BBC

Research & Development

White Paper

WHP 204

September 2012

Techniques for high-performance DAB and DVB-T On-Channel Repeaters

Peter Moss ***and*** **Adam Wiewiorka**

BRITISH BROADCASTING CORPORATION

White Paper WHP 204

Techniques for high-performance DAB and DVB-T On-Channel Repeaters

Peter Moss and Adam Wiewiorka

Abstract

On-channel repeaters are able to receive and re-transmit a broadcast signal on the same frequency and thus can facilitate network expansion, particularly in an SFN, and network repair following the introduction of strong adjacent-channel transmissions. For effective operation, systems must incorporate excellent RF design and high-performance signal processing in order to remain stable and re-radiate a clean, decodable signal.

This paper describes the operation of on-channel repeaters designed for DAB and DVB-T signals by BBC R&D, and goes on to outline a number of key techniques, including correlation compensation for controlling unwanted out-of-band products due to part-Nyquist signal occupancy and adaptive null-steering for pre-echo rejection in a DAB SFN. The former is now a field-proven technique whilst the latter is in development and awaiting a full engineering realisation.

This document is based on that originally published in the International Journal of Sensors, Wireless communications and Control vol. 1 no. 1 June 2011 ISSN:2210-3279

Additional key words: Active deflector, DAB, DVB-T, On-channel repeater, SFN, transposer.

White Papers are distributed freely on request.

Authorisation of the Chief Scientist or General Manager is required for publication.

© BBC 2012. All rights reserved. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at http://www.bbc.co.uk/rd/pubs/whp for a copy of this document.

White Paper WHP 204

Techniques for high-performance DAB and DVB-T On-Channel Repeaters

Peter Moss and Adam Wiewiorka

List of Abbreviations

BER	Bit error ratio
DAB	Digital Audio Broadcasting
DMB	Digital Multimedia Broadcasting
DVB-T	Digital Video Broadcasting (Terrestrial)
DVB-T2	Digital Video Broadcasting (Terrestrial), second generation
DVB-H	Digital Video Broadcasting (Handheld)
IF	Intermediate frequency
IP3	Third order intercept
OCR	On-channel repeater
MER	Modulation Error Ratio
MMSE	Minimum Mean Squared Error
PDP	Power delay profile
PSD	Power spectral density
RF	Radio frequency
SFN	Single Frequency Network
UHF	Ultra High Frequency

White Paper WHP 204

Techniques for high-performance DAB and DVB-T On-Channel Repeaters

Peter Moss and Adam Wiewiorka

1 Introduction

Within a network of broadcasting transmitters operating in the VHF or UHF bands, the means of delivering the payload signal to each individual node may often incur considerable cost, particularly where rented land-lines or fixed microwave links are required. An alternative approach for low-cost relatively low power transmitters is to obtain the excitation signal 'off-air', amplify and re-transmit. This is a straightforward method if a frequency change is also acceptable. But in a Single Frequency Network (SFN), which is desirable for reasons of spectral efficiency, the option of frequency changing is not open to the network designer. Therefore in this case an off-air repeater must re-transmit on the same frequency as it receives.

In the past such an approach was usually precluded by the problem of unwanted coupling from the output to the input, causing frequency response aberrations or indeed instability. To obtain sufficient isolation at VHF or UHF frequencies when a low-level input is to be amplified to a useful re-broadcast signal power, a very large and impractical physical separation is required between input and output antennas if this alone is to provide the necessary attenuation.

The approach to the design of an on-channel repeater (OCR) described in this paper is based on suitably adapted echo cancellation techniques used extensively in digital signal processing for audio applications. Although the principles of these methods have been known for a considerable time, a commercially viable implementation using digital hardware running at the speed required for digital broadcasting only became possible around 2002. The theory of operation is explained in the first parts of this paper, together with the mathematical basis of the adopted solution. Following that, some specific techniques designed to address particular performance limitations of the basic architecture are presented. These techniques include some which have already been field-proven, and others which at this stage are at the design evaluation and simulation phase.

2 On-channel re-transmission

2.1 RF repeaters

Broadly speaking, there are two classes of application for RF repeaters in a broadcast network. These may be loosely termed coverage extension and network repair. Coverage extension refers to an increase of the total served area by extending the network at the edges, whereas network repair, as the name implies, is concerned with filling a gap in coverage caused, for instance, by interference from a high power adjacent channel transmitter.

A conceptual architecture is shown in figure 1 below, where RF filtering is not shown for clarity. In practice however, such a simple arrangement can rarely be used, unless the output/input power ratio only needs to be very small, or the isolation has been made very high by some means, thereby reducing the unwanted feedback to acceptable levels.

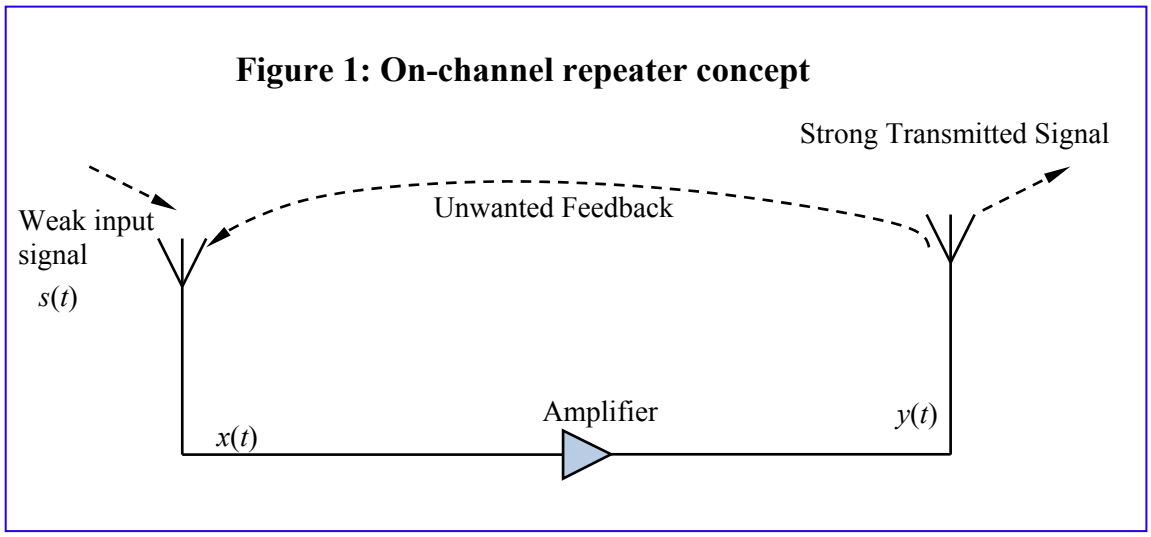

Figure 1: On-channel repeater concept

2.2 Adaptive cancellation architecture

A more sophisticated architecture based on adaptive path cancellation and capable of useful performance in a variety of scenarios is shown in figure 2.

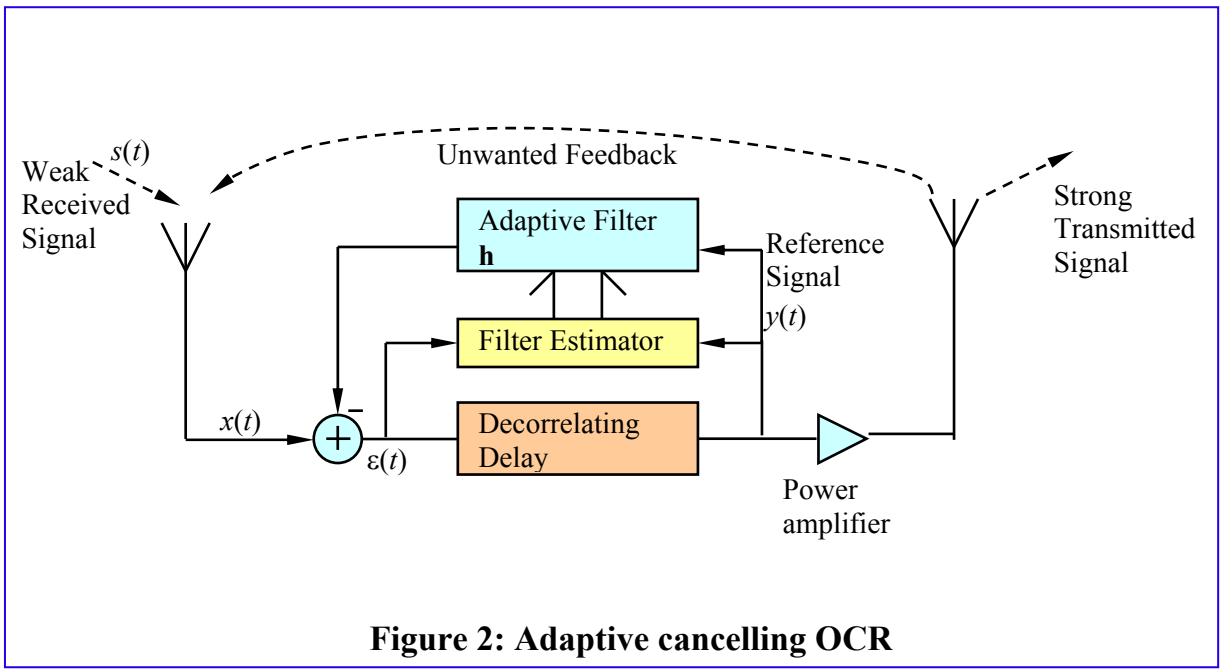

Figure 2: Adaptive cancelling OCR

The cancellation of parasitic feedback is performed by an adaptive finite impulse response filter which models the path between the antennas. The filter impulse response **h** can be estimated using correlation methods, provided that the delay through the system is sufficiently long to avoid the effects of autocorrelation. In other words, the retransmitted signal, and hence the feedback, must be uncorrelated with the received signal, so that it can be unambiguously identified.

The simplest, and in most cases entirely sufficient, method of estimating filter taps is the Least Mean Square (LMS) algorithm. Given the input $x(t)$ and output $y(t)$, the objective is to minimize the error $\varepsilon(t)$, leaving only the received signal $s(t)$. This signal is recovered by subtracting from $x(t)$ the output $y(t)$ filtered by **h**, as shown below:

$$\varepsilon(t) = x(t) - \mathbf{h}^{\mathrm{T}}\mathbf{y}(t) \quad \text{.....(1)}$$

Vector $\mathbf{y}(t)$ contains $y(t)$, $y(t\text{-}1)$ to $y(t\text{-}K\text{-}1)$ where K is the length of $\mathbf{h}$. Ideally all remnants of $y(t)$ are removed from $x(t)$, and the energy in $\varepsilon(t)$ is minimized. In order to achieve this goal the filter taps $\mathbf{h}$ must be estimated. First the expectation of squared magnitude of $\varepsilon(t)$ is defined

$$\mathrm{E}\,|\,\varepsilon(t)\,|^2 = \mathrm{E}\,|\,x(t) - \mathbf{h}^{\mathrm{T}}\mathbf{y}(t)\,|^2 \quad \text{.....(2)}$$

and differentiated with respect to $\mathbf{h}$:

$$\frac{d\,\mathrm{E}\,|\,\varepsilon(t)\,|^2}{d\mathbf{h}} = -\mathrm{E}\left\{\mathbf{y}(t)\left[x(t) - \mathbf{h}_t^{\mathrm{T}}\mathbf{y}(t)\right]^*\right\} = -\mathrm{E}\left[\mathbf{y}(t)\varepsilon(t)^*\right]$$

.....(3)

The taps are then updated with an appropriately scaled conjugate of the instantaneous value of the derivative:

$$\mathbf{h}_t = \mathbf{h}_{t-1} + \mu\varepsilon(t)\mathbf{y}(t)^* \quad \text{.....(4)}$$

The value of μ is chosen as a compromise between the speed of convergence and signal-to-noise ratio. In practice, faster convergence allows the algorithm to follow rapid changes in the feedback path caused by the time-varying nature of the coupling, at the expense of increased noise in the filter tap estimates $\mathbf{h}$, which affects feedback cancellation. Similarly, in the case of slow convergence and in the presence of Doppler, $\varepsilon(t)$ may contain a significant remnant of $y(t)$, potentially causing instability. Because the level of this residual feedback is proportional to the loop gain of ε, the amount of Doppler variability in $x(t)$ imposes an upper limit on the gain of the repeater.

As discussed in Marple [1], convergence also depends on the eigenvalues of the covariance matrix $\mathbf{R}_{yy}$=E(**yy***). Optimal behaviour occurs when all eigenvalues are equal, but if $\mathbf{R}_{yy}$ is not full rank, or some eigenvalues are close to zero, convergence is significantly impaired. One of the consequences of poor convergence might be large amounts of noise generated in unused areas near the edges of the Nyquist band. For this reason, the choice of sampling rate has to be carefully addressed and best results are achieved if the input signal is noise-like in nature.

3 Optimising the performance

3.1 Residual correlation

3.1.1 Origin

Although for reasons of practical implementation it is not possible to allow the wanted signal to fully occupy the full Nyquist bandwidth of the baseband DSP signal, the consequence of having an 'unused' part of the available Nyquist range is unwanted input-output cross-correlation, even in the presence of an otherwise sufficient decorrelating delay. This residual correlation perturbs the ideal coefficient values of the echo cancelling filter and thereby creates unwanted effects, as discussed below. Although mitigating techniques will also be described, it is good design practice to minimise this problem 'at source' by maximising the Nyquist occupancy whilst maintaining realistic demands on associated filtering. An occupancy of 90%+ is desirable and still allows realisable filters.

3.1.2 Unwanted effects

As the LMS equaliser is attempting to minimise correlation, whatever the source, there is a tendency for parasitic energy, in the form of tones, to appear in the unused Nyquist spectrum. A typical spectrum of the error signal $\varepsilon(t)$ illustrating the problem is shown below in figure 3 which

shows both the out-of-band tones and excessive ripple in the signal itself. To exaggerate the problem the Nyquist occupancy has been reduced to around 50% in this simulation.

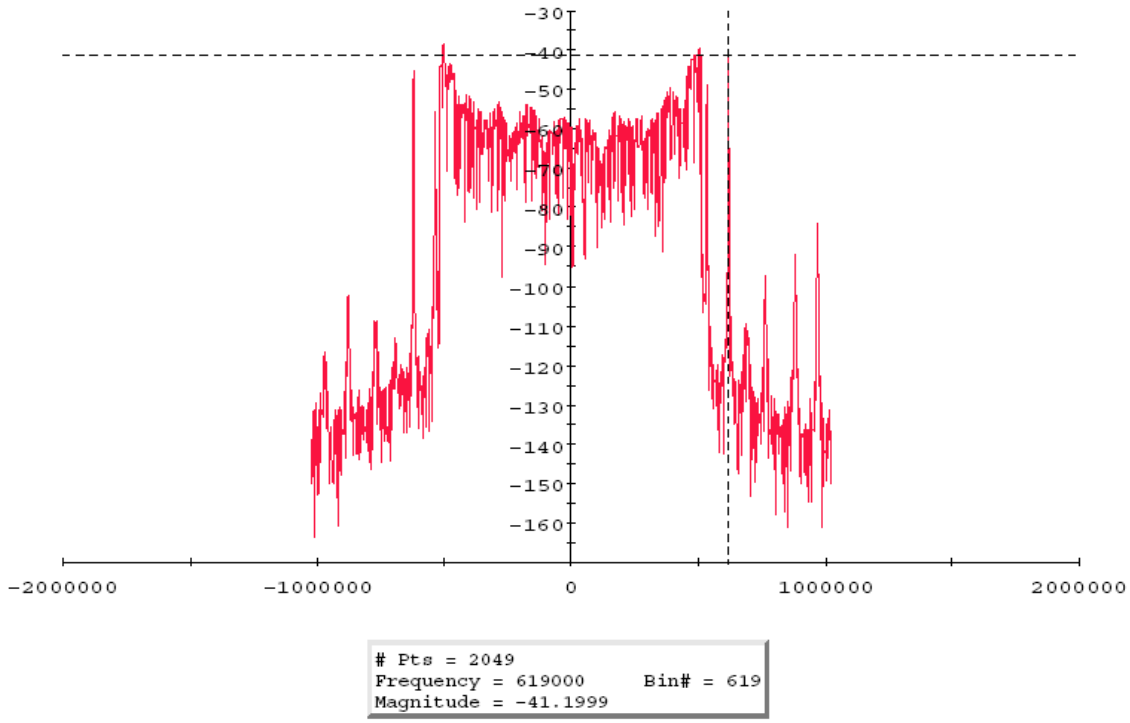

Figure 3: Spectrum of system error (i.e the OCR output)

3.1.3 Correlation compensation

An expression for the tap perturbation caused by the unwanted (part-Nyquist) correlation is derived below, and a method of compensating for it described.

3.1.3.1 Derivation of convergence error term

Each elemental correlator/tap combination within the LMS equaliser structure shown in figure 4 reaches steady state when the average correlator output is zero. This condition is required by the presence of the integrator steering each tap, as illustrated by the k^{th} tap of a multi-tap system.

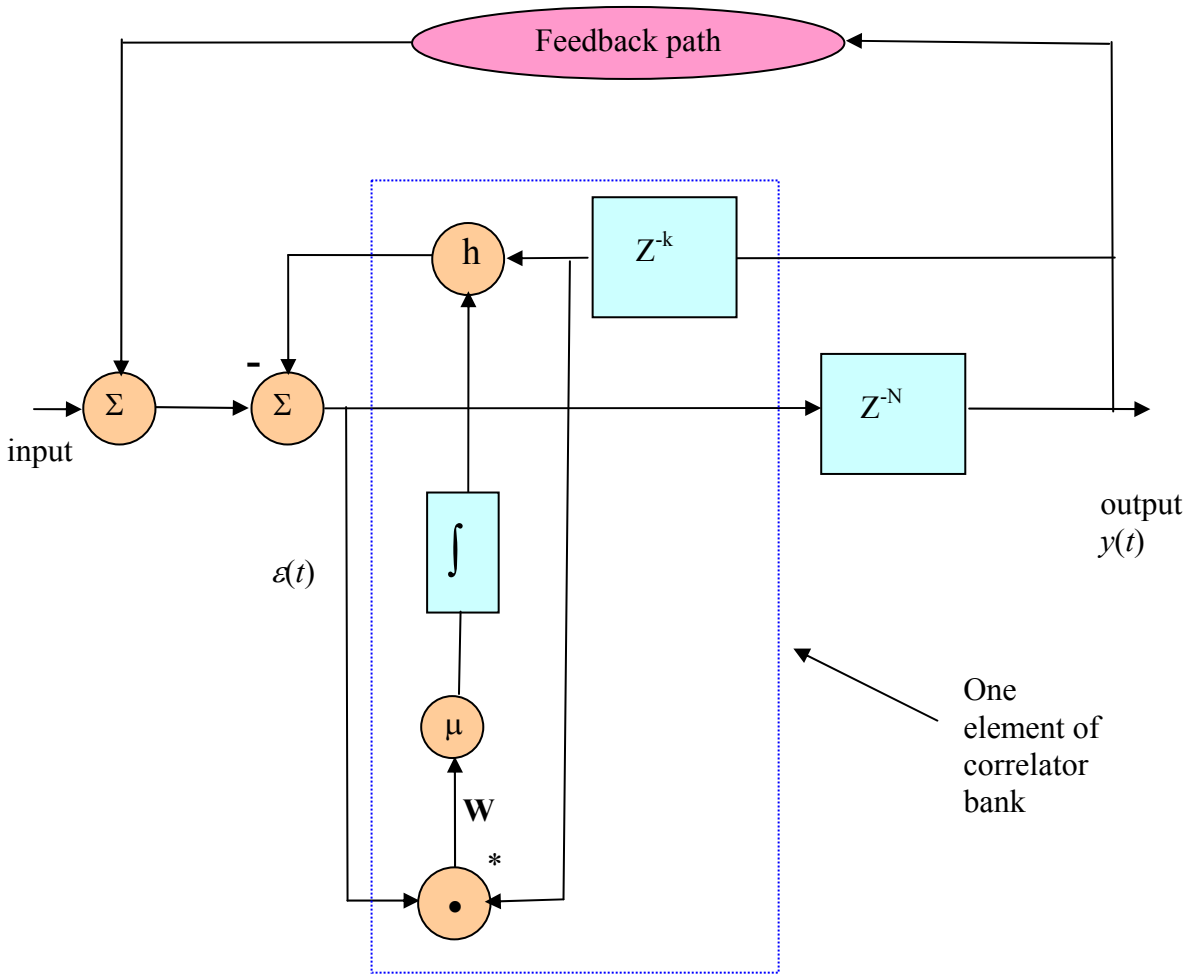

Figure 4: k^{th} tap of multitap architecture

Referring to figure 5 for nomenclature, we can express this condition (over the entire tap vector) as:

$$E\left\{\boldsymbol{\varepsilon}(t)\mathbf{y}^{\mathbf{H}}(t)\right\} = 0_{1xN} \quad \text{.....(5)}$$

where $\mathbf{y}(t)$ is the vector of the last N output values of the system (with indices 0 (current sample) to N-1).

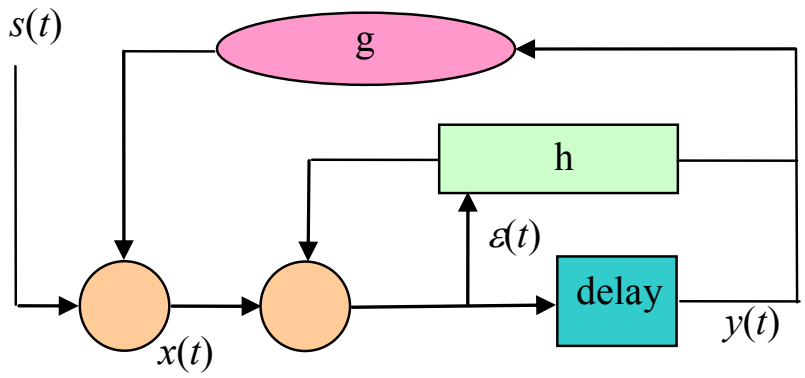

Figure 5:Signal nomenclature

Now

$$\varepsilon(t) = -\mathbf{h}^{\mathrm{T}}\mathbf{y}(t) + x(t) \quad \text{.....(6)}$$

where **h** is the adaptive filter tap vector, and

$$x(t) = \mathbf{g}^{\mathrm{T}}\mathbf{y}(t) + s(t) \quad \text{.....(7)}$$

where **g** is the unwanted feedback vector.

So

$$\varepsilon(t) = s(t) + \mathbf{g}^{\mathrm{T}}\mathbf{y}(t) - \mathbf{h}^{\mathrm{T}}\mathbf{y}(t) = s(t) + \left(\mathbf{g}^{\mathrm{T}} - \mathbf{h}^{\mathrm{T}}\right)\mathbf{y}(t) \quad \text{.....(8)}$$

combining eqs. (8) and (5) yields

$$E\left\{\left(s(t) + \left(\mathbf{g}^{\mathrm{T}} - \mathbf{h}^{\mathrm{T}}\right)\mathbf{y}(t)\right)\mathbf{y}^{\mathrm{H}}(t)\right\} = 0 \quad \text{.....(9)}$$

Let us define a tap difference vector as

$$\left(\mathbf{g}^{\mathrm{T}} - \mathbf{h}^{\mathrm{T}}\right) \equiv \mathbf{\Delta}^{\mathrm{T}} \quad \text{.....(10)}$$

So that (9) becomes, recognising $E\left\{\mathbf{y}\mathbf{y}^{\mathrm{H}}\right\} = \mathbf{R}_{\mathbf{yy}}$, the output signal *N*-by-*N* covariance matrix,

$$\mathbf{\Delta}^{\mathrm{T}}\mathbf{R}_{\mathbf{yy}} + E\left\{s(t)\mathbf{y}^{\mathrm{H}}(t)\right\} = 0 \quad \text{.....(11)}$$

The second term in (11) is an input-output cross-correlation vector which shall be written as $\mathbf{r}_{\mathbf{sy}}^{\mathbf{T}}$. So (11) can be re-written in a compact form as

$$\mathbf{\Delta}^{\mathbf{T}}\mathbf{R}_{\mathbf{yy}} + \mathbf{r}_{\mathbf{sy}}^{\mathbf{T}} = 0 \quad \text{.....(12)}$$

Hence, the value of the tap difference vector, which gives rise to the residual loop gain, is given by

$$\mathbf{\Delta}^{\mathbf{T}} = -\mathbf{r}_{\mathbf{sy}}^{\mathbf{T}}\mathbf{R}_{\mathbf{yy}}^{-1} \quad \text{.....(13)}$$

We see immediately that this vector is zero-valued if the input-output cross-correlation vector is zero-valued. So one solution to the problem is to increase the delay of the system sufficiently to render $\mathbf{r}_{\mathbf{sy}}^{\mathbf{T}}$ negligible, even for a non-white signal. This is undesirable in practice though, because minimising the overall system delay is beneficial in terms of guard interval erosion.

Fortunately, an alternative solution is available which takes the form of an intentional bias introduced into the correlator output. We shall discuss this below.

3.1.3.2 Applying the compensation

Let us introduce a fixed biasing vector $\mathbf{z}^{\mathbf{T}}$ to the output of our correlator bank which evaluates $E\left\{\varepsilon(t)\mathbf{y}^{\mathbf{H}}(t)\right\}$. This point for the k^{th} tap is marked **W** in figure 4. The condition expressed by equation (5) changes to the following

$$E\left\{\varepsilon(t)\mathbf{y}^{\mathbf{H}}(t)\right\} + \mathbf{z}^{\mathbf{T}} = 0 \quad \text{.....(14)}$$

Hence eq. (9) becomes

$$E\left\{\left(s(t) + \left(\mathbf{g}^{\mathbf{T}} - \mathbf{h}^{\mathbf{T}}\right)\mathbf{y}(t)\right)\mathbf{y}^{\mathbf{H}}(t)\right\} = -\mathbf{z}^{\mathbf{T}} \quad \text{.....(15)}$$

and so

$$\mathbf{\Delta}^{\mathbf{T}} = -(\mathbf{r}_{\mathbf{sy}}^{\mathbf{T}} + \mathbf{z}^{\mathbf{T}})\mathbf{R}_{\mathbf{yy}}^{-1} \quad \text{.....(16)}$$

Choosing

$$\mathbf{z} \equiv -\mathbf{r}_{\mathbf{sy}} \quad \text{.....(17)}$$

then sets $\mathbf{\Delta}^{\mathbf{T}}$ to zero as desired. This can be done since $\mathbf{r}_{\mathbf{sy}}$ is known a priori from the input and output signal characteristics.

This was the approach adopted in the BBC repeater, albeit with some further elaborations to cope with the responses of various filters present in the OCR architecture, and the consequences of having an external reference feedback node.

The simulated effect of the technique on the LMS system error $e(n)$ (i.e. the OCR output signal) is shown in figure 6.

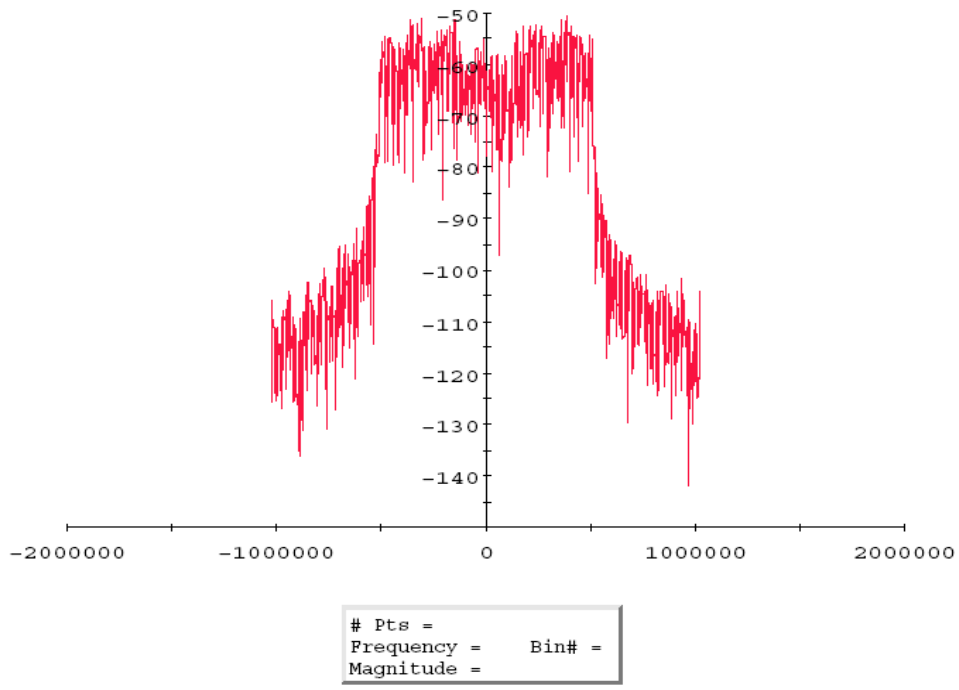

Figure 6: System error $\varepsilon(t)$ after correlation compensation

3.2 Power amplifier distortion

3.2.1 Manifestation

Referring back to figure 2, the power amplifier block will inevitably exhibit non-linear behaviour to some degree. With OFDM, the resulting spectrum from a third-order distortion process is illustrated below in figure 7:

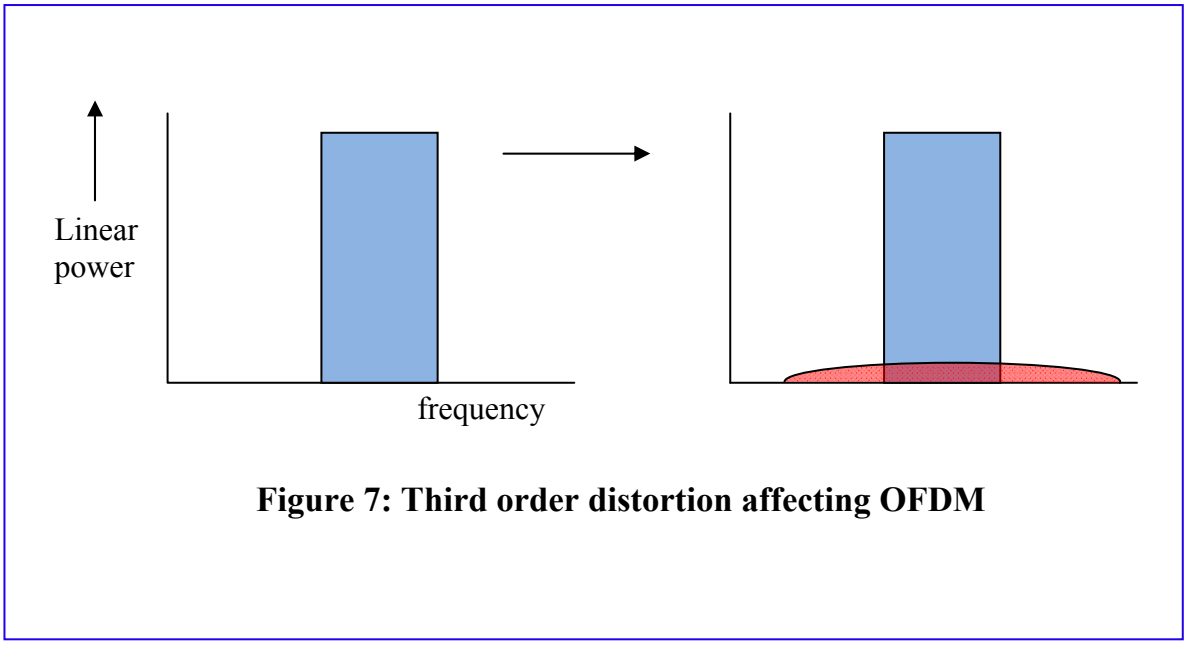

Figure 7: Third order distortion affecting OFDM

The distorted spectrum is made up of the sum of the original signal, and a term which essentially results from a double convolution of the signal PSD with itself, producing a triangular PSD after the first convolution and a softened inverted bowl-shape from the second convolution.

3.2.2 Impact on cancellation performance

The intermodulation phenomenon is well-understood by system designers, and the operating point of the amplifier would normally be chosen such that non-linear distortion introduces only a small overall loss of system margin. However, in an OCR context the intermodulation products can impact cancellation performance because when configured as figure 2, the cancelling filter, being linear, cannot synthesise the necessary higher-order terms of the power amplifier input signal to cancel the non-linear components of the coupled signal, in-band or out-of-band. As a result there is a limit on the available cancellation, and/or a demand that the amplifier is heavily 'backed off', i.e. run at a power well below its nominal rating, to minimise the intermodulation.

If neither of these options is acceptable, an alternative approach is to include the distortion terms in the reference signal, as we shall now describe.

3.2.3 Revised feedback tap position

The idea is quite straightforward in principle as the reference coupler position is simply moved to a point after the power amplifier, as shown in figure 8. In this way the reference signal contains the distortion products and cancellation performance is rendered almost independent of amplifier performance. However there is a price to pay: whereas in the architecture of figure 2 the reference signal is taken from a point within the DSP part of the system, in figure 8 it is taken from a point in the analogue domain. This necessitates an extra analogue to digital convertor, so in practice it may still be preferable to use the original tap position for low-cost systems of more limited performance, whilst reserving the external-tap option for high-power high-performance installations.

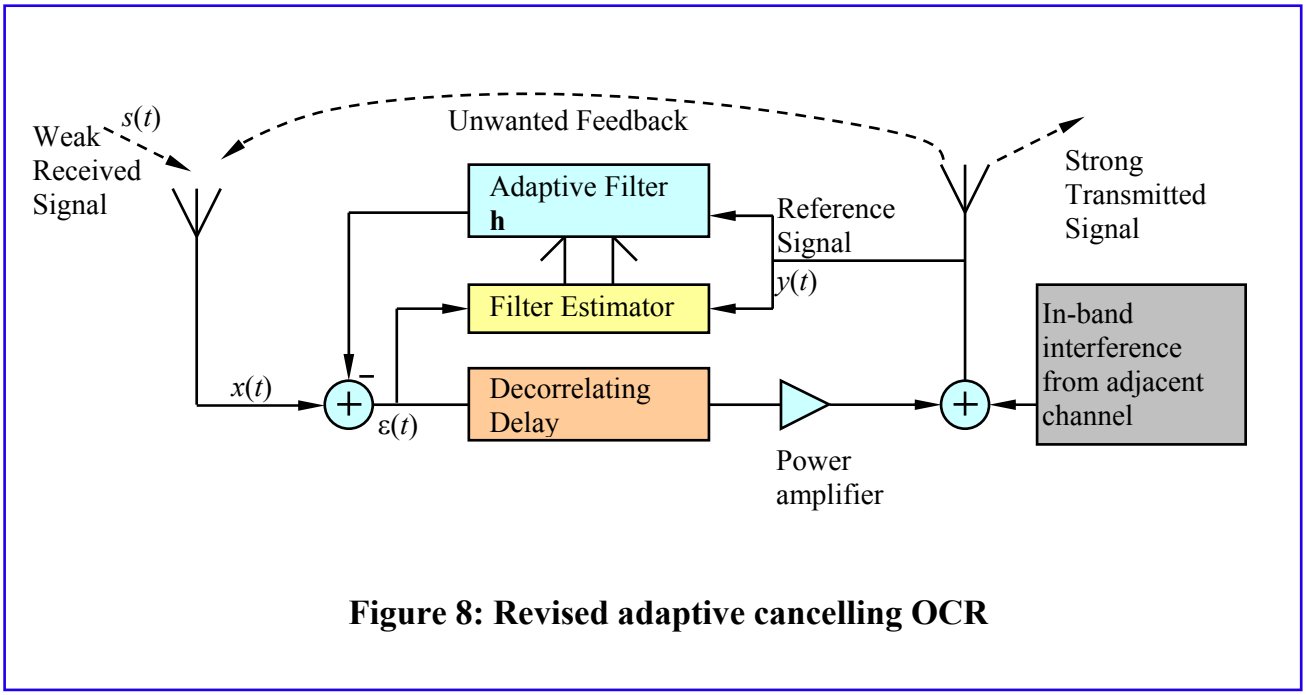

Figure 8: Revised adaptive cancelling OCR

An additional benefit of the post-power amplifier tap is alluded to in figure 8, namely the possibility of including out-of-band products from an adjacent transmission sharing the same antenna. This is very useful in 'hole-filling' applications where an OCR is to be deployed to restore reception in a 'hole' in the coverage area punched by a new adjacent transmitter, which is not co-sited with the

existing source of the 'victim' signal. A directional coupler fitted in the common antenna feed satisfies the requirement for the reference signal to contain both intermodulation from the host OCR amplifier and out-of-band products from the adjacent high-power amplifier. This arrangement mitigates against the target received signal being rendered unusable at the OCR input port by the adjacent transmission.

3.3 Post-echoes on the OCR input

3.3.1 Origin

In a practical deployment the input to the OCR is rarely made up of a single temporal component; the path from the source features impulse response components not only as a result of multipath propagation but also, where an SFN is in use, from other transmission nodes. In a lot of cases we will be faced with post-echoes, i.e. impulse response components later and smaller than the earliest, largest term. Pre-echoes can also occur and are more troublesome. These will be discussed later in this paper.

3.3.2 Post-echo reduction

If there are no pre-echoes, or they are very small, and certain other conditions are satisfied, an LMS echo canceller can be used to remove unwanted post-echo paths. To that end, additional digital filters with arbitrarily adjustable delays can be implemented in the OCR firmware. While observing the channel impulse response of the input and the re-transmitted signals, it is possible to target and eliminate echoes by introducing additional appropriately positioned taps in the FIR structure. Although these filters can add a considerable number of filter taps to the cancellation circuit, their noise contribution can be minimised by reducing their associated value of λ in equation 4. This is justified as these post-echo components vary much more slowly than the terms associated with the repeater feedback path.

It must be emphasised that although this method ensures that the guard interval will not be exceeded by the unwanted impulse response terms, the SNR of those OFDM carriers that are attenuated by mutual cancellation before they enter the OCR may still be poor. For this reason antenna installation plays a very important part in the design of the repeater system.

3.3.3 Example: performance in an SFN environment

As mentioned above, the complete removal of unwanted paths (SFN paths in this case) is possible only if certain conditions are met. If there are N SFN components, the sufficient, as well as the simplest, condition to apply is:

$$|a_0| > \sum_{n=1}^{N-1} |a_n| \quad \text{.....(18)}$$

where $|a_n|$ is the amplitude of the n-th path, with a_0 denoting the wanted parent signal. This is the stability condition for an IIR filter with N-1 coefficients of arbitrary phases. If it is not satisfied, the adaptive filter converges to an inappropriate but stable solution, that usually produces an output with a large number of delayed versions of the wanted signal. Such a signal, although often unusable for broadcast purposes due to the violation of the guard interval, is nonetheless spectrally flat, because the stable solution is an all-pass transformation of the unstable and thus unrealisable idealised filter. Therefore it is very important to monitor the channel impulse response of the re-transmitted signal, as its frequency response alone is not sufficient to assess its suitability for re-transmission.

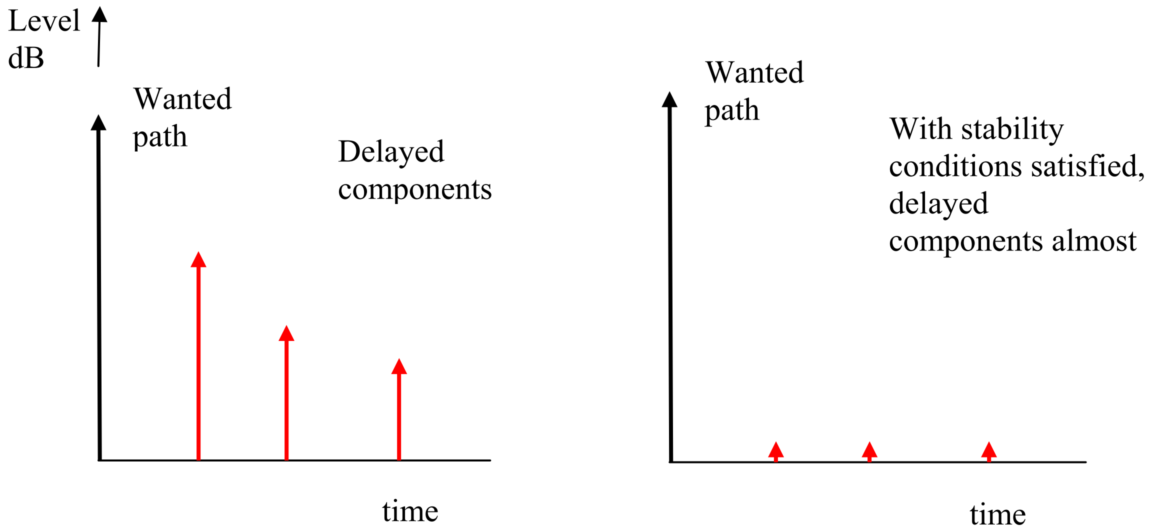

Figure 9: removal of post-echoes in channel impulse response

4 Future developments

4.1 Additional filtering of out-of-band tones

The correlation compensation technique described earlier in §3.1.3 can produce a considerable improvement in unwanted out-of-band tones, particularly in the presence of dynamic feedback channels. Further gains, however, can be obtained from a supplementary method, to be described in this section. Here we address the issue of spurious components by introducing an additional side-chain with two coefficients that supplements the LMS estimator and targets the out-of-band regions of the spectrum specifically and improves tone rejection and speed of convergence in this region.

The on-channel repeater recovers the signal $\varepsilon(t)$ from the corrupted input signal $x(t)$ and the reference signal $y(t)$ using an FIR filter of length N with coefficients $\mathbf{h}$ (as equation 1):

$$\varepsilon(t) = x(t) - \mathbf{y}(t)^{\mathrm{T}}\mathbf{h}(t) \qquad (19)$$

where $\mathbf{y}(t) = [y(t)\ \ y(t\text{-}1)\ \ y(t\text{-}2)\ \ \ldots\ \ y(t\text{-}N\text{+}1)]^{\mathrm{T}}$

The general on-channel repeater architecture is shown below. The input processing block of the reference signal is optional as $y(t)$ can be simply a delayed version of $\varepsilon(t)$.

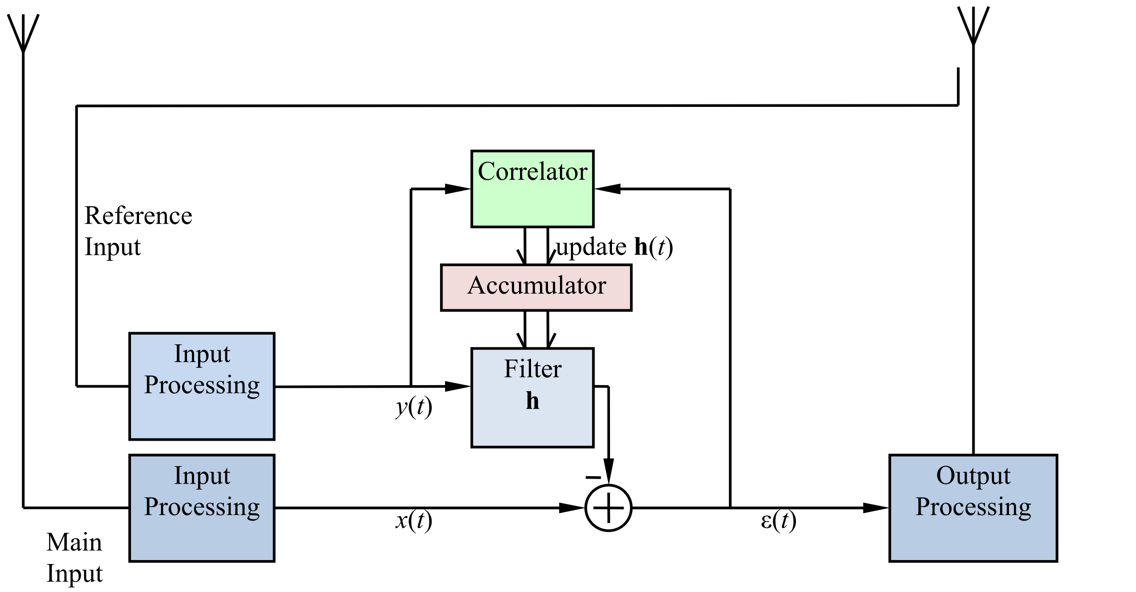

Figure 10: simplified OCR architecture

The conventional LMS algorithm determines the filter coefficients by successive updates as follows (as equation 4):

$$\mathbf{h}(t+1) = \mathbf{h}(t) + \mu\varepsilon(t)\mathbf{y}(t)^* \quad (20)$$

where μ controls the speed of convergence. However μ applies to the entire Nyquist bandwidth regardless of the spectrum of $y(t)$, resulting in slow convergence in the regions where $y(t)$ has low energy, that is the out-of-band regions. One way to deal with this problem is to boost those out-of-band frequencies in $y(t)$ by an additional filter. The result is a version of the transform-LMS algorithm [2]:

$$\varepsilon(t) = x(t) - [\mathbf{T}\mathbf{y}(t)]^{\mathrm{T}}\mathbf{h}(t) \quad (21)$$

with the transformation of the form:

$$\mathbf{T} = \mathbf{I} + \mathbf{w}\mathbf{w}^{\mathbf{H}} \quad (22)$$

where **w** is a frequency selective filter. The filter update becomes

$$\mathbf{h}(t+1) = \mathbf{h}(t) + \mu\varepsilon(t)[\mathbf{T}\mathbf{y}(t)]^* \quad (23)$$

This method does address the convergence problem but at the cost of introducing noise in the targeted out-of-band regions, proportional to the gain of **w**. Since the recovered signal ε(t) that is used to estimate filter taps contains the entire signal bandwidth, most of the filter update value for the region of interest is noise. Therefore it would be better to filter ε(t) as well so that the correlation is performed only within the targeted bandwidth. However, such filtering also introduces additional correlation into ε(t) and the filter update vector itself becomes correlated. The solution requires the insertion of another filter in the reference signal path as explained below.

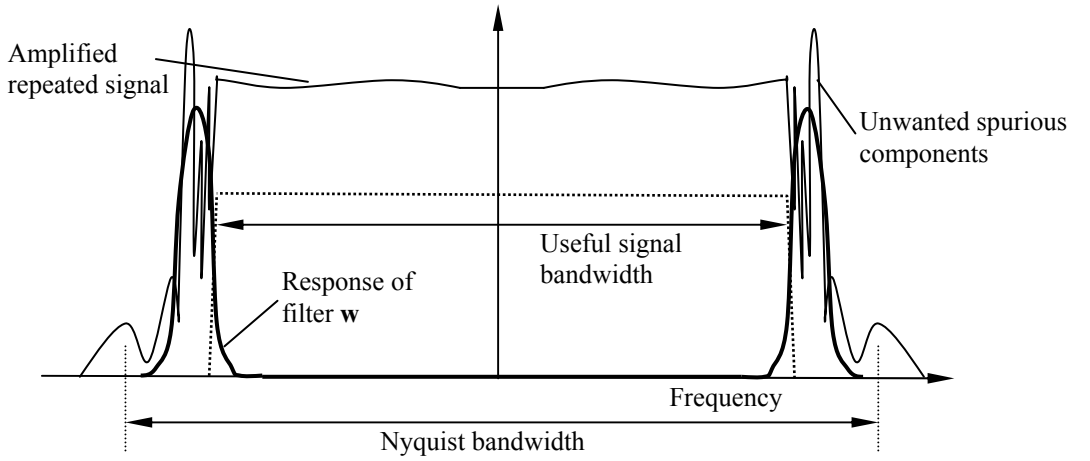

Figure 11: Unwanted Nyquist band-edge products

An additional adaptive filter $\mathbf{h}_w$ will be used to minimise the energy within a bandwidth defined by **w**, so that the recovered signal becomes:

$$\varepsilon(t) = \left[x(t) - \mathbf{y}(t)^{\mathrm{T}}\left[\mathbf{h}(t) + \mathbf{h}_w(t)\right]\right] \quad \text{.....(24)}$$

By defining the original LMS error as:

$$\varepsilon_o(t) = x(t) - \mathbf{y}(t)^{\mathrm{T}}\mathbf{h}(t) \quad \text{.....(25)}$$

the new error becomes:

$$\varepsilon(t) = \varepsilon_o(t) - \mathbf{y}(t)^{\mathrm{T}}\mathbf{h}_w(t) \quad \text{.....(26)}$$

In order to estimate $\mathbf{h}_w$ the filtered error signal is defined as follows:

$$\varepsilon_w(t) = \mathbf{w}^{\mathrm{T}}\left[\boldsymbol{\varepsilon}_o(t) - \mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_w(t)\right] \quad \text{.....(27)}$$

where $\mathbf{Y}(t)$ is a Hankel matrix (where $y_{k,n} = y_{k+1,n-1}$) of reference signal samples. It should be noted that the above is approximately equal to:

$$\varepsilon_w(t) \approx \mathbf{w}^{\mathrm{T}}\boldsymbol{\varepsilon}(t) = \sum_n w(n)\varepsilon(t-n) \quad \text{.....(28)}$$

if $\mathbf{h}_w$ changes slowly. The minimisation of $\varepsilon_w(t)$ is performed by finding the derivative of its energy with respect to the filter $\mathbf{h}_w$.

$$\frac{d}{d\mathbf{h}_w}\left|\varepsilon_w(t)\right|^2 = \left[\boldsymbol{\varepsilon}_o(t) - \mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_w(t)\right]^{\mathrm{H}}\mathbf{w}^*\mathbf{w}^{\mathrm{T}}\mathbf{Y}(t)^{\mathrm{T}} \quad \text{.....(29)}$$

Out of a number of possible forms of the update for $\mathbf{h}_w$ one looks particularly useful:

$$\mathbf{h}_w(t+1) = \mathbf{h}_w(t) + \mu\mathbf{w}^{\mathrm{T}}\left[\boldsymbol{\varepsilon}_o(t) - \mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_w(t)\right]\mathbf{Y}(t)^*\mathbf{w}^* \quad \ldots\ldots(30)$$

By assuming that $\mathbf{h}_w$ changes slowly, a further simplification of the update becomes possible:

$$\mathbf{h}_w(t+1) = \mathbf{h}_w(t) + \mu\left[\mathbf{w}^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{Y}(t)\mathbf{w}\right]^* \quad \ldots\ldots(31)$$

Thus the update is a correlation between a filtered error and a vector of filtered reference samples. The disadvantage of using coefficients that are applied separately from the original adaptive filter **h** lies in the fact that **h** may under some circumstances try to compensate for the change in response due to $\mathbf{h}_w$. The estimated coefficients may start to acquire ever larger values until numerical limits are reached. Instead, the tap update of $\mathbf{h}_w$ can be added to the update of the original filter **h**.

$$\mathbf{h}(t+1) = \mathbf{h}(t) + \mu\varepsilon(t)\mathbf{y}(t)^* + \mu_w\left[\mathbf{w}^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{Y}(t)\mathbf{w}\right]^* \quad \ldots\ldots(32)$$

The result is a combined adaptive filter with the desired characteristics of fast convergence and relatively low noise in the out-of-band regions. The example implementation shown below in figure 12 implements Equations (19) and (30).

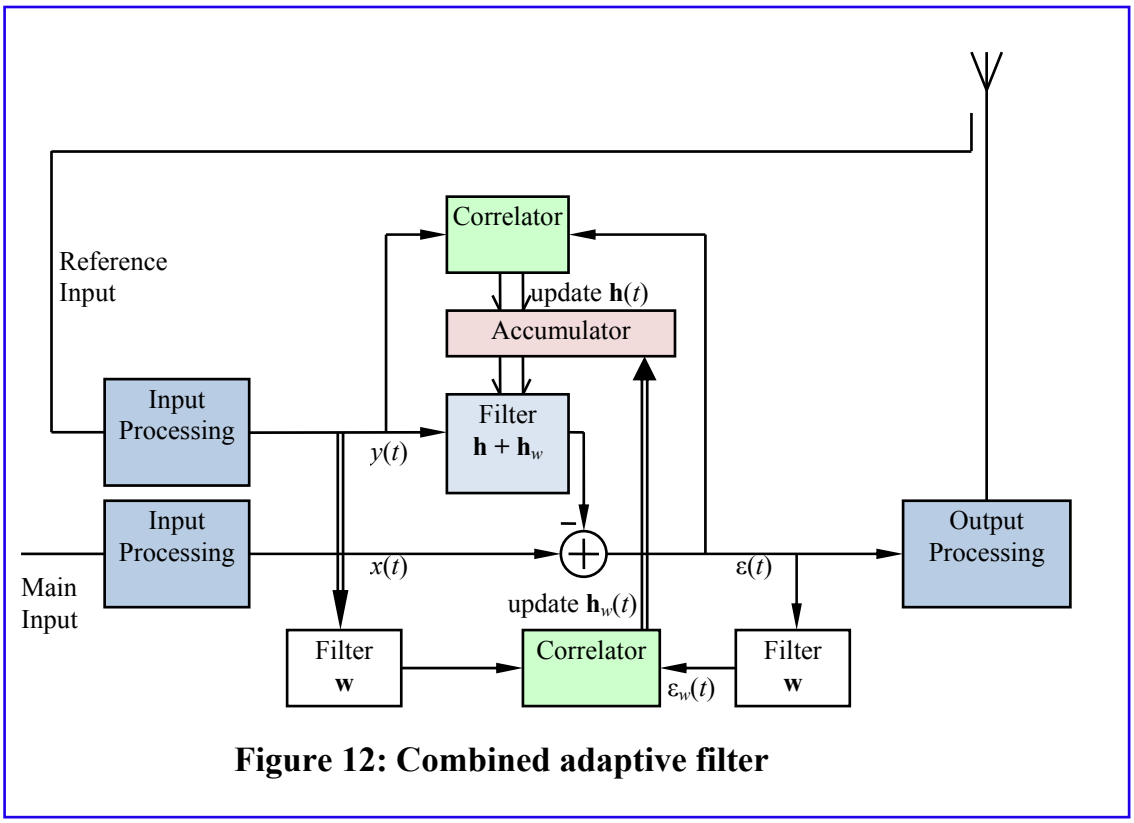

Figure 12: Combined adaptive filter

However, the additional adaptive filter $\mathbf{h}_w$ can be constrained to reduce noise even further. Instead of trying to estimate individual taps of $\mathbf{h}_w$, just two coefficients can be used, one for the lower and one for the upper out-of-band region. This is done by combining the above method of figure 12 with the transform-LMS algorithm. The error can be redefined as:

$$\varepsilon(t) = \left[x(t) - \mathbf{y}(t)^{\mathrm{T}}\mathbf{h}(t) - c_l(t)\mathbf{y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{y}(t)^{\mathrm{T}}\mathbf{h}_u\right] \quad \ldots\ldots(33)$$

and its filtered versions:

$$\varepsilon_l(t) = \mathbf{w}_l^{\mathrm{T}}\left[\boldsymbol{\varepsilon}_o(t) - c_l(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u\right] \quad \text{.....(34)}$$

$$\varepsilon_u(t) = \mathbf{w}_u^{\mathrm{T}}\left[\boldsymbol{\varepsilon}_o(t) - c_l(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u\right] \quad \text{.....(35)}$$

where $\mathbf{w}_l$ and $\mathbf{w}_u$ are filters that target the lower and the upper out-of-band region respectively. There are also two transformations, $\mathbf{h}_l$ for the lower out-of-band region and $\mathbf{h}_u$ for the upper. The objective is to derive two coefficients c_l and c_u that minimise the filtered error ε_c. The derivatives are:

$$\frac{d}{dc_l}\left|\varepsilon_l(t)\right|^2 = \left[\boldsymbol{\varepsilon}_o(t) - c_l(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u\right]^{\mathrm{H}}\mathbf{w}_l^{*}\mathbf{w}_l^{\mathrm{T}}\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l(t) \quad \text{.....(36)}$$

and

$$\frac{d}{dc_u}\left|\varepsilon_u(t)\right|^2 = \left[\boldsymbol{\varepsilon}_o(t) - c_l(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u\right]^{\mathrm{H}}\mathbf{w}_u^{*}\mathbf{w}_u^{\mathrm{T}}\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u(t) \quad \text{.....(37)}$$

As before, assuming that c_l and c_u change slowly, one can write:

$$\mathbf{w}_l^{\mathrm{T}}\left[\boldsymbol{\varepsilon}_o(t) - c_l(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u\right] \approx \mathbf{w}_l^{\mathrm{T}}\boldsymbol{\varepsilon}(t) \quad \text{.....(38)}$$

and

$$\mathbf{w}_u^{\mathrm{T}}\left[\boldsymbol{\varepsilon}_o(t) - c_l(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_l - c_u(t)\mathbf{Y}(t)^{\mathrm{T}}\mathbf{h}_u\right] \approx \mathbf{w}_u^{\mathrm{T}}\boldsymbol{\varepsilon}(t) \quad \text{.....(39)}$$

The coefficient updates can be simplified:

$$c_l(t+1) = c_l(t) + \mu_w\left[\mathbf{w}_l^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{h}_l^{\mathrm{T}}\mathbf{Y}(t)\mathbf{w}_l\right]^{*} \quad \text{.....(40)}$$

$$c_u(t+1) = c_u(t) + \mu_w\left[\mathbf{w}_u^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{h}_u^{\mathrm{T}}\mathbf{Y}(t)\mathbf{w}_u\right]^{*} \quad \text{.....(41)}$$

This results in a block diagram as shown below in figure 13 which implements Equations (33), (40) and (41).

It can be seen that using $\mathbf{h}_l$ and $\mathbf{h}_u$ applied just to to $\mathbf{Y}$ is equivalent to transform LMS where $\mathbf{h}_l$ and $\mathbf{h}_u$ define the transforms.

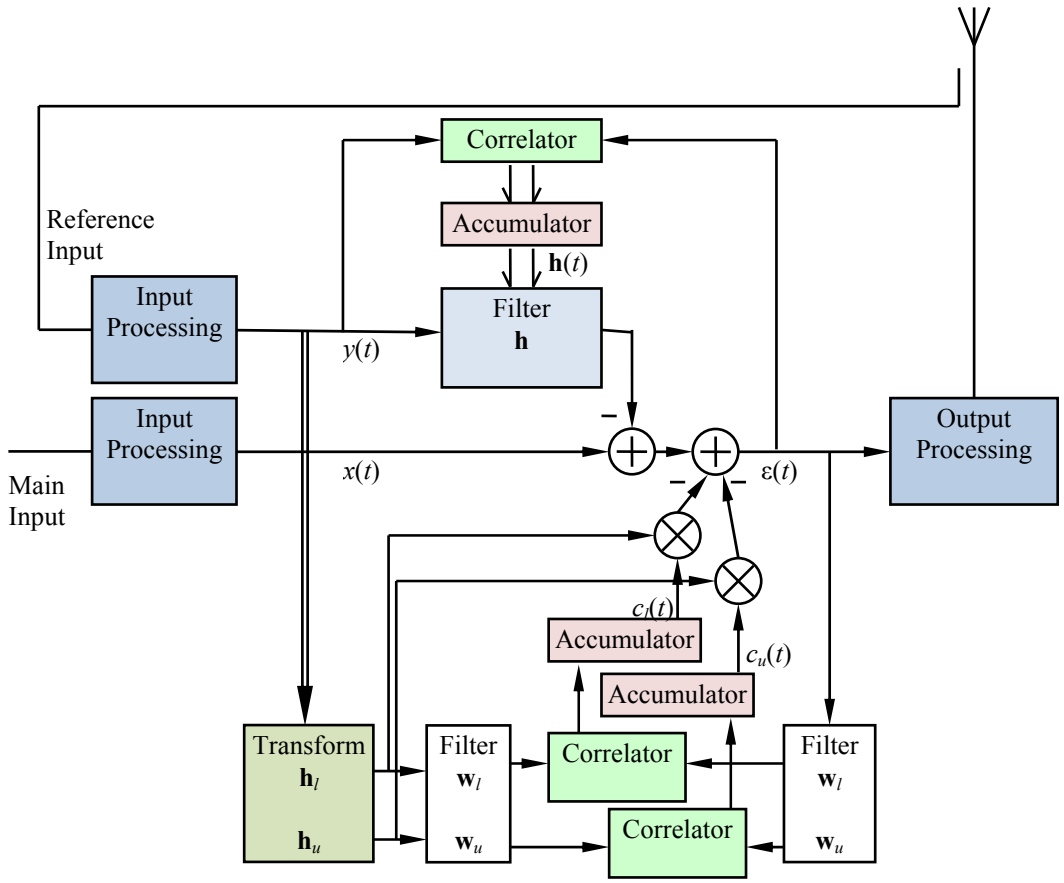

Figure 13: Simplification with scalar coefficients *c*

Again, in order to eliminate the possibility of divergence between $\mathbf{h}$, c_l and c_u the accumulators that hold the computed c_l and c_u can be combined with the accumulator for $\mathbf{h}$ as follows:

$$\mathbf{h}(t+1) = \mathbf{h}(t) + \mu\varepsilon(t)\mathbf{y}_r(t)^* + \mu_w\left[\mathbf{w}_l^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{h}_l^{\mathrm{T}}\mathbf{Y}_r(t)\mathbf{w}_l\right]^*\mathbf{h}_l + \mu_w\left[\mathbf{w}_u^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{h}_u^{\mathrm{T}}\mathbf{Y}_r(t)\mathbf{w}_u\right]^*\mathbf{h}_u \quad \text{.....(42)}$$

This method is more expensive than using c_l and c_u explicitly as it requires additional vector multiplications by $\mathbf{h}_l$ and $\mathbf{h}_u$. The example implementation based on Equations (19) and (42) is shown in figure 14.

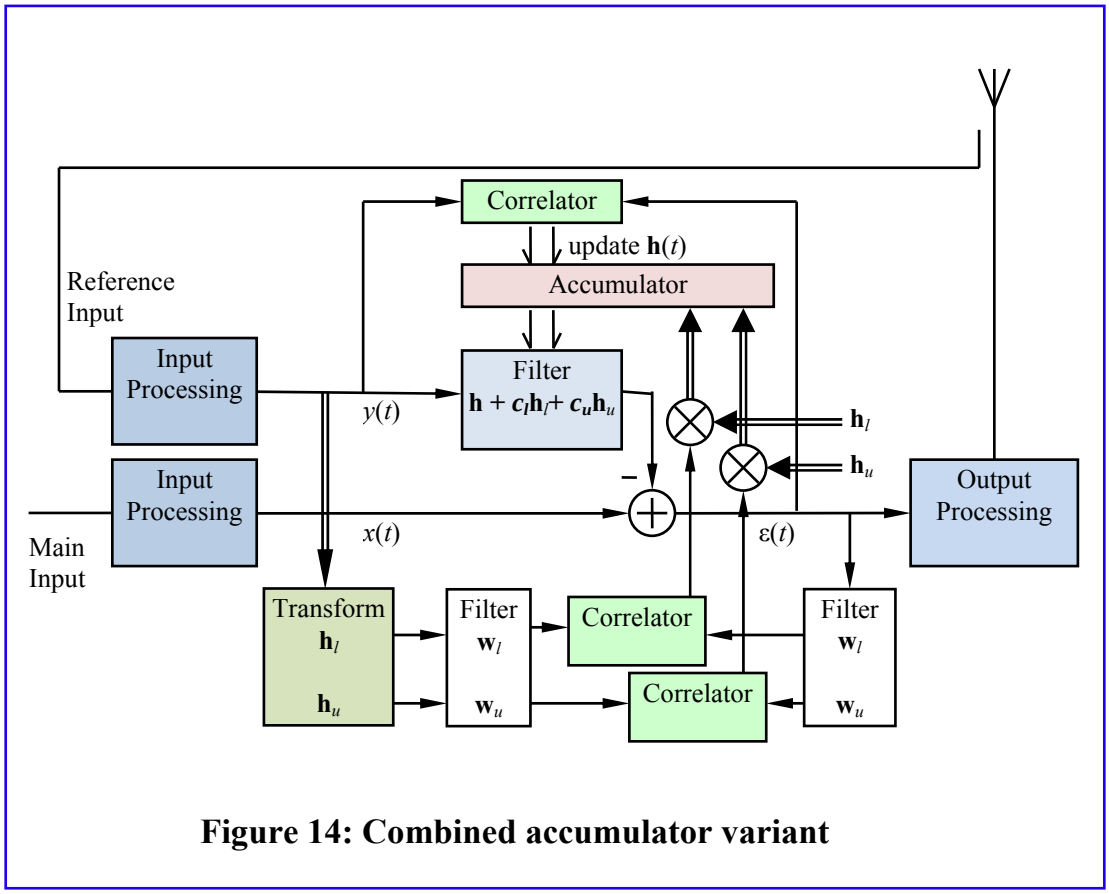

Figure 14: Combined accumulator variant

If the out-of-band regions are symmetrical:

$$\mathbf{h}_u = \mathbf{h}_l^* \quad \text{and} \quad \mathbf{w}_u = \mathbf{w}_l^* \qquad \text{.....(43)}$$

This means that the number of multiplications required for the two coefficient updates as well as the error signal extraction is the same as that needed for only one coefficient. It is suggested that the filter values are chosen as follows:

$$\mathbf{w}_l = \mathbf{h}_l \quad \text{and} \quad \mathbf{w}_u = \mathbf{h}_u \qquad \text{.....(44)}$$

and

$$w_{ln} = a_n \exp\left(-\frac{j\pi f_l n}{N f_s}\right) \qquad \text{.....(45)}$$

$$w_{un} = w_{ln}^* \qquad \text{.....(46)}$$

where

f_l – centre frequency of the upper out-of-band region targeted by the filter
f_s – sampling frequency
a_n – coefficients of a window function, for example the Hamming window

The final expression is:

$$\mathbf{h}(t+1) = \mathbf{h}(t) + \mu\varepsilon(t)\mathbf{y}_r(t)^* + \mu_w\left[\mathbf{w}_l^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{w}_l^{\mathrm{T}}\mathbf{Y}_r(t)\mathbf{w}_l\right]^*\mathbf{w}_l + \mu_w\left[\mathbf{w}_u^{\mathrm{T}}\boldsymbol{\varepsilon}(t)\right]\left[\mathbf{w}_u^{\mathrm{T}}\mathbf{Y}_r(t)\mathbf{w}_u\right]^*\mathbf{w}_u \qquad \text{.....(47)}$$

4.2 Pre-echoes

4.2.1 The pre-echo problem

Pre-echoes are likely to arise in situations where an OCR is also a useful tool, e.g. in SFNs. A transmitter local to the OCR site may not be the source of the strongest signal, so it may be desirable to use a more distant, higher power source as the OCR input. This is a typical scenario where a pre-echo may arise. The remainder of this section discusses this problem and suggests a possible solution based on multiple receive antennas for the OCR.

Suppose a source signal $s(t)$ to be recovered in a receiver suffers linear distortion which can be represented in a finite impulse response (FIR) filter:

$$x(t) = \sum_{\tau=0}^{(M-1)T_S} h(\tau)s(t-\tau) \quad \text{(48)}$$

Here $x(t)$ is the received signal, $h(\tau)$ is the channel impulse response, T_s is the sampling interval and M the number of taps.

In the z-domain, the convolution of (48) can be expressed as a product

$$X(z) = H(z)S(z) \quad \text{(49)}$$

If we normalise $h(0)$ to be unity, then (49) can be written

$$X(z) = \left(1 + \widetilde{H}(z)\right)S(z) \quad \text{(50)}$$

Where $\widetilde{H}(z)$ represents the response of the remaining terms of the filter.

Now the general form of an IIR equaliser is

$$W(z) = \frac{1}{1+E(z)} \quad \text{(51)}$$

So choosing $E(z) = \widetilde{H}(z)$ allows exact equalisation of the FIR channel response (49) assuming, of course, that the equaliser $\frac{1}{1+\widetilde{H}(z)}$ is stable.

Consider first the case of a single post echo, which by our definition of post-echo requires the coefficient $|h_1| < 1$. The IIR equaliser is then simply

$$W(z) = \frac{1}{1+h_1 z^{-1}} \quad \text{(52)}$$

The requirement for stability is that the pole of $W(z)$ lies inside the unit circle on the z-plane, which is satisfied in this case.

But now consider a pre-echo, with the main path $|h_1| > 1$. The pole of $W(z)$ now lies outside the unit circle, and the equaliser is not stable. If there are many coefficients in the impulse response, and it

starts with a pre-echo, a stable structure may theoretically exist but may become unstable if the coefficients are slightly perturbed, as may happen in an RF path. An example would be the stable response shown below as equation (53) which equalises a pre-echo but relies on the term in z^{-2} maintaining its magnitude and phase to prevent one of the poles drifting outside the unit circle.

$$W(z) = \frac{1}{\left(1+1.8z^{-1}+0.81z^{-2}\right)} = \frac{1}{\left(1+0.9z^{-1}\right)\left(1+0.9z^{-1}\right)} \quad \text{.....(53)}$$

4.2.2 Summary of OCR behaviour with pre-echoes

The IIR equalisation in the current OCR is primarily intended to provide cancellation of unwanted coupling of the transmitted signal back into the input. Within the temporal limits imposed by its cancellation widow, however, input signal post-echoes are also cancelled. Input pre-echoes in general cannot be equalised by a stable IIR filter and so the OCR cannot be expected to deal with these. What happens in practice depends on whether the pre-echo to second term separation is within the cancellation temporal window. If a single pre-echo needing an unstable equaliser is within the window, the steady state solution the OCR adopts has the same frequency response as the 'ideal' unstable equaliser, but a different phase response. The net result is a flat output spectrum, but containing many impulse response components.

To illustrate this, if the required unstable response is

$$W_U(z) = \frac{1}{1+h_1 z^{-1}} \quad \text{.....(54)}$$

With $|h_1| > 1$, the OCR settles in steady state to

$$W_{OCR}(z) = \frac{1}{h_1\left(1+\frac{1}{h_1}z^{-1}\right)} \quad \text{.....(55)}$$

which still provides a flat output spectrum because if z is on the unit circle the following holds:

$$\left|1+h_1 z^{-1}\right| = \left|h_1\left(1+\frac{1}{h_1}z^{-1}\right)\right| \quad \text{.....(56)}$$

Equation (56) also shows *why* the OCR converges to that value; as follows: the square of each side is the power spectral density (PSD) of two signals. One is the pre-echoed path excited by a white input, the other a scaled time-reversed post-echoed path. The PSDs are the same, and hence so are the autocorrelation functions[1]. The OCR taps stop changing (i.e. steady-state is achieved) when the autocorrelation of the output signal is zero, which (55) satisfies and is stable.

4.2.3 Twin-antenna reception

If we are restricted to a single receive antenna, then our only spatial degree of freedom is to adjust the antenna beam orientation to obtain the best compromise between wanted and pre-echo levels, which in general will come from somewhat differing directions. In some instances this may be

[1] Recall that the PSD is the Fourier transform of the autocorrelation function.

sufficient. Better results may be obtained, however, if a pair of antennas is used to obtain, for instance, higher gain in the direction of the wanted source whilst attempting to minimise signal from the early source. Once established, the optimum antenna arrangement can be permanently maintained by mechanical means.

Allowing for a pair of antennas in fact opens up a more advanced alternative, in the form of 'null-steering'. This technique aims to automate the process of finding and maintaining a low response to the pre-echo whilst maintaining a good response to the wanted input, and will now be described.

4.2.4 Null steering linear equaliser

4.2.4.1 Channel model

We shall first consider a wanted signal and a single unwanted interferer which arrive at the antenna array from different directions. This leads to a model, valid for a single impulse response term flat-fading case[2], in the form of a 2-by-2 matrix. This is shown below in figure 15.

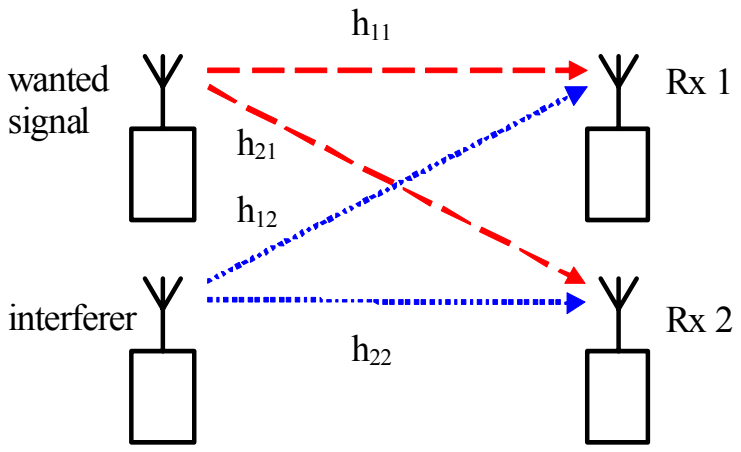

Figure 15: Interference model

We can write the channel matrix as **H** as follows

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \quad \text{.....(57)}$$

The matrix elements are complex scalars, accounting for the channel magnitude and phase.

The received signal, including noise, is given by

$$\mathbf{r} = \mathbf{Hs} + \mathbf{n} \quad \text{.....(58)}$$

[2] i.e. i.e. all terms in **H** have are characterised by a co-timed single impulse response component

Where **r** is the received vector and **s** is the transmitted vector comprising signal s_1 and pre-echo s_2:

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \quad \text{.....(59)}$$

The vector **n** is a noise vector, assumed to be made up of two independent complex Gaussian terms each of variance σ^2.

$$\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \quad \text{.....(60)}$$

4.2.4.2 Channel estimation – single temporal term

We shall first consider the case where the interferer is a pre-echo characterised by a single impulse response term. We can hence still make use of the flat-fading channel model as described, as the temporal offset of the pre-echo is included in its definition as an interferer. We will of course see frequency-selective fading where the pre-echo and the main signal combine.

The receiver must of course gain knowledge of the channel, i.e. obtain the terms of **H**. Ordinarily the on-channel repeater does not require a signal demodulator for normal operation, since the architecture is not that of a regenerative 'demod-remod' system. It is in fact simply an amplifier with associated bandpass filtering plus a parallel echo canceller. The 'demod-remod' approach, which has the benefit of signal regeneration, is not compatible with low-delay operation as required for SFNs.

However the need for channel estimation in the present context dictates that a partial demodulator is necessary to extract the impulse response components seen by each receiver. This demodulator can be in a side-chain and will not therefore introduce delay into the main signal path. Once an impulse response is obtained, a suitable algorithm must be used to identify both the main 'wanted' path and any pre-echo to be rejected. This could be done in a simple way by setting the largest component as the reference and looking for an earlier component above a certain threshold[3], which is then deemed to be the target pre-echo, as shown in figure 16 below. Dealing with the more practical case of multiple components of the pre-echo or main path will be dealt with in §4.2.4.6.

In addition a more sophisticated parsing of the impulse response will also be necessary, which might result in advice to the user that the signal is unsuitable and antenna re-alignment is therefore recommended. An indication of the effectiveness of the equalisation could also be presented to the user, in the form of a single figure of merit or an estimate of the net equalised response.

[3] In practise a smooth weighting function performing a similar function to a threshold is likely to be used to avoid the effects of a hard discontinuity.

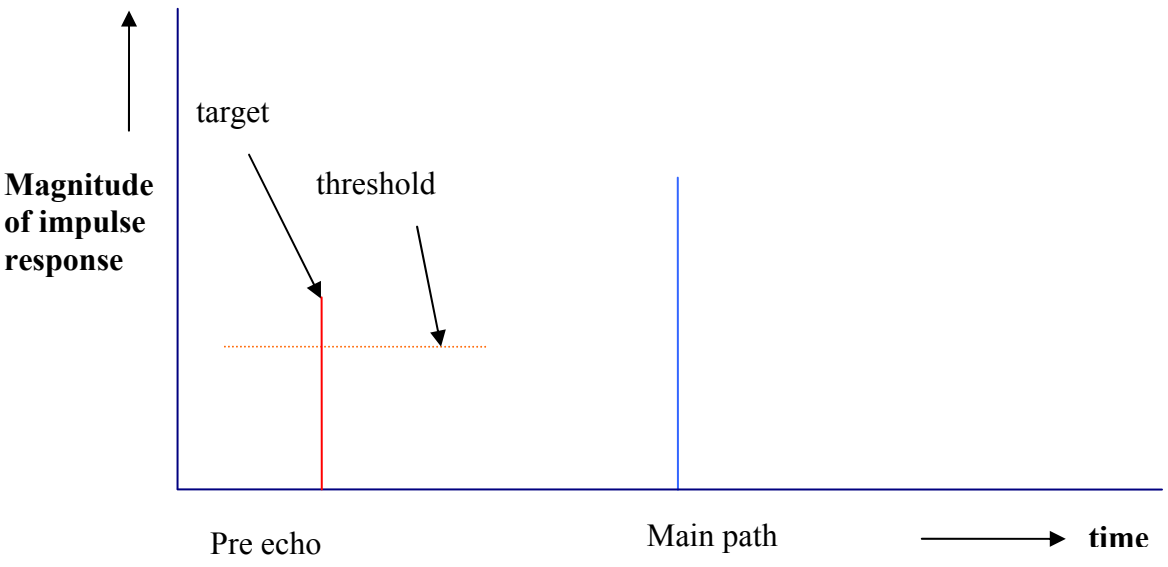

Figure 16: Pre-echo analysis

4.2.4.3 Receiver architecture and zero-forcing equalisation

As described above we shall regard the twin-antenna input of the OCR input as a complex vector of length two, where each element is the complex scalar derived from the associated antenna. We wish to design a vector equaliser which derives a wanted signal with a reduced pre-echo as one element of its output vector. Now given the matrix form of the channel, we can see immediately that a simple pre-echo rejection algorithm is simply to invert the matrix to recover the signal and pre-echo separately. In fact if we only want the signal, we need only use one row of the inverse matrix. We call this row the *steering vector*, which, as its name implies, works by steering a null towards the interfering pre-echo. Mathematically, the steering vector is orthogonal to the interfering vector in the complex space $\mathbb{C}^2$. After this processing, the scalar 'de-echoed' signal passes to the feedback-path cancelling part of the OCR for the usual processing.

This front-end concept is shown in figure 17 below. The received signals, which both contain the pre-echo, are shown as r_1 and r_2. The separated output signal, in the absence of noise, is s_1. The pre-echo has been rejected by the steering vector.

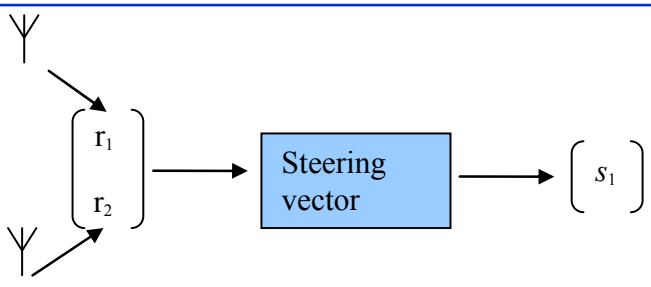

Figure 17: Vector processing OCR front end

This simple type of equalisation is more generally known as *zero-forcing* (ZF) equalisation, since it forces a deep null onto the interferer. However the disadvantage is that in the presence of noise, the output can suffer from poor signal-to-noise ratio because the ZF nulling vector may cause large noise gain. The degree to which the system suffers from this depends on the characteristics of the channel matrix. If the matrix is poorly conditioned, by which we mean it has a high ratio of singular values[4], then significant noise peaking results. In the extreme case, where the matrix is close to being singular, the noise peaking tends to infinity.

We can visualise this by considering the two requirements that the ZF nulling vector must satisfy; firstly it must be orthogonal to the received interference vector and secondly the scalar product with the basis vector[5] of the wanted transmission must be unity. Figure 18 below illustrates this for the simplified case of real signals and channel:

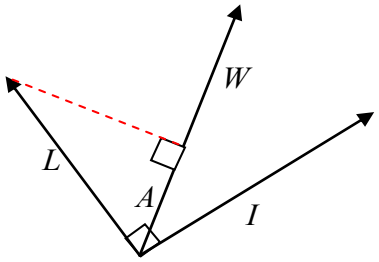

Figure 18: Origins of noise peaking

In the figure, L is the steering vector, I is the interfering vector, and W is in the direction of the basis vector of the wanted signal. To reject the interferer, L must be orthogonal to I, as shown, and the projection of L onto W, i.e. the length of A, must be unity. Clearly as the angle between W and I diminishes, L must increase in magnitude to maintain this condition. The variance of the output noise is $|L|^2$ times that at the receiver input, which therefore increases accordingly.

4.2.4.4 Steering vector for MMSE equalisation

A better approach in most circumstances is to find a compromise between nulling the interference and raising the noise. One such compromise is to minimise the mean-square error of the output; the steering vector which has this characteristic is equal to one row of the MMSE inverse of the channel matrix.

It can be shown that the MMSE inverse **W** of a matrix **H** is given by

$$\mathbf{W} = (v\mathbf{I} + \mathbf{H}^{\mathbf{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathbf{H}} \quad \text{.....(61)}$$

[4] Singular values make up the diagonal matrix **Σ** in the Singular Value Decomposition of a matrix M: $\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathbf{T}}$, where **U** and **V** are unitary matrices.

[5] The wanted transmission has basis vector $\begin{pmatrix} h_{11} \\ h_{21} \end{pmatrix}$ in the current example. That is to say, the received signal due to wanted signal s is $\begin{pmatrix} h_{11} \\ h_{21} \end{pmatrix} s$.

where ν (in the context of **H** being a channel matrix) is the ratio of the added noise power at the receiver to the transmitted power. If the desired component of the transmission **s** is s_1, then the first row of **W**, which we shall denote by $\mathbf{W}_1$, is the required MMSE steering vector which produces the optimum combination of null-steering (towards s_2) and the avoidance of noise.

Hence, formally, the recovered estimate of s_1 ($\hat{s}_1$ say) is given by

$$\hat{s}_1 = \mathbf{W}_1\mathbf{r} = \left((\nu\mathbf{I} + \mathbf{H}^{\mathbf{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathbf{H}}\right)_1\mathbf{r} \quad \text{.....(62)}$$

Of course, if we want to obtain the interfering term itself for some reason, e.g. monitoring purposes we can use the analogous expression

$$\hat{s}_2 = \mathbf{W}_2\mathbf{r} = \left((\nu\mathbf{I} + \mathbf{H}^{\mathbf{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathbf{H}}\right)_2\mathbf{r} \quad \text{.....(63)}$$

4.2.4.5 Multi-tap vector MMSE equalisation

So far we have only considered a channel characteristic where a sufficient channel description is a matrix **H** with elements which are simply complex variables. In practice, however, we are likely to face selectively fading channels with multiple impulse response components and so our null-steering equaliser should be able to deal with these. One possible approach would be to use filtering to divide the selectively fading channel into a number of contiguous quasi-flat fading channels and assign each its own matrix $\mathbf{H}_k$. The drawbacks of this architecture are principally the complexity of the necessary filter banks and the signal delay and degradation which will inevitably occur. A compromise would have to be struck between the number of filterbank sub-channels and the ability to deal with strongly selective fading. So although conceptually straightforward, a more elegant approach is desirable in practice and time-domain equalisation can offer such a possibility.

As well as the IIR equaliser discussed above, an FIR equaliser structure can be used for correcting the response of a single-ended input in the time domain, where the input has suffered linear distortion in the form of an FIR filter (i.e. as in equation (48)). This structure is inherently preferable because of its unconditional stability. The correction is necessarily imperfect due to its FIR nature but in practice good results can be obtained with quite modest equaliser tap lengths in many circumstances.

Suppose we define a vector of the most recent *N* input samples to the transmit/receive system as follows:

$$\mathbf{y}(k) = \left[y(k) \quad y(k-1) \quad y(k-2)\ldots\ldots\ldots\ldots y(k-N+1)\right]^{\mathbf{T}} \quad \text{.....(64)}$$

Now if we choose *N* to be *M*+*L*-1 where *M* is the *source* FIR length (as eq. (48)), and *L* is the proposed FIR equaliser length, we can write the equalised output of the system as

$$z(k) = \mathbf{w}^{\mathbf{H}}\mathbf{H}\mathbf{y}(k) \quad \text{.....(65)}$$

Where **w** is the vector of equaliser taps and **H** is a Toeplitz matrix with source impulse response taps on the upper diagonals (see below).
For example this expands in a system with *L*=*M*=4 to

$$z(k) = \begin{pmatrix} w_0^* & w_1^* & w_2^* & w_3^* \end{pmatrix} \begin{pmatrix} h_0 & h_1 & h_2 & h_3 & 0 & 0 & 0 \\ 0 & h_0 & h_1 & h_2 & h_3 & 0 & 0 \\ 0 & 0 & h_0 & h_1 & h_2 & h_3 & 0 \\ 0 & 0 & 0 & h_0 & h_1 & h_2 & h_3 \end{pmatrix} \begin{pmatrix} y(k) \\ y(k-1) \\ y(k-2) \\ y(k-3) \\ y(k-4) \\ y(k-5) \\ y(k-6) \end{pmatrix} \quad \text{.....(66)}$$

Now we need to work out the coefficients of the optimum equaliser $\mathbf{w}_{opt}$. It can be shown that for MMSE equalisation the following expression holds:

$$\mathbf{w}_{opt}^{\mathbf{H}} = \mathbf{b}^{\mathbf{H}} \left(\mathbf{R}_{\mathbf{yy}}^{-1} + \mathbf{H}^{\mathbf{H}} \mathbf{R}_{\mathbf{nn}}^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^{\mathbf{H}} \mathbf{R}_{\mathbf{nn}}^{-1} \quad \text{.....(67)}$$

Where $\mathbf{R_{nn}}$ is the L x L covariance matrix of the added noise and $\mathbf{R_{yy}}$ is the N x N covariance matrix of the source signal[6]. $\mathbf{b}^{\mathbf{H}} = \begin{bmatrix} 0 & 0 & \ldots\ldots & 1 & \ldots\ldots & 0 & 0 \end{bmatrix}$ is a vector of length N containing a single 'one' and the remainder zeros; the position of the 'one' sets the overall delay through the source and equaliser FIR sections to which we are optimising the equaliser design. This is typically set to about half the total number of taps in the source and equaliser. The formal approach is to find

$$\arg\min_{\mathbf{b}} \left\{ \left(\mathbf{b}^{\mathbf{H}} - \mathbf{w}_{opt}^{\mathbf{H}}(\mathbf{b}) \mathbf{H} \right) \mathbf{R}_{\mathbf{yy}} \left(\mathbf{b} - \mathbf{H}^{\mathbf{H}} \mathbf{w}_{opt}(\mathbf{b}) \right) + \mathbf{w}_{opt}^{\mathbf{H}}(\mathbf{b}) \mathbf{R}_{\mathbf{nn}} \mathbf{w}_{opt}(\mathbf{b}) \right\} \quad \text{.....(68)}$$

which minimises the expectation of the equalisation error against the value of **b**, subject to **b** having the form defined above.

So far we have seen how to deal with the single-tap vector and multi-tap scalar cases of the MMSE equalisation problem. But if we are to deal with selective fading in our vector receiver, we need to combine the two capabilities and build an FIR equaliser whose taps are matrices, of dimension 2 x 2 in the present context, to act upon our 2 x 1 signal vectors. An L-tap vector equaliser therefore has a 1 x L block matrix structure **W** with overall dimension 2 x 2L. This variable is analogous to variable **w** above, and the vector **b** becomes an array **B** containing a single identity matrix in an appropriate position.

It can be shown that for a given choice of reference delay matrix **B** the optimum value of **W** is given by

$$\mathbf{W}_{opt} = \mathbf{B} \left(\mathbf{R}_{\mathbf{YY}}^{-1} + \mathbf{H}^{\mathbf{H}} \mathbf{R}_{\mathbf{NN}}^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^{\mathbf{H}} \mathbf{R}_{\mathbf{NN}}^{-1} \quad \text{.....(69)}$$

where **H** is a block Toeplitz matrix, defined in the appendix, and the structure of **B** is as follows

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & \ldots & \ldots & 1 & 0 & \ldots & \ldots & 0 & 0 \\ 0 & 0 & \ldots & \ldots & 0 & 1 & \ldots & \ldots & 0 & 0 \end{bmatrix}.$$

$\mathbf{R}_{\mathbf{YY}}$ and $\mathbf{R}_{\mathbf{NN}}$ are signal and noise covariance matrices as defined in the appendix. The matrix **B** is established by finding

[6] i.e. the covariance matrix of the vectors formed of the last L and N samples of the input noise and signal respectively.

$$\arg\min_{\mathbf{B}} tr\left\{\left(\mathbf{B}-\mathbf{W}_{opt}(\mathbf{B})\mathbf{H}\right)\mathbf{R}_{\mathbf{YY}}\left(\mathbf{B}^{\mathbf{H}}-\mathbf{H}^{\mathbf{H}}\mathbf{W}_{opt}^{\mathbf{H}}(\mathbf{B})\right)+\mathbf{W}_{opt}(\mathbf{B})\mathbf{R}_{\mathbf{NN}}\mathbf{W}_{opt}^{\mathbf{H}}(\mathbf{B})\right\}\ldots\ldots\ldots\ldots\ldots\ldots(70)$$

with **B** constrained to be in the form described above.

4.2.4.6 Channel estimation – multiple temporal terms

To deal with selective fading of the pre-echo, we must identify more than one component of the associated impulse response. In the example shown below, the pre-echo consists of four components in total each separated by one sampling period. However, as previously in the flat fading scenario, a level threshold has been imposed and only three elements are above this. But we are now also concerned with the time extent of the impulse response components, and for a given equaliser tap length must impose a limit in time in addition to the level threshold. In the current example, one component exceeds our temporal equalisation window and is discarded. So two components remain to contribute to our estimate of **H**, an L x (L+1) block matrix in this case[7]. Clearly the resulting equalisation will not be perfect, since we have ignored some components of the pre-echo.

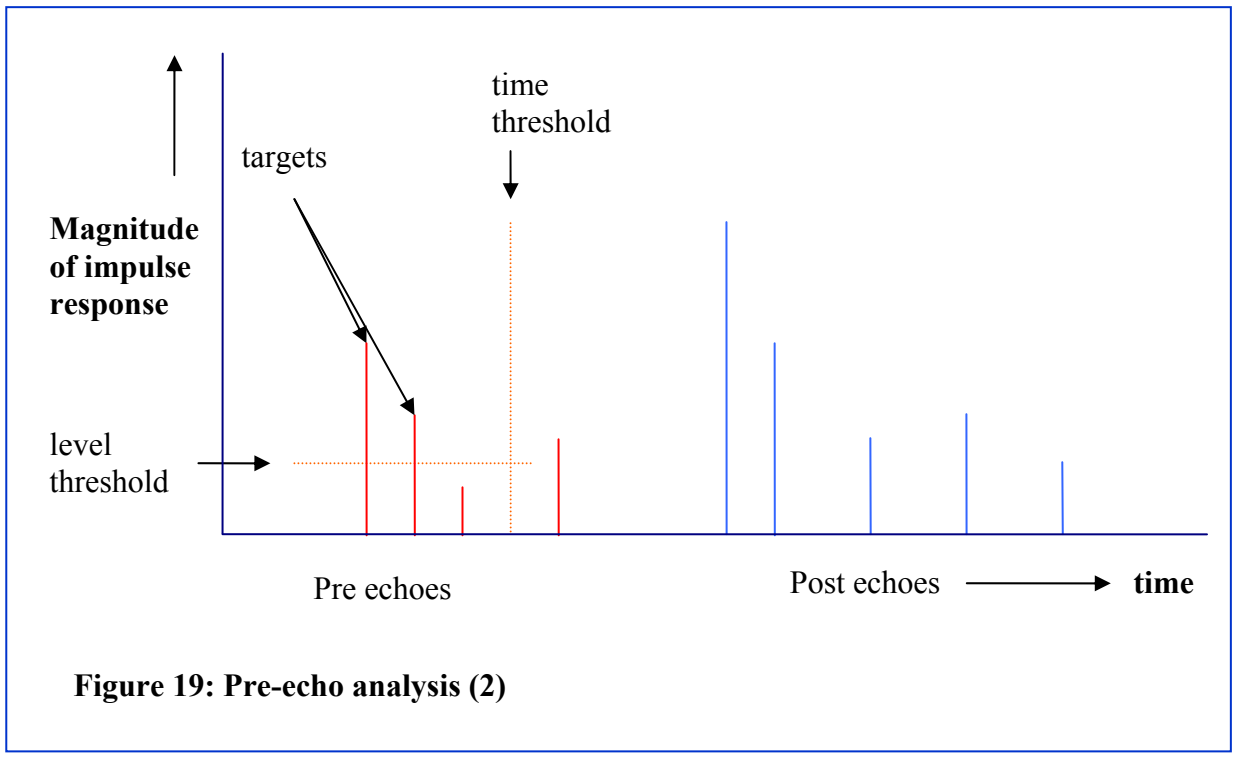

Figure 19: Pre-echo analysis (2)

5 Hardware realisation

5.1 Block diagram

The block diagram of the physical realisation of BBC R&D's on-channel repeater, for DAB is shown in figure 20.

The wanted DAB signal in Band III is fed into the main input of the repeater unit. The first module is a radio frequency (RF) filter designed primarily to reject an image frequency before downconversion. A mixer is then used to convert to a 20MHz intermediate frequency (IF) where a

[7] i.e. 2L x 2(L+1) scalar elements

high order LC IF filter is applied to remove most of the energy of the adjacent channels. The received signal may of course be buried under a feedback signal up to 30dB greater in level.

In addition, the unfiltered remnants of the adjacent channels must also be accommodated in order to avoid overload. Consequently, the design of the mixer and IF filter must ensure the lowest distortion and the highest dynamic range within the wanted bandwidth before the signal is digitised. The signal is sampled by a 16-bit analogue-to-digital converter running at 80MHz so that the quantisation noise is spread over the entire Nyquist bandwidth of 40MHz resulting in the effective signal to noise ratio of 93dB with respect to full scale within the useful 1536kHz bandwidth.

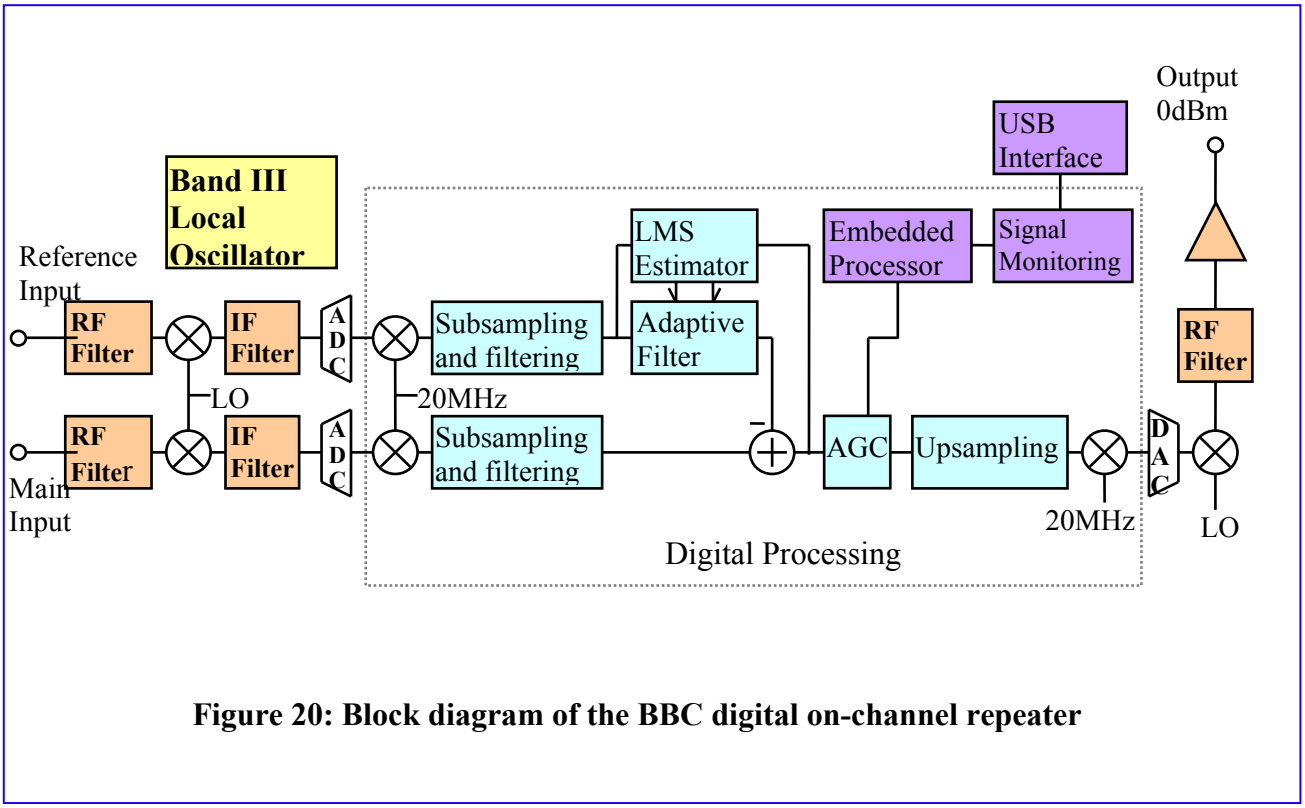

Figure 20: Block diagram of the BBC digital on-channel repeater

Once in the digital domain, the signal is converted to complex baseband, filtered and subsampled to 1.66MHz. There follows a feedback cancellation block, which renders the signal suitable for re-broadcasting. The output level is stabilised by the AGC module, which is appropriately gated to avoid gain changes during the DAB null symbol. The signal is then upsampled, mixed up to an IF of 20MHz and converted back to analogue in a 16-bit converter. All digital processing is performed in a single Field Programmable Gate Array (FPGA) device, which executes approximately 3 billion fixed-point multiplications per second. A further up-conversion to Band III is followed by an image rejection band-pass filter and a 0dBm driver stage. This signal can then be routed to an external RF power amplifier and RF channel filter, and combined with the main transmitter output carrying an adjacent channel or directly applied to the transmitting antenna.

The reference signal obtained from a coupler in the antenna feed is processed in an identical manner to the main signal, as described above, up to and including the digital down-conversion and filtering. The next stage is the adaptive filter which produces the cancellation signal subtracted from the main input. Filter taps are calculated in an LMS estimator module.

The signal levels are monitored by an embedded microprocessor in the FPGA that controls the behaviour of the AGC circuitry. If an abnormal condition is detected, such as an overload or signal disconnection, the output of the repeater is reduced or muted.

6 Conclusions

The on-channel repeater is a most useful device in broadcast network engineering where spectral efficiency is a key concern. This paper has presented an overview of the operation of the BBC R&D OCR together with some operational issues and their solution. Also presented were methods which have not yet been subject to practical test and are awaiting realisation, field trial and eventual inclusion in realisable systems. It is hoped that in the not too distant future the full benefit of all the techniques described here, and some which have not been described in order to limit the paper to a reasonable length, are available to broadcast network planners through the availability of a new generation of high-performance commercial on-channel repeaters.

7 Acknowledgements

The authors would like to thank the BBC for permission to publish this paper.

8 References

[1]. Marple, S. L. Jr., Digital Spectral Analysis with Applications, Prentice Hall,1987, ISBN 0-13-214149-3.

[2]. D. F. Marshall, W. K. Jenkins and J.J. Murphy "The use of Orthogonal Transforms for Improving Performance of Adaptive Filters" *IEEE Trans. Circuits Sys*, vol. 36 no. 4 Apr 1989, pp. 474-484