1 Assignment: Echo Cancellation

We observe a primary signal $d(n)$ (`signal.asc`) that consists of a local signal (`local.asc`, *theoretically unknown*) mixed with an interference. This interference consists of an echo, which is a hybrid-circuit-filtered version of a remote signal $u(n)$ (`remota.asc`). The sampling rate of these signals is 8 kHz. We want to design a **normalized LMS filter** canceling such an echo.

1.1 Draw the block diagram of the system

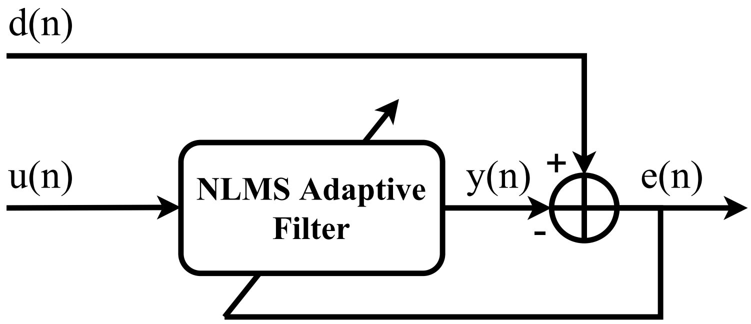


Figure 1: Block diagram of the system.

In the above diagram, $d(n) = x(n) + v(n)$, where $x(n)$ represents the target signal (`local.asc`) and, $v(n)$, the interference. Therefore, $y(n)$ is an estimate of the interference $v(n)$ from $u(n)$, namely, $y(n) = \hat{v}(n)$. The error signal, which is also the output of interest, is $e(n) = d(n) - y(n) = x(n) + [v(n) - \hat{v}(n)]$, that is, an estimate $\hat{x}(n)$ of the target signal.

1.2 What is the *a priori* signal-to-noise ratio (SNR) of the observed primary signal $d(n)$?

The *a priori* SNR of $d(n)$ is

$$10 \log_{10} \left(\frac{\sum_{n=1}^{N} x^2(n)}{\sum_{n=1}^{N} \left(d(n) - x(n)\right)^2} \right) = 7.74 \text{ dB}, \tag{1}$$

where $N = 15970$ is the total number of samples of each of the signals.

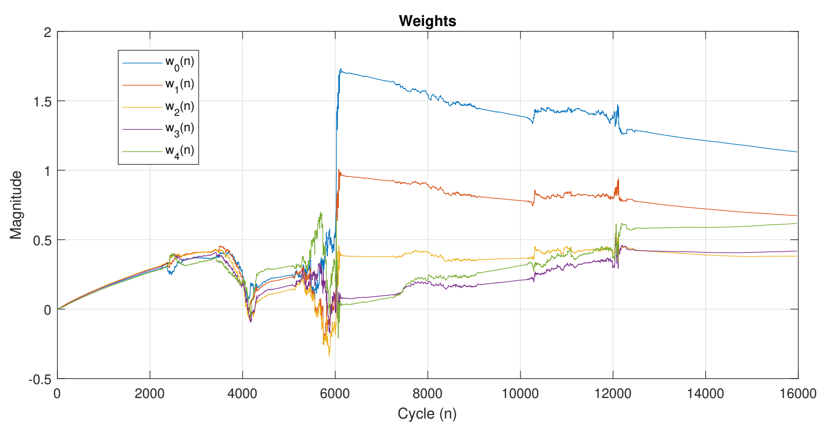
1.3 Implement an echo canceler using NLMS, where $\hat{\mathbf{w}}(0) = \mathbf{0}$. Search for the step-size parameter $\hat{\tilde{\mu}}$ and filter order $\hat{M}$ maximizing the SNR of the signal after echo cancellation. The search ranges are $\tilde{\mu} \in \{0.0001,\ 0.0002,\ 0.0004,\ ...,\ 0.0512\}$ and $M \in \{1,\ 2,\ 3,\ ...,\ 10\}$. What are the SNR, $\hat{\tilde{\mu}}$ and $\hat{M}$ values?

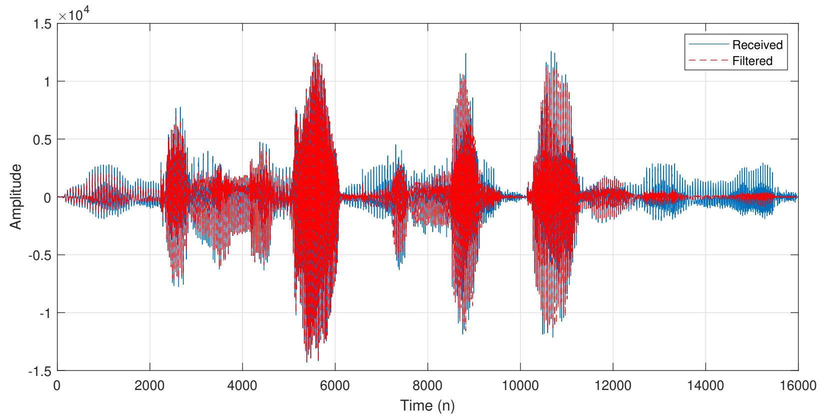
The SNR is

$$10\log_{10}\left(\frac{\sum_{n=1}^{N} x^2(n)}{\sum_{n=1}^{N}\left(e(n)-x(n)\right)^2}\right) = 15.01 \text{ dB}. \qquad (2)$$

In addition, $\hat{\tilde{\mu}} = 0.0004$ and $\hat{M} = 5$.

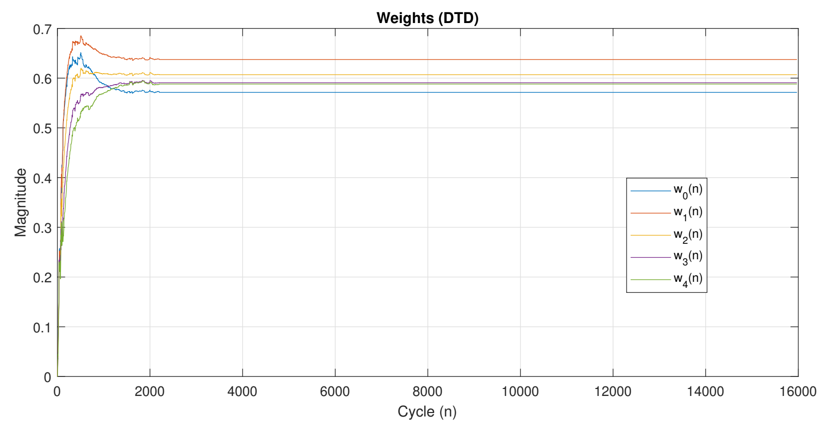
1.4 Plot the time evolution of the filter weights for the best SNR configuration. Also, plot both $d(n)$ and the signal after echo cancellation. Very briefly, comment the results

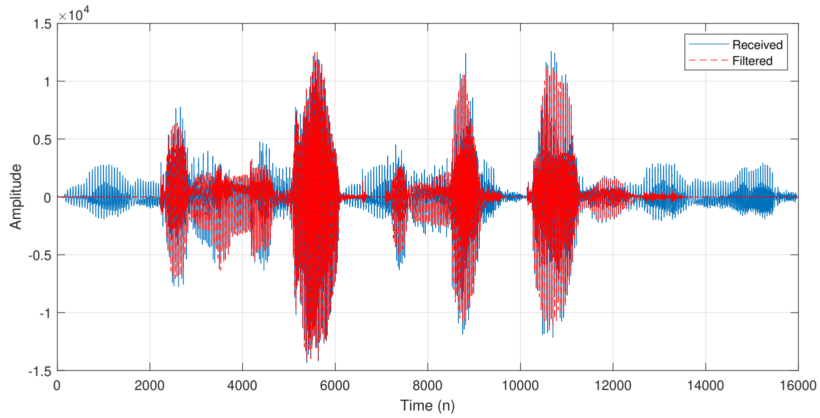




1.5 Repeat *3)* and *4)* by assuming a *double talk detector* (DTD) spots local speech from sample 2200, in such a manner that the filter adapts up to that sample only. Compare the results with those obtained in *3)* and *4)*. What kind of filter is obtained upon sample 2200?

In this case, the SNR is 47.80 dB. Furthermore, $\hat{\hat{\mu}} = 0.0128$ and $\hat{M} = 5$.





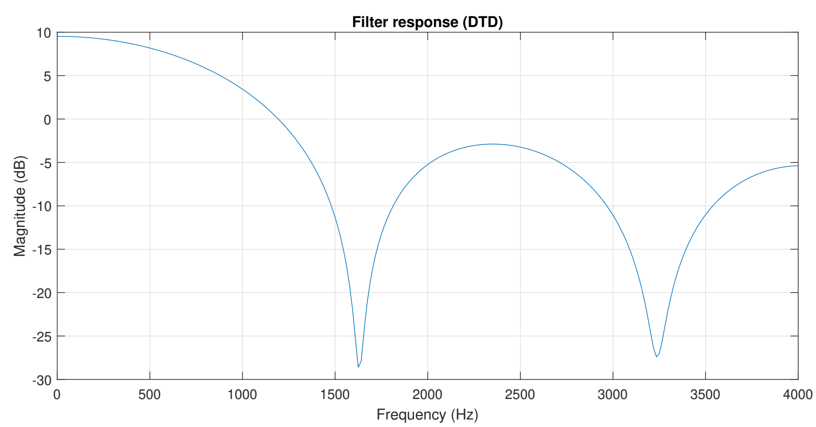
To figure out what kind of filter we obtain upon sample 2200, we can proceed as follows. We can simply apply the Z-transform to the FIR filter response in the time domain, i.e., to

$$y(n) = \sum_{i=0}^{M-1} w_i u(n-i), \tag{3}$$

in order to get its magnitude response in the frequency domain as

$$20\log_{10}|H(z)| = 20\log_{10}\left|\frac{Y(z)}{U(z)}\right| = 20\log_{10}\left|\sum_{i=0}^{M-1} w_i z^{-i}\right|. \tag{4}$$

The magnitude response in the frequency domain of the FIR filter is plotted down below:



We can see that a **comb filter** is obtained.