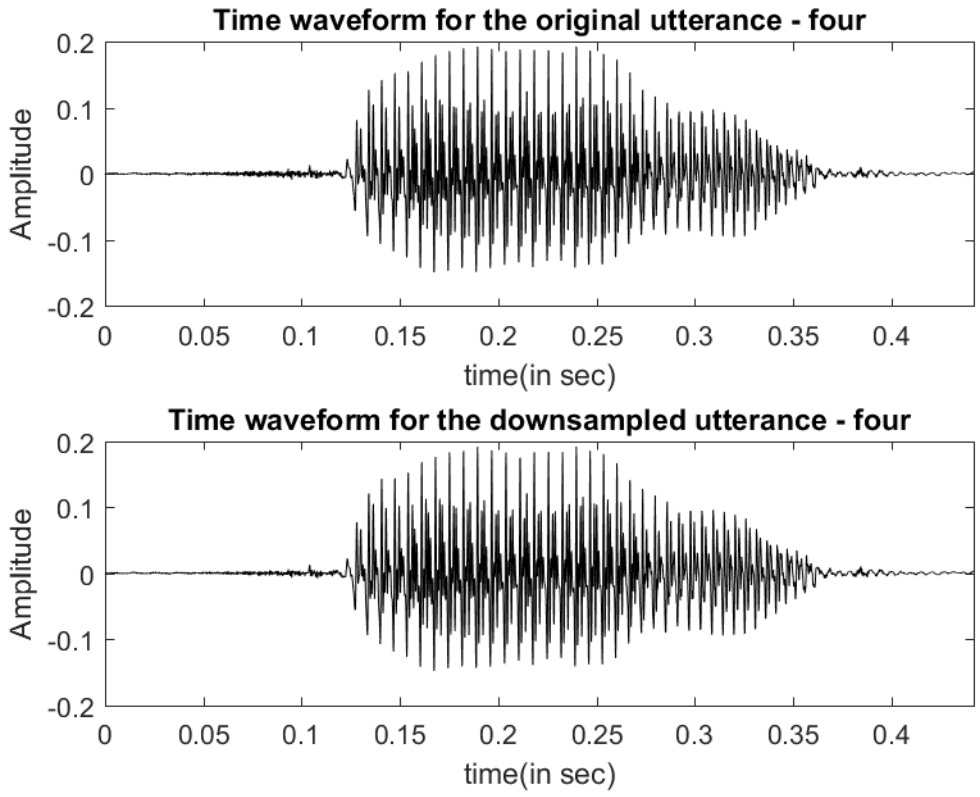
Assignment 4 CRL707

Akashdeep Bansal (2016ANZ8049)

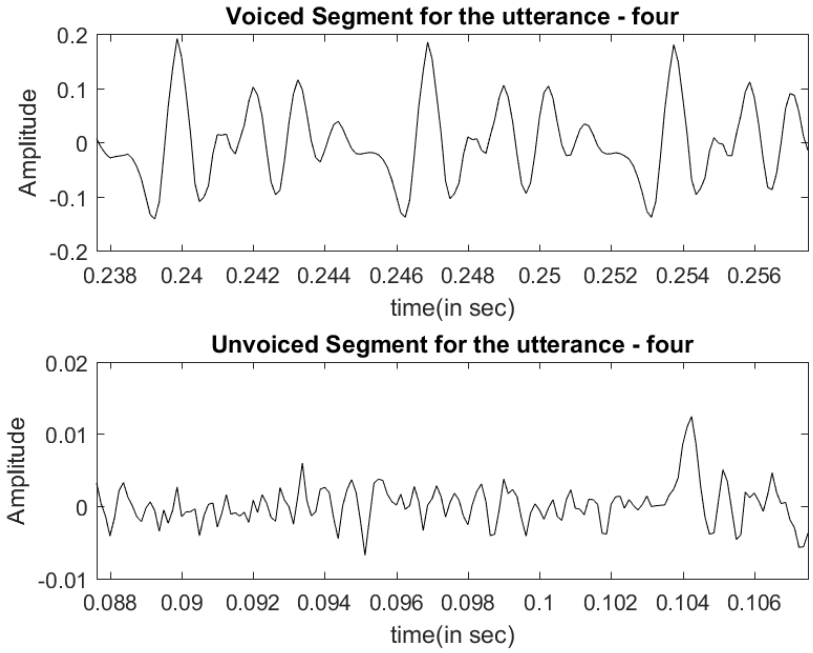
Q. As we are supposed to take two voiced segments and two un-voiced segments. We have chosen two digits: Four and Six. First, we will show results on voiced and unvoiced segment of digit four and then on Six.

The time waveform on 16k and 8k sampling are as shown below

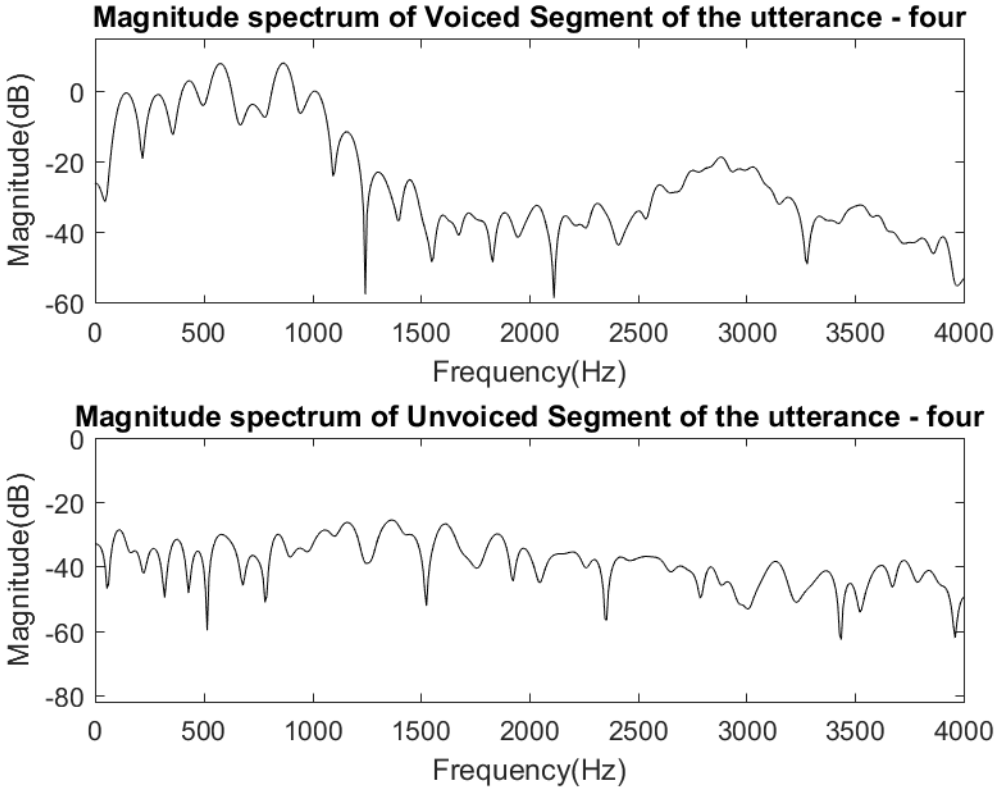


We can clearly see the loss of information due to down-sampling. A significant amount of loss can be observed in the period 0-0.15 sec and 0.38-0.43 sec.

The next figure shows the time waveform of the voiced and unvoiced segment chosen for the further analysis.



i) The magnitude spectrum of the selected segments are shown below -



ii) The 10^{th} order linear predictor coefficients and gain G are as follows –

For Voiced Segment

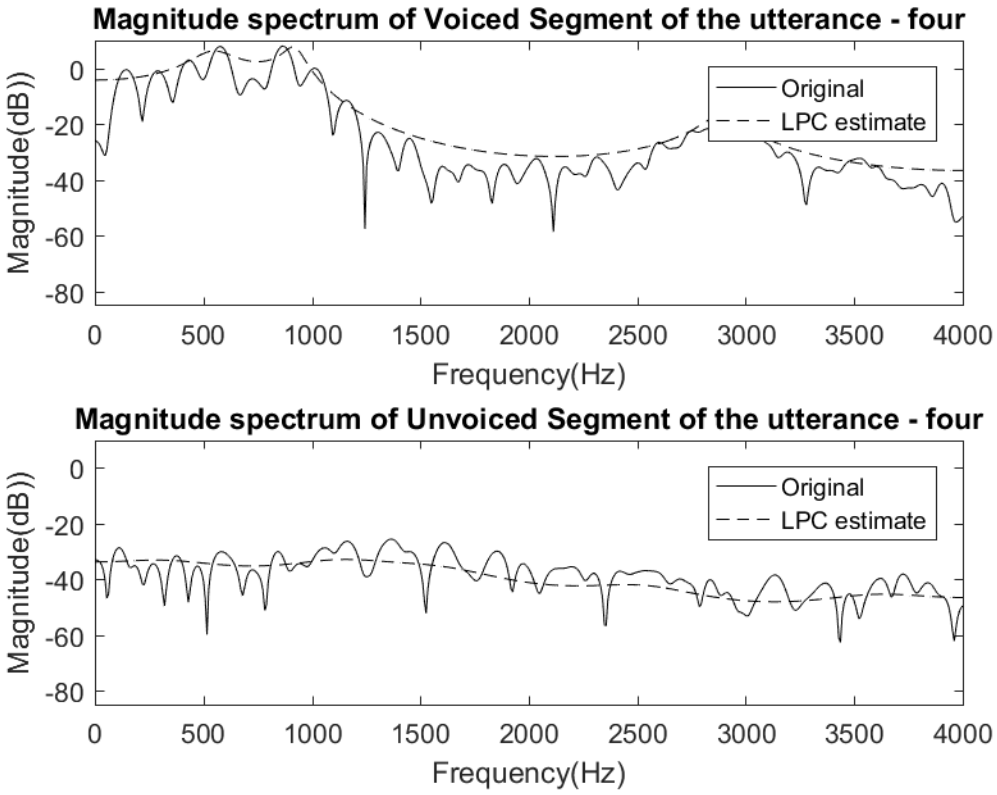
A(z) = [1.0000 -1.8892 1.0072 -0.0809 1.0425 -1.5123 0.4102 0.2448 0.0305 -0.0824 -0.0090]

G = 0.0097

For Unvoiced Segment

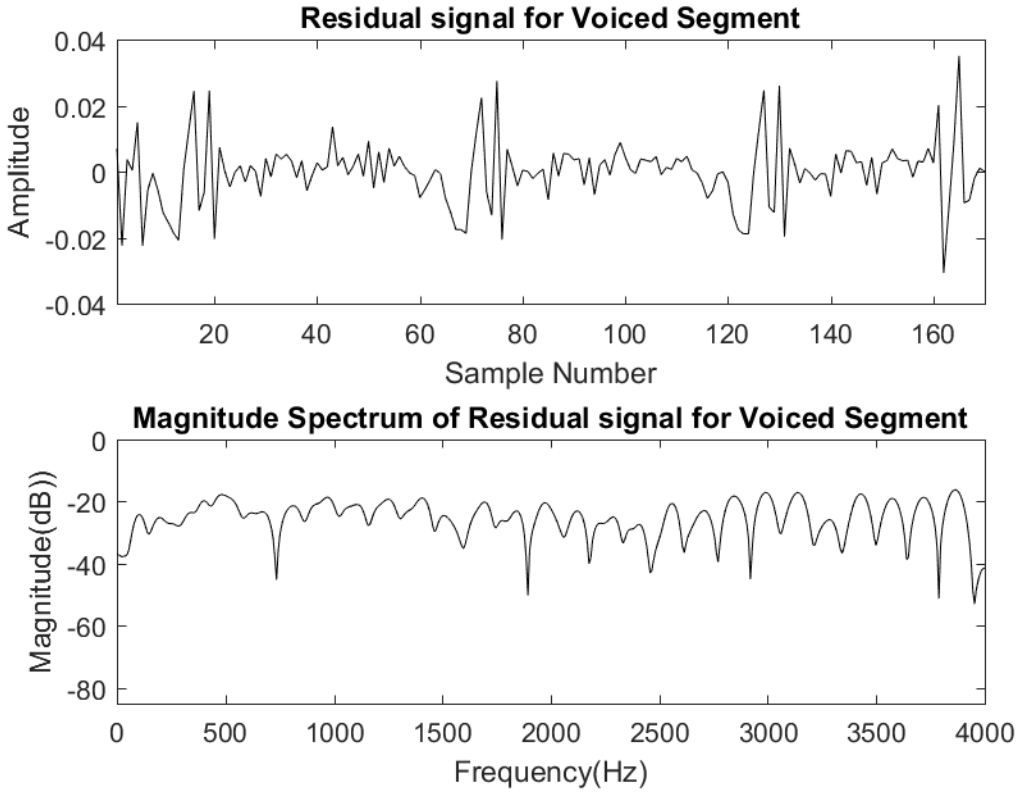
A(z) = [1.0000 -0.8677 0.4090 0.0407 -0.1424 0.0876 -0.1385 0.0483 0.0885 -0.1348 0.0993]

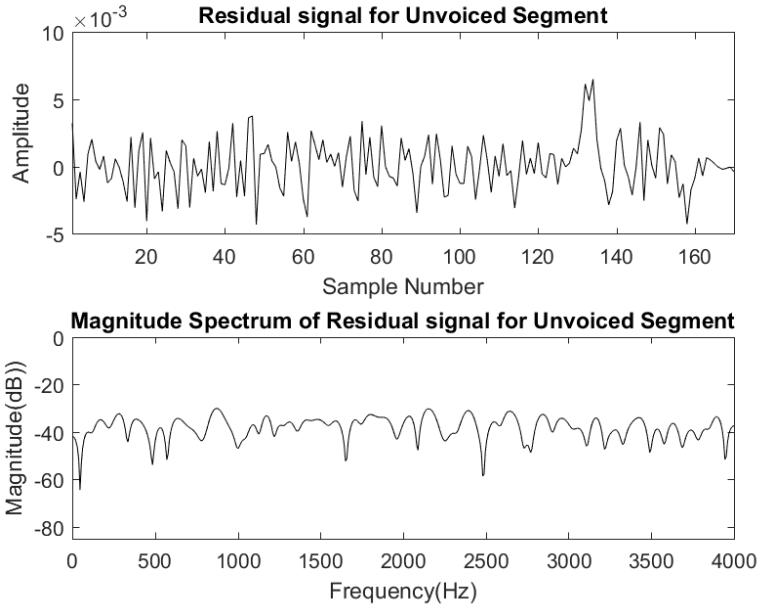
G = 0.0019



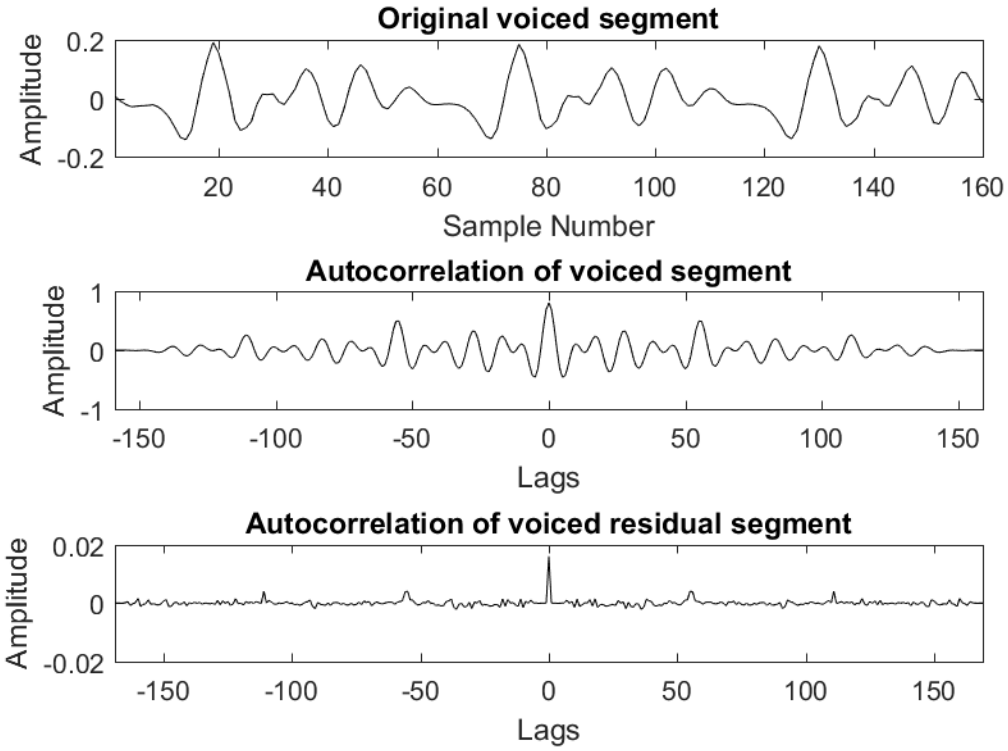
From the figure, we can observe that spectrum of A(z) overlaps with the spectrum of respective segments. In the voiced region, we can clearly see the pitch frequency and the formants. Whereas, this information is missing in the unvoiced. The spectrum of the unvoiced is highly random.

iii) The residual signal and their respective spectrum are shown below. The residual signal in case of voiced is higher in magnitude than unvoiced signal, because our predictor is just based on the previus 10 samples, not on the basis of the complete pitch period. Which is a tradeoff between bit rate required vs quality. Unvoiced segments is random, so relatively it gives less residual error.



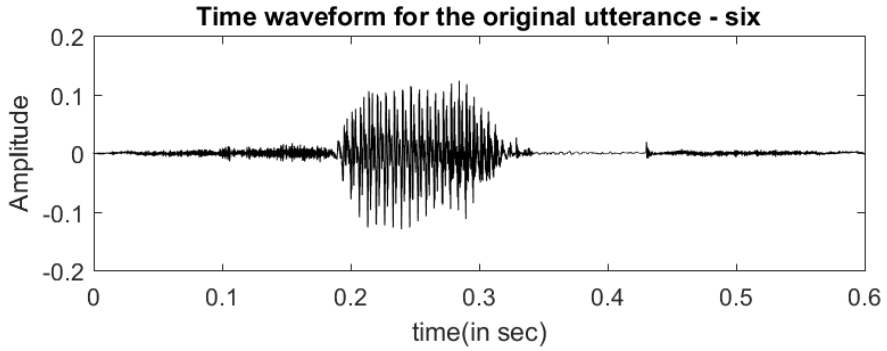


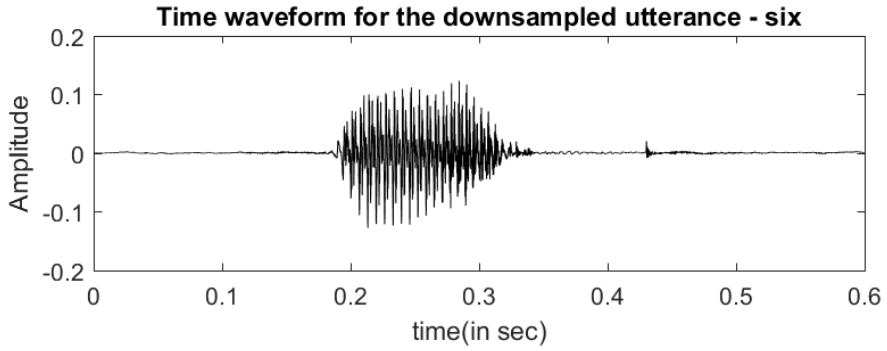
iv) The auto-correlation for the voiced segment and residual signal is shown in the figure below. We can clearly observe the peaks, which provide information about the pitch period.

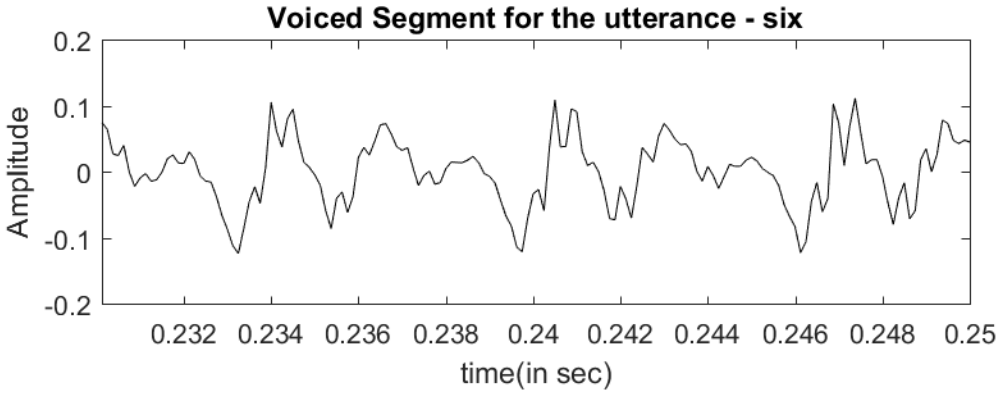


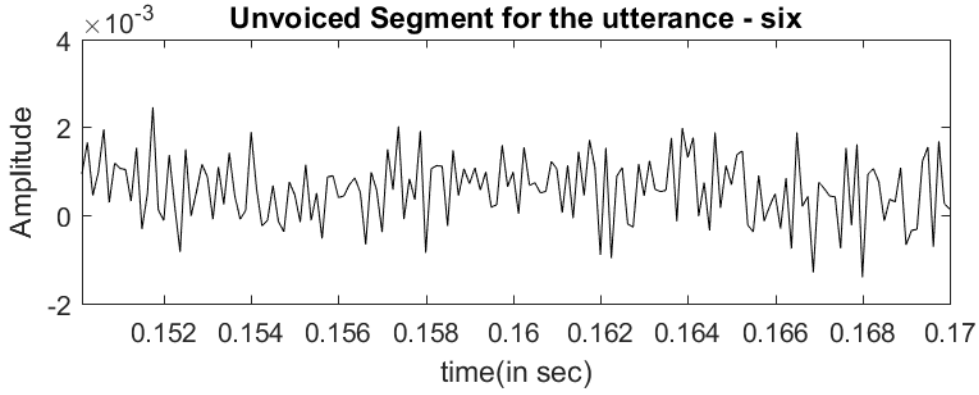
Now we have similar results for the voiced and unvoiced segments of the digit six. The observations are similar. So, explicit mentioning is not required.

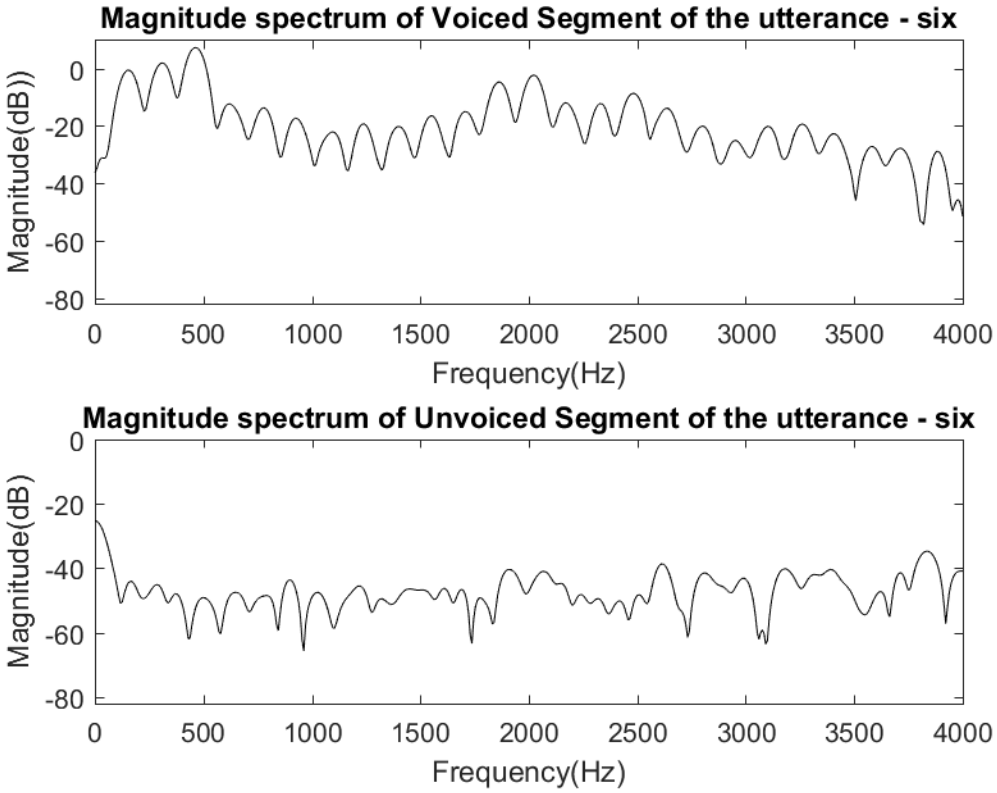
For another digit (Six)











For voiced segment,

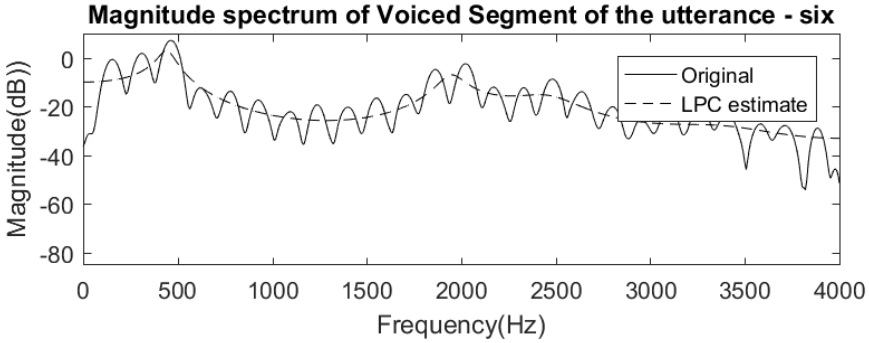
A(z) = [1.0000 -0.9460 0.5820 -0.9785 0.2630 0.1106 0.4244 0.0233 0.0490 -0.3045 0.0909]

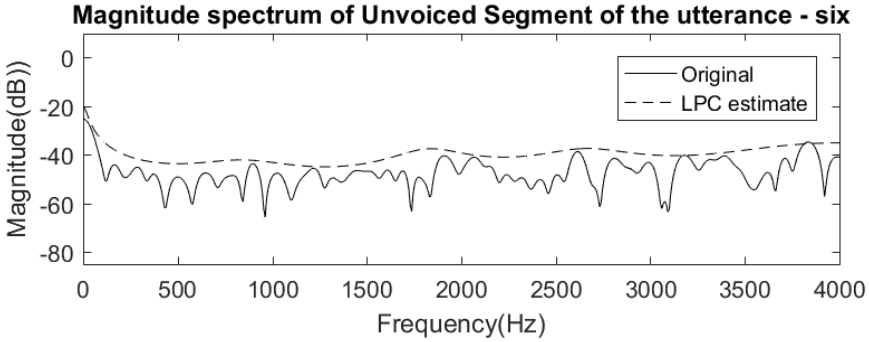
G = 0.0172

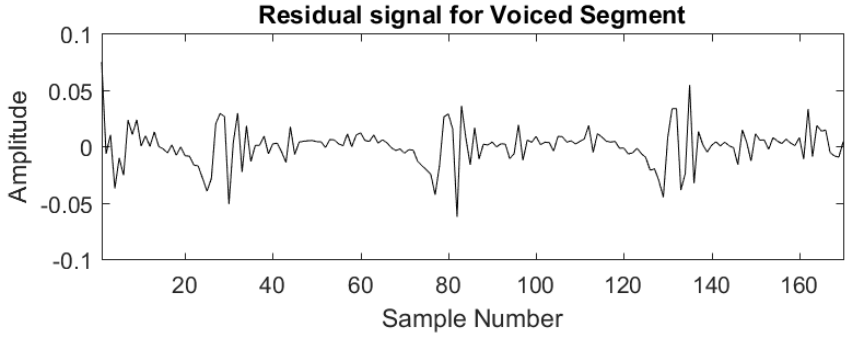
For Un-voiced segment,

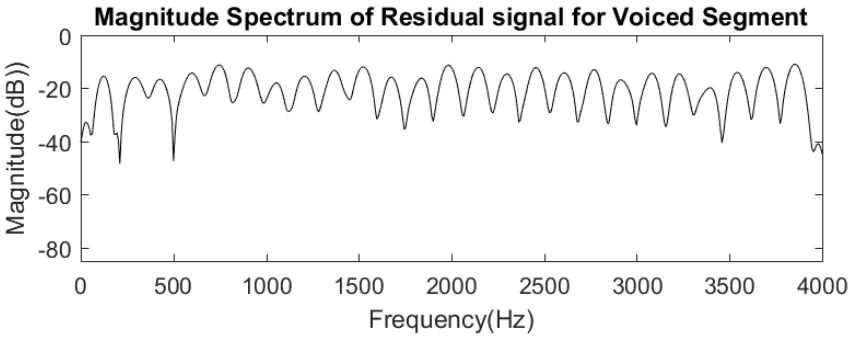
A(z) = [1.0000 0.2313 -0.1233 -0.1727 -0.2939 -0.0942 -0.0773 0.0283 -0.0737 -0.2268 -0.0948]

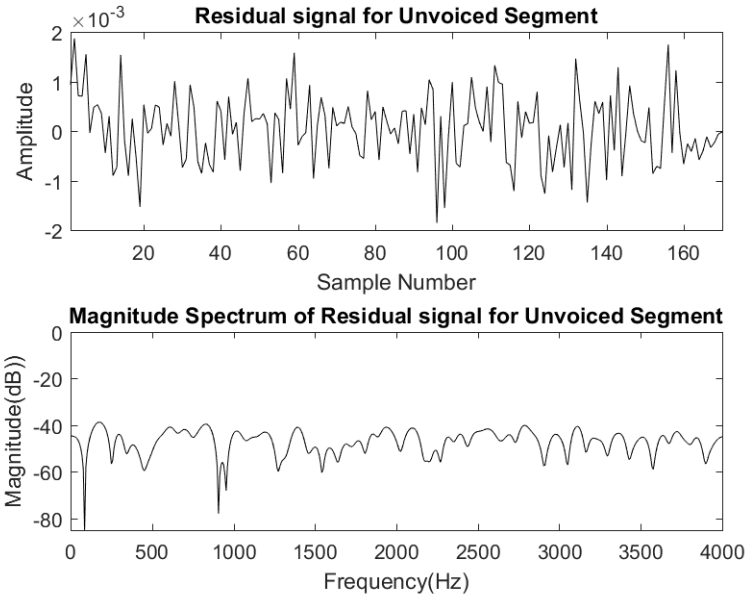
G = 6.9566e-04

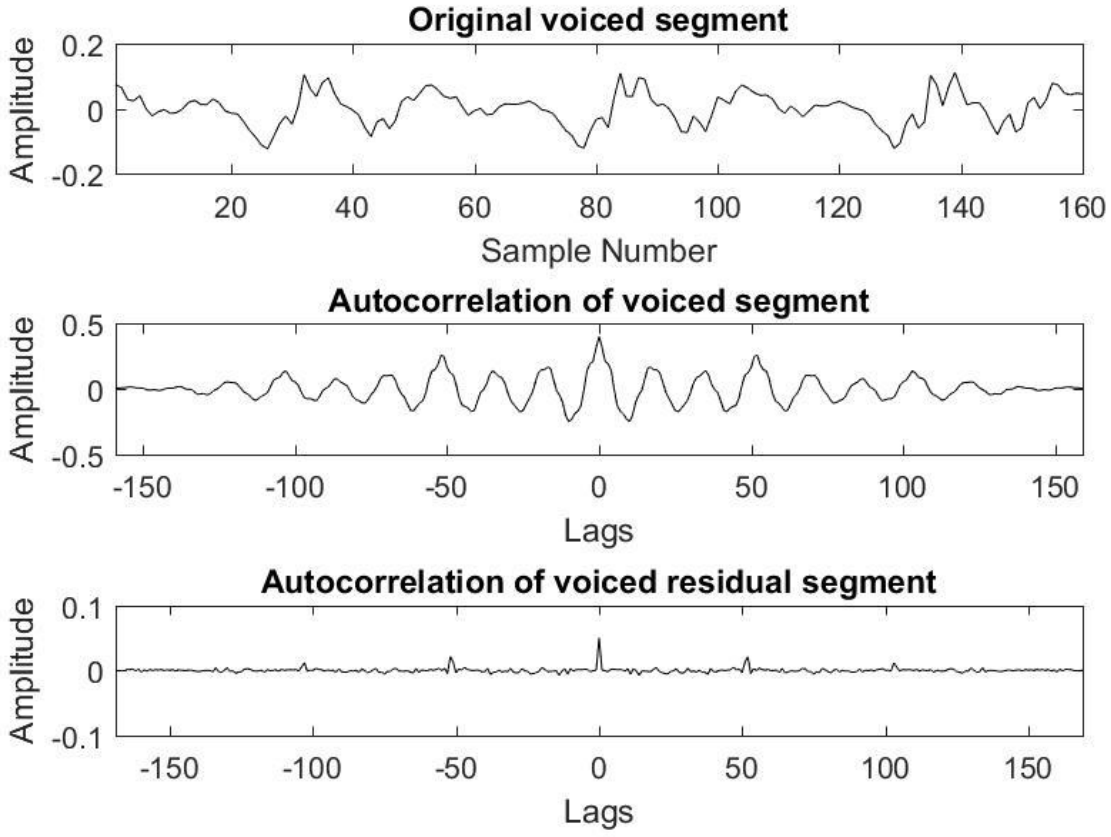

Magnitude spectrum of Voiced Segment of the utterance - six
Original
LPC estimate
Magnitude(dB))
Frequency(Hz)


Magnitude spectrum of Unvoiced Segment of the utterance - six
Original
LPC estimate
Magnitude(dB))
Frequency(Hz)


Residual signal for Voiced Segment
Amplitude
Sample Number


Magnitude Spectrum of Residual signal for Voiced Segment
Magnitude(dB))
Frequency(Hz)


Residual signal for Unvoiced Segment
×10^-3
Amplitude
Sample Number
Magnitude Spectrum of Residual signal for Unvoiced Segment
Magnitude(dB))
Frequency(Hz)


Original voiced segment
Amplitude
Sample Number
Autocorrelation of voiced segment
Amplitude
Lags
Autocorrelation of voiced residual segment
Amplitude
Lags