Spectral Shaping Using Wiener Filter and Reduction of MMSE using NLMS

Ashish Rawat
Electrical Engineering
The University of Texas at Dallas
Dallas, USA
ayr170130@utdallas.edu

Dr. Issa M. S. Panahi
Department of Electrical and Computer Engineering
University of Texas at Dallas
Dallas, USA
imp015000@utdallas.edu

***Abstract*— This project asks for filtering out an estimate of the desired signal from the noisy speech signal. Desired signal and the speech signal are different. System which is identified using wiener filter has an optimal impulse response calculated along with the length of the filter. This filtering operation is performed on whole noisy speech signal taken together and on frames taken at once. Performing this helps us understand the nature of speech as a non-stationary signal and gaussian noise as a wide sense stationary. Here an analysis of the independent Wiener Filter and independent NLMS Filtering and Wiener Filter followed by NLMS is performed. Observations for a all these cases are made in the paper to follw.**

Keywords*— *Wiener filtering; Adaptive filtering; satationary process; MMSE.

I. PROBLEM DISCRIPTION

The problem given involves a noisy speech signal of 3.12325 seconds with 49972 sample called x(n). Another signal provided is the desired signal d(n) which is also 3.12325 seconds with 49972 samples. The sampling frequency for both the signals is 16000Hz.

The desired signal is given as a weighted summation of cosines of 3 different frequencies with specific phase shifts.

Task:

- Find an optimal impulse response of the required FIR filter which estimates the system.
- Evaluate how well the obtained signal y(n) estimates the desired signal d(n).
- Reduce the MMSE using adaptive filtering.

Issues:

Speech is a non-stationary signal which must be verified. Speech is mixed with the gaussian white noise and the signal desired signal obtained from the signal is different from speech.

Given : $x(n)= s(n)+w(n)$

Assume $x(n)= d(n)+ (s(n)-d(n)) +w(n)$

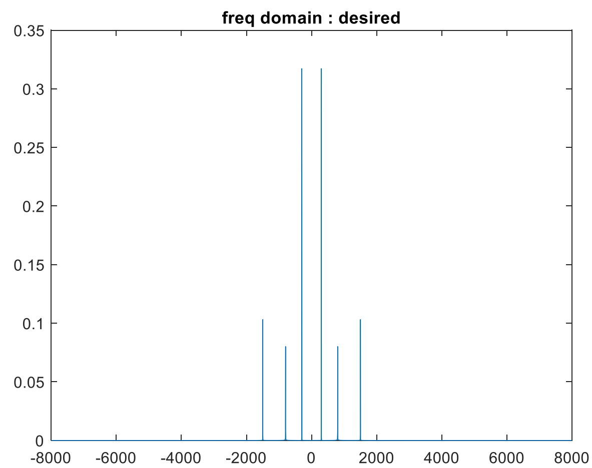
Under this assumption we shall proceed with taking just one frame.

For verifying that the speech is a non-stationary process, we shall consider the whole noisy speech at once.

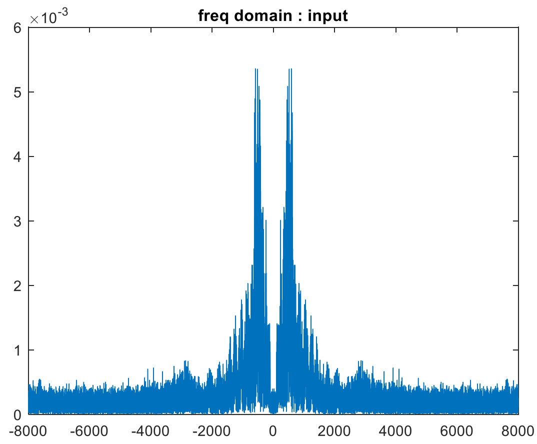
Several questions to solve:

- What should be the order of the filter which we are going to predict?
- Which system model would be the best and why?
- How should the framing of the signal be done?

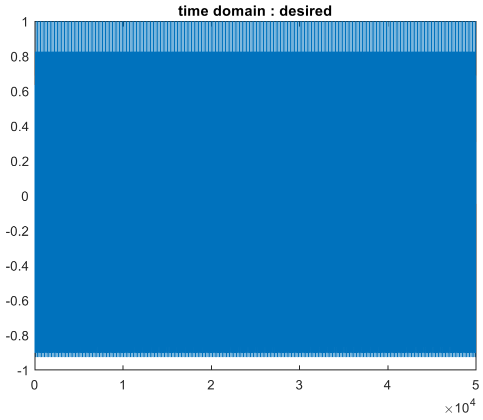
We shall explore these questions and solve them in the following paragraphs.



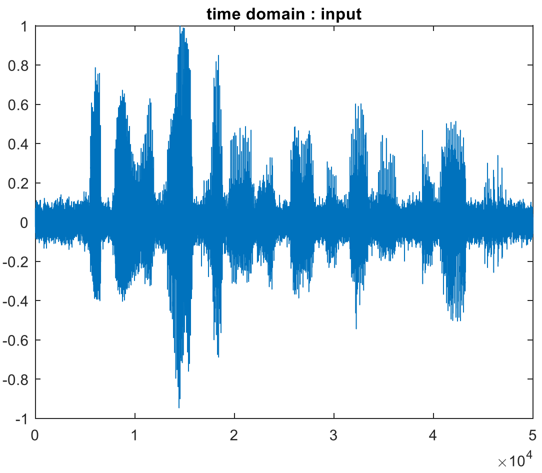
Fig(1). Frequency domain representation of desired signal



Fig(2). Frequency domain representation of the input signal.



Fig(3). Time domain representation of the desired signal



Fig(4). Time domain representation of the input signal

II. Wiener Filter Theory and Observations

To design a system that filters out additive noise and estimates the desired system which produces the desired signal for the given input. Wiener filter is used for signals that are WSS. So the basic assumption is that x(n), d(n) and w(n) are all wide sense stationary with zero mean.

$y(n)= \sum h(k)\, x(n-k)$; k=0 to M-1 ; M is order of the filter.

The Mean square error is calculated by using:

$$MSE= \varepsilon = E[(|e(n)|^2)] \quad \text{Eq. 1}$$
$$= E|\, d(n)- \sum h(k)\, x(n-k)\, |$$

Note:
$e(n)=d(n)-y(n)$
$y(n) = h_{opt}(n) * x(n)$

Steps to find the Optimum Filter Response:
1.Find the autocorrelation of the input sequence : r_{xx}
2. Find the cross correlation of the input signal and the desired signal: r_{dx}
3. Form a Toeplitz matrix from r_{xx}.
4. Use the Wieiner Hopff equation to calculate the $h_{opt}(n)$

Wiener Hopf equation: $\Gamma_{xx}\, h(n) = P$; P is the matrix of r_{xy}

$$h_{opt} = (\Gamma_{xx})^{-1}\, r_{xy} \quad \text{Eq. 2}$$

From the following equation we understand that the optimum filter is unique.

$$J = (h- h_{opt})^T \,.\, \Gamma_{xx} \,.\, (h- h_{opt}) + J_{opt} \quad \text{Eq. 3}$$

J is error for impulse response a filter. whereas J_{opt} is the error for optimum impulse response filter.

Minimum mean square error:

$$\varepsilon_{min} = \sigma^2 - (r_{xy})^{*T}(\Gamma_{xx})^{-1}\, r_{xy} \quad \text{Eq. 4}$$

For the assumption that d(n)=s(n) :

$$r_{xx}(k) = r_{ss}(k) + r_{ww}(k) \quad \text{Eq. 5}$$

$$r_{dx}(k) = r_{ss}(k) = r_{dd}(k) \quad \text{Eq. 6}$$

We will also verify if the above assumption holds true for our case. We can do analysis of the Fig(hmmmm.2) and Fig(hmmm.1). Not just the frequency response of the correlation matrices must be compared but also the matrices themselves.

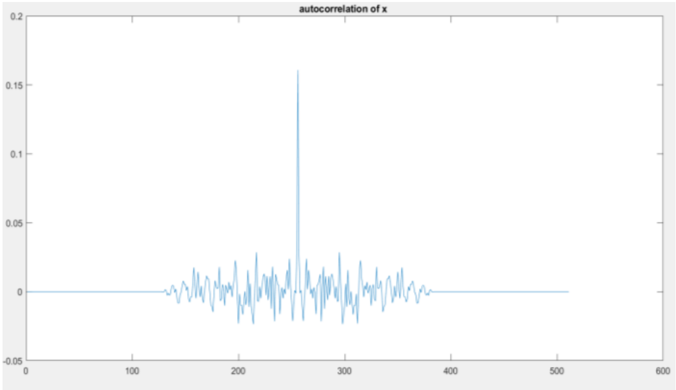
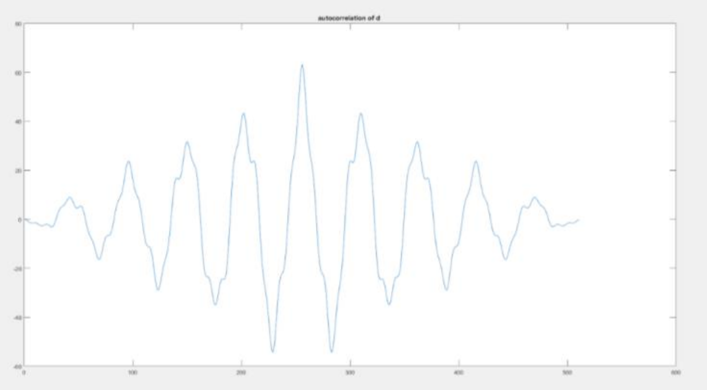


Fig (5). Autocorrelation of x(n).

Autocorrelation of x(n) appears symmetric, but there are subtle difference in the values that are observed. This symmetry is due to high amount of white noise in the signal.



Fig(6). Autocorrelation of 256 samples of the desired signal

Now we will observe the outputs for the each frame at a time. Here the information of just one frame is too less for obtaining y(n). Hence we will observe the frequency response of the cross correlation of desired signal and input signal ($F(r_{dx}(k))$).

Frame number 250 is selected for finding the cross correlation with desired signal. The frequency response of this cross correlation is observed to be similar to frequency response of autocorrelation of the desired signal.

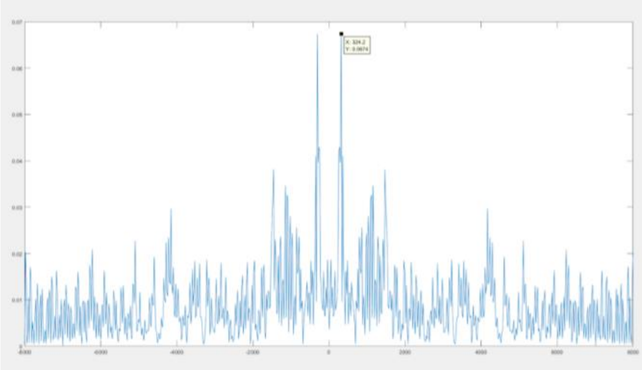
On applying the wiener filter to the whole signal segment at once we get the result shown in Fig (7). Here we see that the obtained response doesn't enhance the desired frequency components. Filter length obtained is 8001.
Here we are not specifically enhancing or filtering out any particular frequency. This output is a result of considering a non-stationary signal s(n) for wiener filtering. Hence we reaffirm the requirement of wiener filter that it works only for WSS signals and also that speech signal is a non-stationary signal.

We will evaluate MMSE and MSE for various frame lengths and overlaps:

We are calculating y(n) and its frequency response for the whole noisy speech signal but frame wise. Frame wise analysis helps us make the speech frames to behave like stationary signals.

Fig(7) Frame length 256 samples with 50 samples overlap was chosen to observe that frequency response of y(n). Lesser suppression of unwanted high frequency components also lesser enhancement of desired frequency components.

MSE : 0.33112



Fig(7) Frequency response of y(n) where x(n) has Frame length 256 samples with 50 samples overlap

Fig(8) has Frame length 256 samples with 200 samples overlap was chosen to observe that frequency response of y(n) has 300 Hz enhanced most followed by 800 Hz and 1500Hz. Slight deviations were also observed. Further a lot of high frequency components are also observed which can be made more accurate using NLMS approach.

MSE : 0.3350

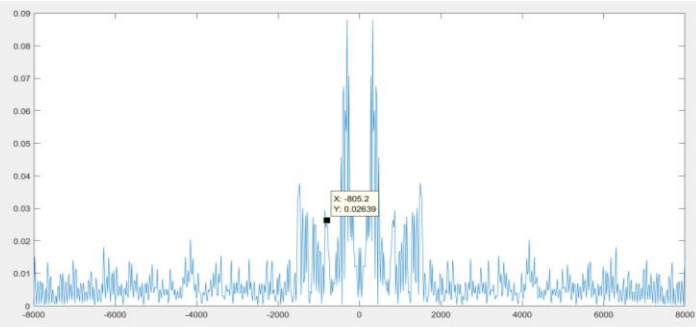
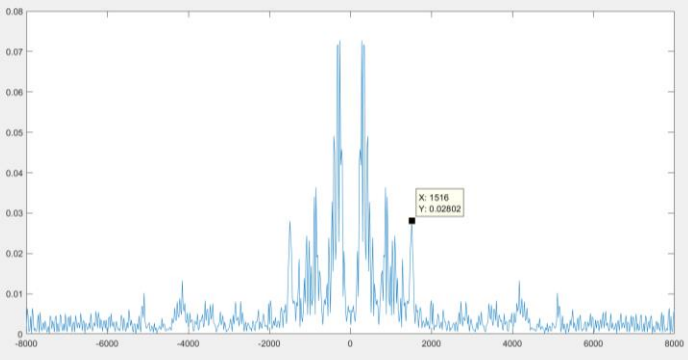


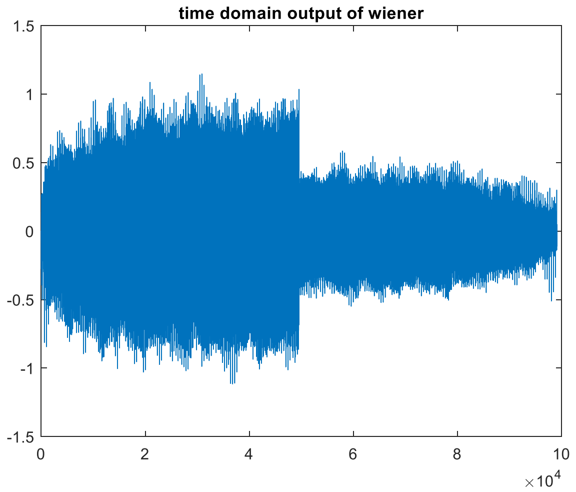
Fig (8) Frequency response of y(n) where x(n) has Frame length 256 samples with 200 samples overlap

Fig(9) Frame length 256 samples with 128 samples overlap was chosen to observe that frequency response of y(n). We observe the best result for frame overlap of 50%. Here we observe high attenuation of unwanted frequency components.
MMSE : 0.0608 ; MSE : 0.267412



Fig(9). Frequency response of y(n) where x(n) has Frame length 256 samples with 128 samples overlap

We have observed the obtained signals for frame wise wiener filtering.



Fig(9). Time domain representation of the y(n) from wiener

III. Using NLMS for Reduction of MMSE

Execution of this algorithm shall be carried out in the second phase of the project. This phase focusses on reduction of MMSE using Gradient decent by NLMS.Gradient decent approach is a stochastic gradient approach which is extended as LMS and NLMS.

LMS is used conventionally when it is easier for us to predict the learning rate. Mostly in real time applications of filtering, learning rates can be tedious to choose. Hence NLMS is a better method since it normalises the power of the input.

NLMS allows us to evaluate optimal learning rate. It is called the step size of the NLMS algorithm. This step size determines the pace at which we reach the optimal output. In our case reduced MMSE.

$$e(n) = d(n) - \hat{h}^{H}(n)\, X(n) \qquad \text{Eq. 7}$$

$$\hat{h}(n+1) = \hat{h}(n) + (\mu .\, e(n)\, X(n)) \,/\, X^{H}(n)\, X(n) \qquad \text{Eq. 8}$$

Optimal learning rate is given by μ_{opt}

$\mu_{opt} = 1$; For no interference, optimal learning rate.

$$\mu_{opt} = E[\,|\, y(n) - y\,(n)\ |^{2}] \,/\, E\,[\,|\, e(n)\,|^{2}] \qquad \text{Eq. 9}$$

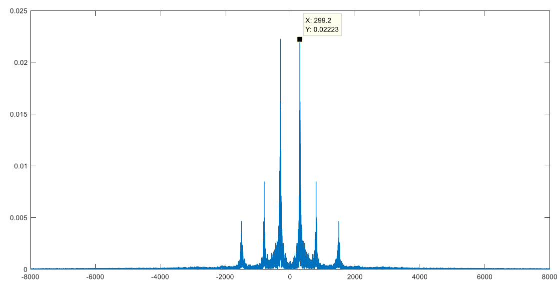
For signals with noise, learning rate.

Note: for filter order p :

$$X(n) = [x(n), x(n-1), \ldots, x(n-p+1)]^{T}$$

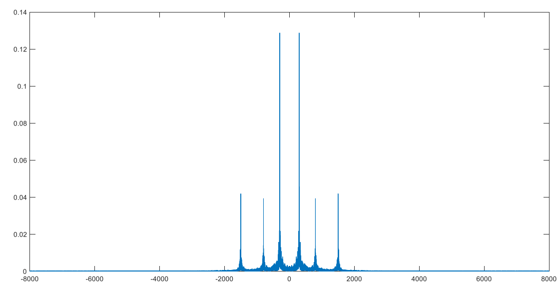
NLMS shall be performed in the second Phase of the project where the emphasis would be on obtaining an optimized result for the MMSE.

IV. Observations for Pure NLMS Filtering



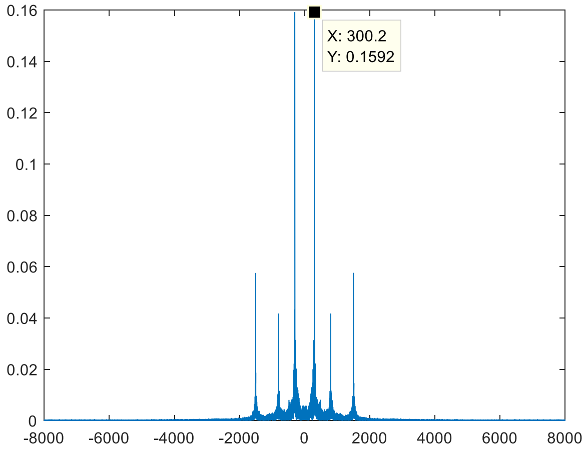
Fig(11). Frequency response of y(n) where x(n) has Filter length 256 samples for NLMS step size 0.05

In Fig(11), we see an exceptional reduction of noise and a great regeneration of the desired signal from the given input signal. The step size chosen here is the 0.05. We observe an MSE of 0.2523. The ratio of amplitudes for the desired frequencies observed in the output signal is 2.6:0.8:0.4. On listening to the output, we observe that some amount of speech signal is still present in the end.



Fig(12). Frequency response of y(n) where x(n) has Filter length 256 samples for NLMS step size 0.5

In Fig(12), we see an exceptional reduction of noise and a great regeneration of the desired signal from the given input signal. The step size chosen here is the 0.5. We observe an MSE of 0.2266. The ratio of amplitudes for the desired frequencies observed in the output signal is3:0.75:1. On listening to the output, we observe that some amount of speech signal is still present in the end, but less than the previous example.



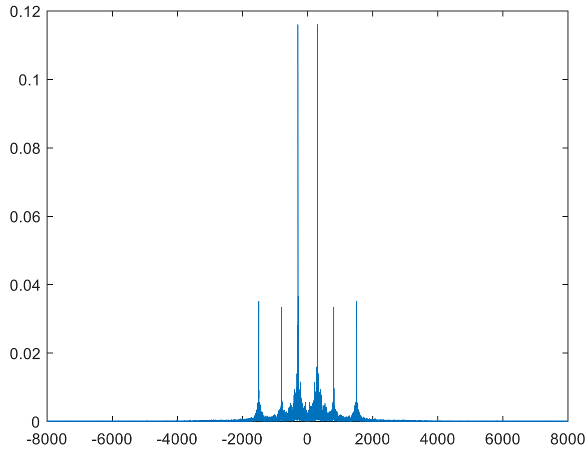
Fig(13). Frequency response of y(n) where x(n) has Filter length 256 samples for NLMS step size 1.0

In Fig(13), we see an exceptional reduction of noise and a great regeneration of the desired signal from the given input signal. The step size chosen here is the 1.0. We observe an MSE of 0.3054. The ratio of amplitudes for the desired frequencies observed in the output signal is 4:1:2.6. On listening to the output, we observe that some amount of speech signal is still present in the end, more than the above two observations. Here clearly, larger step size overshoots the desired signal. Larger step size ensures faster adaption, but lower step size ensures accurate convergence of MSE.

order - Step size:	100	256	500	100
0.005	0.2490	0.2478	0.2467	0.2461
0.05	0.2505	0.2523	0.2522	0.2540
0.4	0.2141	0.2231	0.2237	0.2623
0.5	0.2154	0.2266	0.2260	0.2697
1	0.2839	0.3054	0.3076	0.3506

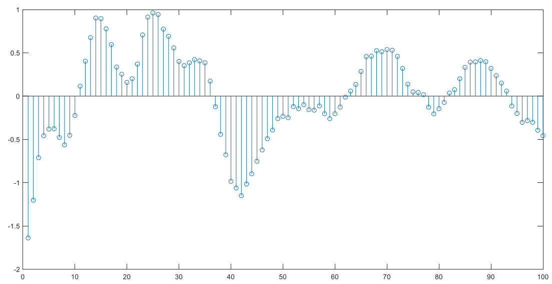
Fig(14). Table of MSE for various step size and order of Filter

The above Table is analyzed using MatLab. This Table indicates that the best result for the MSE is obtained usingThe stepsize of 0.4 with a filter order of 100. This table is result NLMS method.

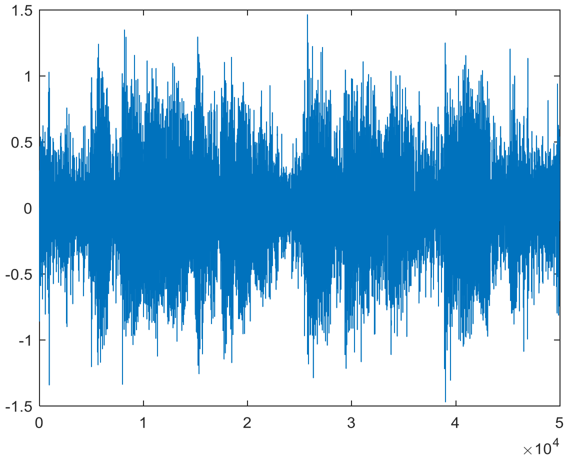


Fig(15). Frequency response of y(n) where x(n) has Filter length 100 samples for NLMS step size 0.4

In Fig(15), we see an exceptional reduction of noise and a great regeneration of the desired signal from the given input signal. The step size chosen here is the 0.4. We observe an MSE of 0.2141. The ratio of amplitudes for the desired frequencies observed in the output signal is3:0.98:1. On listening to the output, we observe that very small amount of speech signal is still present in the end, but this case gives the best result.

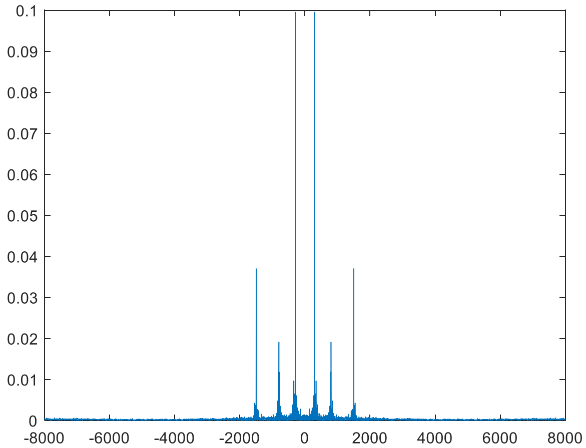


Fig(16). Impulse Response of the optimal filter using NLMS with step size 0.4



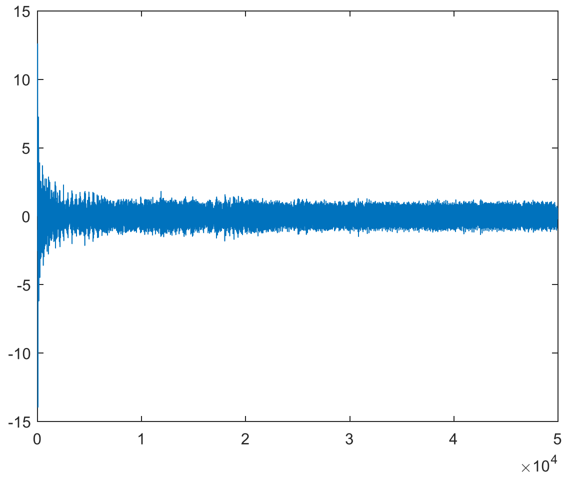
Fig(17). Time domain response of y(n) for NLMS filter using filter length 100 and step size 0.4

V. Observations for Wiener Filter followed by NLMS

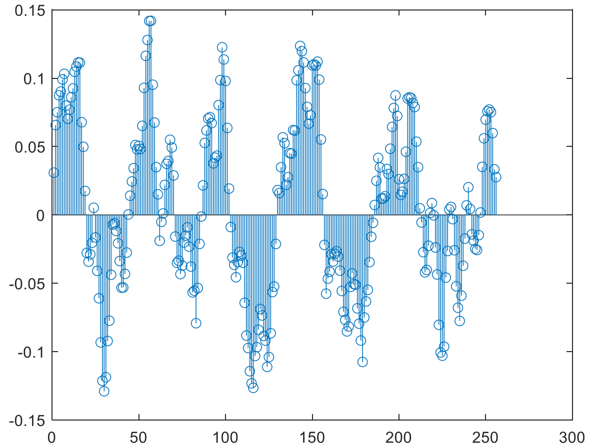


Fig(18). Frequency response of the output from Wiener filter followed by NLMS for best regeneration of the desired signal

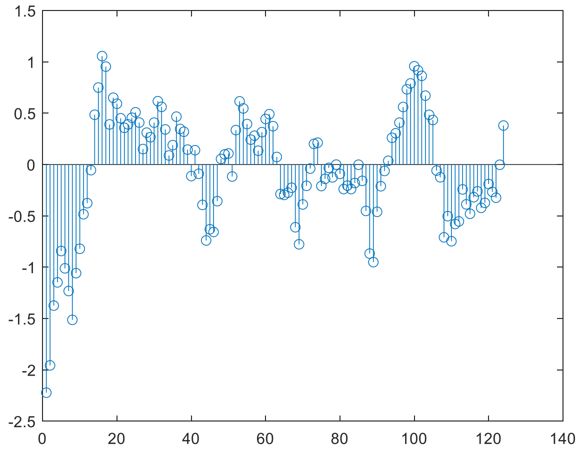
In Fig(18). we see an exceptional reduction of noise and a great regeneration of the desired signal from the given input signal. The step size chosen here is the 0.4. We observe an MSE of 0.2473. The ratio of amplitudes for the desired frequencies observed in the output signal is3:0.98:1.2. On listening to the output, we observe that no amount of speech signal is present in the end, but this case gives the best result.



Fig(19). Time domain representation of best regeneration



Fig(20). Impulse response of the NLMS filter used for the best regeneration.



Fig(21). Impulse response of the Wiener filter used prior to NLMS for best regeneration

Wiener Filter followed by NLMS results in the best regeneration of the desired signal. From the above observed plots, we can easily say that we have obtained output very close to the desired signal.

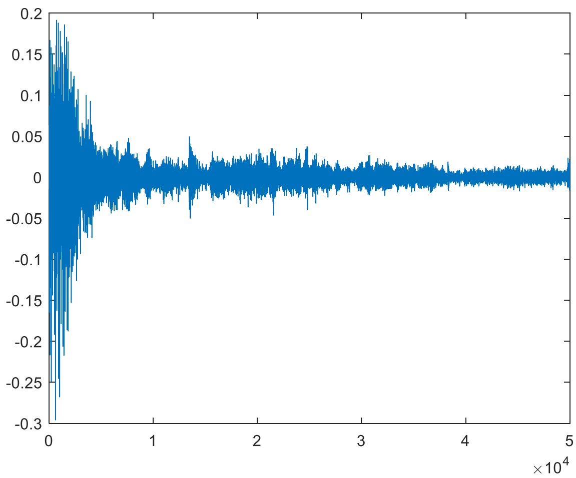
Wiener filter used here uses a filter length of 124. Whereas NLMS uses a length of 256, with a step size of 0.5.

For reduction of the filter order and MSE of the overall system, we can use a filter order of 5, with a step size of 0.00267.
This step size is obtained from equation :

Step size = 2/ max(eigen value of the auto correlation of x(n))

This step size yields us the following results:

MSE = 0.2473. This MMSE is obtained using the matlab function. Least Value of the MSE is observed to be :
2.8194×10^{-15}



Fig(22). Reduction of the MSE with optimized step size.

VI. Conclusions and Prospects

1. Wiener filter is successfully used to find a system that performs the required filtering operation. Weiner filter algorithm gives us an optimal filter which is proven to be unique for the given filter length.
2. Optimum impulse response of the system is found using whole input signal at once. This fails because speech is non-stationary signal. We resort to breaking down the signal into frames. Small chunks of these data will be stationary signal.
3. Length of the filter will be twice the frame length selected. It is independent of the overlap chosen for the frames. In our case it is 512 for a frame length of 256.
4. We have obtained an MMSE of 0.0608 for the filter of length 512. Here MMSE is also independent of the overlap chosen for the frames.
5. NLMS algorithm will be used in second phase of this project. Using wiener filtering we obtain an optimum impulse response and MMSE which will be further reduced using NLMS.
6. NLMS is chosen over LMS because, it can optimize the learning rate (step size). Higher step size will result in faster adaptation at the cost of overshooting the solution. Order of filter will be a challenge.
7. On using Wiener filtering we are not able to produce an acoustically good output. Noise is not completely removed. MMSE observed : 0.3530
8. We use NLMS independently, here we produce an acoustically great output. MSE reduces to the least of 0.2141for the best case.
9. Now we use wiener filter first, followed by NLMS. This gives us best output acoustically. Also, the best result in frequency domain. We produce a very closely similar result to that of the desired signal. MSE = 0.2473. This MMSE is obtained using the matlab function. Least Value of the MSE is observed to be: 2.8194×10^{-15}

Acknowledgment

I would like to extend my gratitude to Dr. Issa M. S. Panahi for the opportuinity to work on this project and his guidance via his lectures. It has been a great experience doing this project.

References

[1] J.G. Proakis, D.G. Manolakis, " Digital Signal Processing, Principle, Algorithms, and Applications", 4Th (2007) Edition,. Prentice-Hall.

[2] Comparison between Adaptive filter Algorithms (LMS, NLMS and RLS) JYOTI DHIMAN1, SHADAB AHMAD2, KULDEEP GULIA3 1 Department of Electronics Engineering, B.M.I.E.T, Sonepat, India 2Asst. Professor, Department of Electronics Engineering, D.R.C.E.T, Panipat, India 3Asst. Professor, Department of Electrical Engineering, B.M.I.E.T, Sonepat, India

[3] Matlab-Routine