive Filtering in Subbands with Critical Sampling : Analysis, Experiments and Application to Acoustic Echo Cancellation,"IEEE Trans. Signal Processing", vol. 40, nr. 8, pp. 1862-1875, August 1992

[3] J. Páes Borrallo and M. García Otero, On the implementation of a partitioned block frequency domain adaptive filter (PBFDAF) for long acoustic echo cancellation, "Signal Processing", vol. 27, pp. 301-315, June 1992

[4] J.-S. Soo and K. Pang, Multidelay Block Frequency Domain Adaptive Filter, "IEEE Trans. Acoust., Speech and Signal Processing", vol. 38, nr. 2, pp. 373-376, February 1990

[5] J. Shynk, Frequency-Domain and Multirate Adaptive Filtering, "IEEE Signal Processing Magazine", pp. 15-37, January 1992

[6] International Telecommunication Union, General Characteristics of International Telephone Connections and International Telephone Circuits : Acoustic Echo Controllers (ITU-T, Recommendation G.167), March 1993