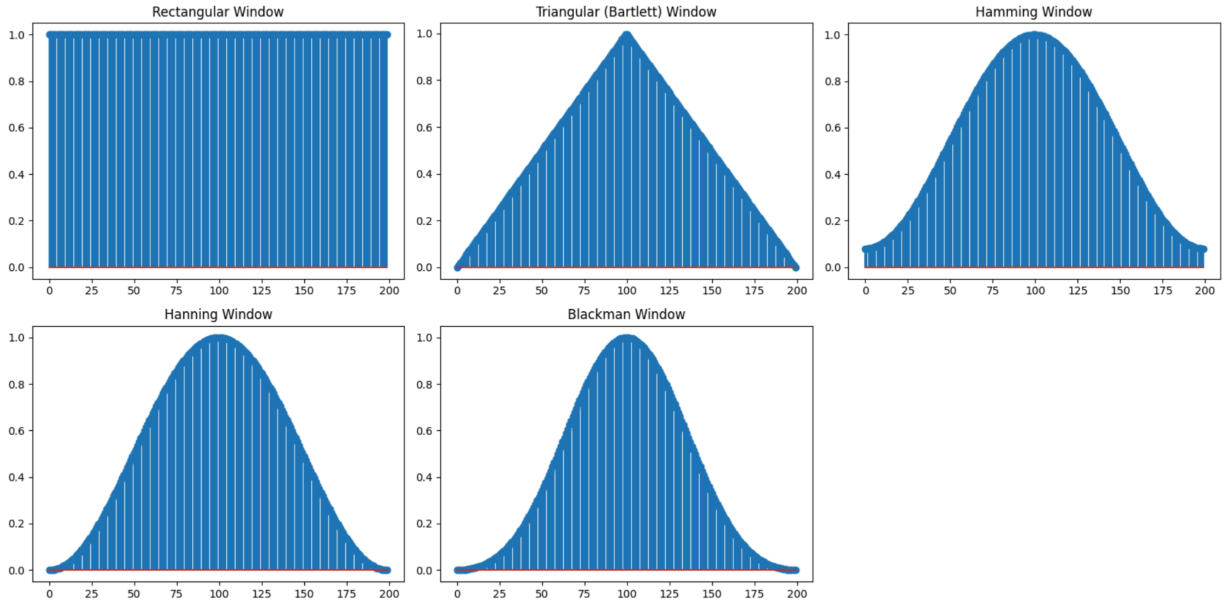
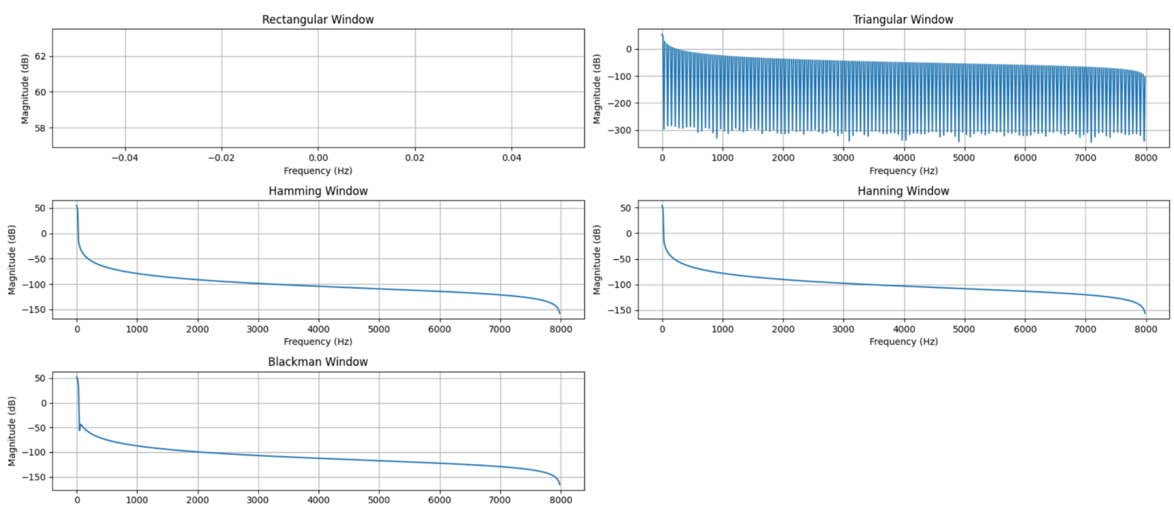
Note: All codes related to Question 1 and Question 3 are present in a single .ipynb file.

Question 1-a) ans:

(i) ans- Plotting the windows. Python code to build this plot is present in a separate .ipynb file.



(ii) ans- Plotting the log magnitude frequency responses.



(iii) ans-
Both Hamming window and Hanning window are considered as a good choice for short time speech processing tasks. But based on the obtained plot Blackman should be considered as best window out of five due to narrower main lobe and lower side lobe which will result in better frequency resolutions and better frequency attenuation.

Q1-b) solve :

Given,

time = 5.3 sec

sampling frequency, $F_s = 16 KHz$

window length, $L = 25$ msec

overlap = 10 msec

Now,

window length, L (in samples) $= 25 \times \frac{16\cancel{000}}{1\cancel{000}}$

$= 400$

Overlap (in samples) $= 10 \times \frac{16\cancel{000}}{1\cancel{000}} = 160$

Number of frames $= \left\lceil \frac{\text{Signal duration}}{\text{window length} - \text{overlap}} \right\rceil$

$= \left\lceil \frac{5.3 \times 16000}{400 - 160} \right\rceil$

$= \left\lceil \frac{848\cancel{0}}{24\cancel{0}} \right\rceil$

$\boxed{\text{No. of frames} = 354}$ ans/

Scanned with CamScanner

Q.2 - a) solve

Given,

$$E[n] = x^2[n] * h[n] \quad ; \quad h[n] = (1-a)a^{n-1}u[n-1]$$

(i) solve

we know

$$E[n] = \sum_{k=-\infty}^{\infty} x^2[k] * h[n-k]$$

$$E(z) = X(z) \cdot H(z)$$ where $X(z)$ and $H(z)$ is z-transform of $x[n]$ & $h[n]$ — (i)

Now,

$$H(z) = \sum_{n=0}^{\infty} h[n] \cdot z^{-n} = \sum_{n=0}^{\infty} (1-a)a^{n-1}u[n-1]z^{-n}$$

$\because$ $u[n-1]$ is 0 for $n<1$.

$$\therefore \quad H(z) = \sum_{n=1}^{\infty} (1-a)a^{n-1}u[n-1]z^{-n} = \sum_{n=1}^{\infty} (1-a)a^{n-1}z^{-n}$$

$$= (1-a)\sum_{n=1}^{\infty} a^{n-1}z^{-n}$$

$$= \frac{1-a}{a}\sum_{n=1}^{\infty}\left(a/z\right)^n$$

$$= \left(\frac{1-a}{a}\right)\left[a/z + \frac{a^2}{z^2} + \frac{a^3}{z^3} + \cdots\cdots\right]$$

$$= \left(\frac{1-a}{\cancel{a}}\right) \times \frac{\cancel{a}}{z}\left[1 + a/z + a^2/z^2 + \cdots\right]$$

$$= \frac{1-a}{z} \times \frac{1}{1-a/z}$$

$$= \frac{1-a}{\cancel{z}} \times \frac{\cancel{z}}{z-a}$$

$$\therefore \quad H(z) = \frac{1-a}{z-a}$$

& $$X(z) = X^2(z)$$

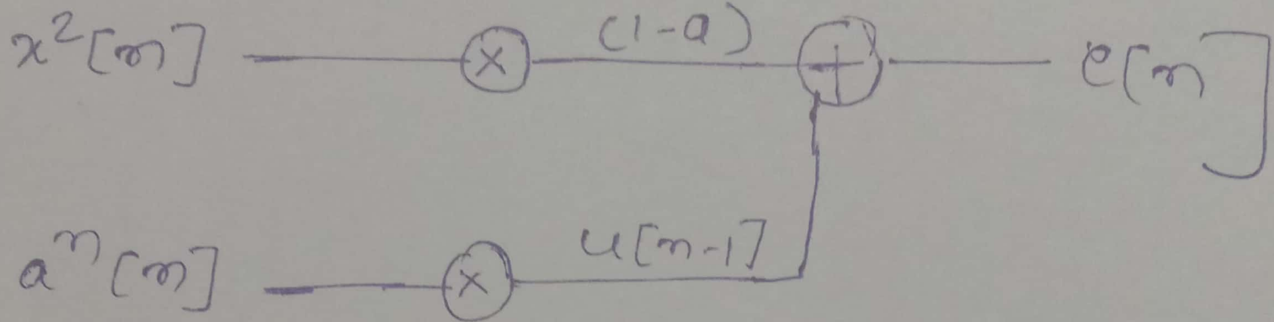
now, using equation (i)

$$E(z) = X^2(z) * \frac{(1-a)}{z-a}$$

to find difference equation take inverse z-transform.

$$\boxed{e[n] = (1-a)\, x^2[n] + a^n u[n-1]}$$ ans/

2 - a) ii



Scanned with CamScanner

2-b) ans – given that,

$$h[n] = a^n \ \forall \ n \geq 0 \ \& \ 0 \leq a \leq 1$$

$h[n]$ is exponential decay window.

This can have some advantages and disadvantages also.

Advantages :-

→ exponential windows can help in spectral analysis resulting better frequency resolution.

→ In feature extraction, this can be helpful in capturing short-term changes.

Scanned with CamScanner

→ This exponential decaying window can be helpful in noise reduction by emphasizing recent speech signal and attenuating older ones.

Disadvantages :-

(i) smearing : sharp features in frequency domain are spread out over neighbouring frequencies. This can reduce the ability to resolve fine spectral details.

(ii) Ringing : exponential decay property can lead to ringing artifacts in the time domain.

2-(C) solve

~~For window duration = 20 msec~~

~~Number of samples = 20/1000 *~~

for $F_s = 8$ kHz :

window duration = 20 msec

$$\text{Number of samples} = \frac{20}{1000} \times 8000 = 160 \text{ samples}$$

$$F_0 = 1000 \text{ Hz}$$

$$\therefore \quad ZCR = (\text{samples}) * ZCR \text{ (per sample)}$$

$$= 160 * 2 \times \frac{F_0}{F_s}$$

Scanned with CamScanner

$$= \cancel{160}^{40} \times \cancel{2} \times \frac{\cancel{1000}}{\cancel{8000}_{4}}$$

$= 40$ ans

for $F_S = 10\ kHz$:

window duration = 20 msec

Number of samples $= \frac{20}{\cancel{1000}} \times \frac{10\ \cancel{000}}{\cancel{1000}}$

$= 200$

$F_0 = 1000\ Hz$

$\therefore$ $ZCR = 20\cancel{0} \times 2 \times \frac{\cancel{1000}}{1\cancel{0}\ \cancel{000}}$

$= 40$

for $FS = 16\ kHz$

Number of samples $= \frac{20}{\cancel{1000}} \times 16\ \cancel{000}$

$= 320$

$$ZCR = \cancel{320}^{40} \times \cancel{2} \times \frac{\cancel{1000}}{\cancel{16000}_{8}}$$

$= 40$

Hence,

ZCR (for 8kHz) = 40

ZCR (for 10kHz) = 40

ZCR (for 16kHz) = 40 ans