Review and Conclusions

Jose Krause Perin

Stanford University

August 15, 2018

Outline

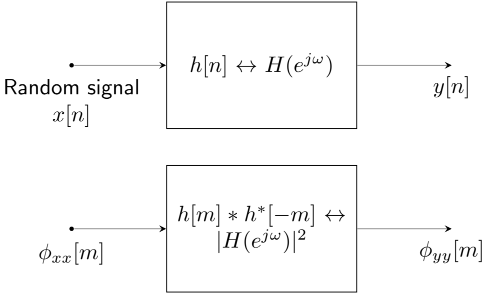
Random processes

- We use random processes to model signals that cannot be easily described by simple equations
- A random process is wide-sense stationary (WSS) if its mean is constant and if its autocorrelation function is only a function of the time difference.
- A random process is ergodic if its time averages are equal to its probability averages
- Autocorrelation functions must be real, even symmetric, and must have non-negative Fourier transform.
- The Fourier transform of the autocorrelation function is called the power spectrum density (PSD). The PSD has units of W/Hz or dBm/Hz.
- Random processes that have PSD constant over all frequencies are called white noise

$$\Phi_{xx}(e^{j\omega}) = \sigma_x^2 \Longleftrightarrow \phi_{xx}[m] = \sigma_x^2\delta[m] \qquad \text{(white noise)}$$

Random processes

Effect of filtering random signals



- The input autocorrelation function is *filtered* by the LTI system defined by $h[n] * h^*[-n] \leftrightarrow |H(e^{j\omega})|^2$
- $c_{hh}[n] = h[n] * h^*[-n]$ is called the deterministic autocorrelation function.
- Average power after filter

$$\sigma_y^2 = \phi_{yy}[0] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{yy}(e^{j\omega}) d\omega$$

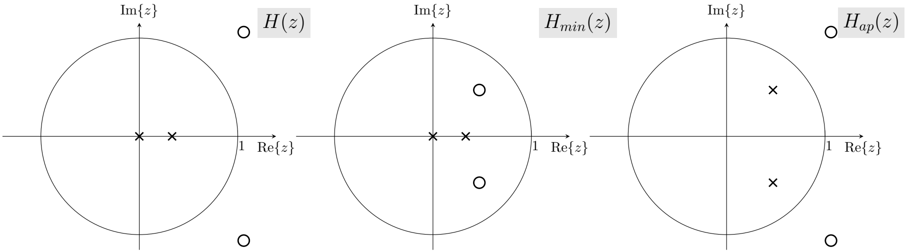
Properties of LTI systems

- The frequency response of a system tell us how much each frequency was scaled (magnitude response), and delayed (phase response) by the system.
- Poles increase magnitude and introduce phase lag (positive group delay)
- Zeros decrease the magnitude and introduce phase lead (negative group delay)
- All-pass systems have constant magnitude response. For each pole at e_k, there is a zero at $1/e_k^*$ (conjugate reciprocal)
- Minimum phase systems have all zeros inside the unit circle. Hence, its inverse is stable.
- Any system can be decomposed into a cascade of a minimum phase system and an all-pass system
- For minimum phase systems, the phase response is given by the Hilbert transform of the log-magnitude response.
- The unwrapped phase response of generalized linear phase systems is piece-wise linear
- FIR systems are linear phase as long as their impulse response is even or odd symmetric
- Linear phase rational IIR systems do not exist

Example of minimum-phase/all-pass decomposition

Consider the non-minimum phase system

$$H(z) = \frac{(1 + \frac{3}{2}e^{j\pi/4}z^{-1})(1 + \frac{3}{2}e^{-j\pi/4}z^{-1})}{1 - \frac{1}{3}z^{-1}}$$

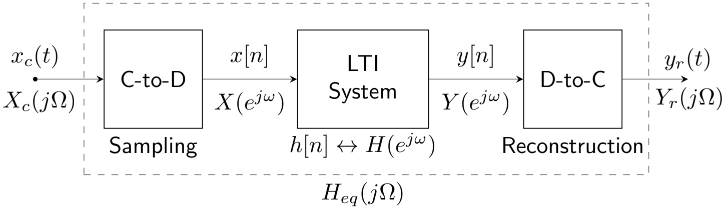


- The zeros of $H(z)$ outside the unit circle are *reflect* inside the unit circle (conjugate reciprocal) in $H_{min}(z)$
- The zeros of $H(z)$ outside the unit circle are now in $H_{ap}(z)$. Consequently, $H_{ap}(z)$ has poles at the conjugate reciprocal
- The poles of $H_{ap}(z)$ will cancel the new zeros of $H_{min}(z)$, yielding $H(z)$

Sampling, reconstruction, and DT filtering

- Sampling a continuous-time signal results in replicas of the spectrum at multiples of the sampling frequency Ω_s (or 2π of the normalized frequency ω)
- A band-limited signal has highest frequency Ω_N $(X_c(j\Omega) = 0, |\Omega| > \Omega_N)$
- If a band-limited signal is oversampled $(\Omega_s > 2\Omega_N)$ there'll be gaps between the spectrum replicas
- If the signal is undersampled $(\Omega_s < 2\Omega_N)$ the spectrum replicas will overlap resulting in aliasing distortion
- We can perfectly reconstruct a signal from its samples, provided that there is no aliasing and that we use the ideal lowpass filter as reconstruction filter
- In practice, we use different reconstruction filters, since the ideal lowpass filter is unfeasible.
- Oversampling relaxes the reconstruction filter specifications
- In theory, we can perform any LTI continuous-time filtering in discrete-time (in DSP), provided that there is no aliasing and that we use the ideal reconstruction filter

Discrete-time processing of continuous-time signals



$$H_{eq}(j\Omega) = H(e^{j\omega})|_{\omega=\Omega T}, |\omega| \leq \pi$$

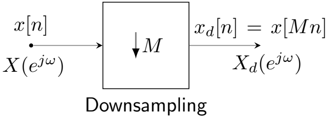
The equation above only holds if two conditions are met:

1. No aliasing
2. The reconstruction filter is the ideal lowpass filter

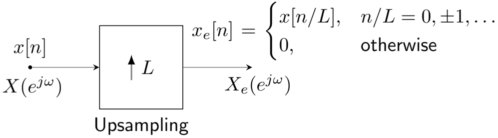
Conclusion: any continuous-time LTI system can be realized in discrete-time (in DSP).
Important: the equivalent continuous-time response $H_{eq}(j\Omega)$ depends on the sampling period T.

Changing the sampling rate in DSP: basic operations

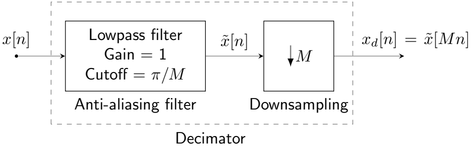
Donwsampling



Upsampling

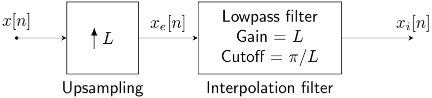


Decimation



Anti-aliasing filtering followed by downsampling.

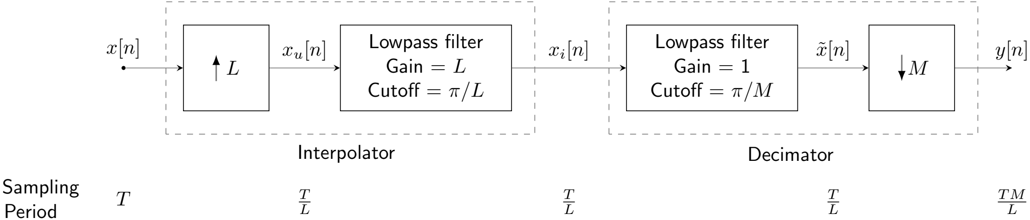
Interpolation



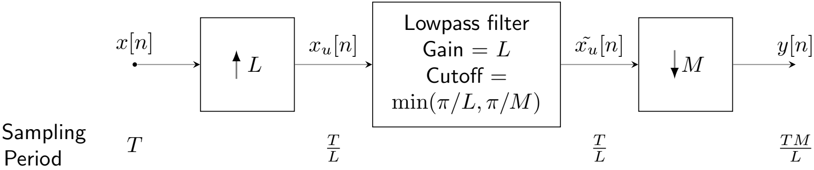
Upsampling followed by interpolation filter with gain L.

Interpolation/Decimation by a non-integer factor

Cascading interpolation and decimation



Equivalent diagram

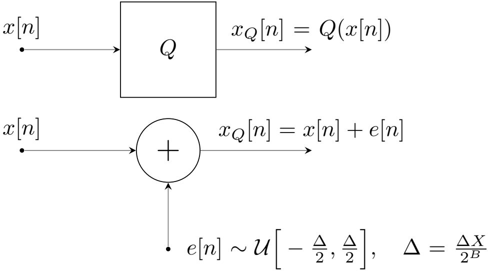


Changing the sampling rate in DSP

- Downsampling by an integer factor M stretches the discrete-time spectrum by a factor M and causes replicas of the spectrum to appear at $2\pi/M$. The amplitude of the spectrum is attenuated by M
- It's often easier to think of downsampling as sampling the original continuous-time signal with a sampling period $T_d = MT$
- Anti-aliasing filtering followed by downsampling is called decimation
- Upsampling by an integer factor L compresses the discrete-time spectrum by a factor L. The interpolation filter is assumed to have gain L, so the spectrum amplitude is scaled by L
- We can achieve non-integer sampling rate changes by cascading interpolation and decimation stages

Quantization: linear noise model

Quantization is modeled as an additive **white noise uniformly distributed** that is independent of the input signal.



ΔX is the quantizer dynamic range and B is the resolution in bits.

$$\sigma_e^2 = \frac{\Delta^2}{12} \qquad \text{(average power)}$$

$$\phi_{ee}[n] = \sigma_e^2 \delta[n] \qquad \text{(autocorrelation function)}$$

$$\Phi_{ee}(e^{j\omega}) = \sigma_e^2, |\omega| \leq \pi \qquad \text{(PSD)}$$

Quantization

- Quantization is unavoidable in DSP systems
- Using the linear noise model, we simply replace quantizers by noise sources of average power $\sigma_e^2 = \Delta^2/12$
- Quantization noise is assumed white (samples are uncorrelated)
- Every extra bit of resolution in a quantizer improves the SNR by 6.02 dB
- The signal amplitude must be matched to the dynamic range of the quantizer, otherwise there'll be excessive clipping or some bits won't be used
- Noise shaping is a strategy that minimizes quantization noise in A-to-D and D-to-A converters. The goal is to shape the quantization noise PSD, so that most of the noise power falls outside the signal band
- Noise shaping requires oversampling

Digital filter structures

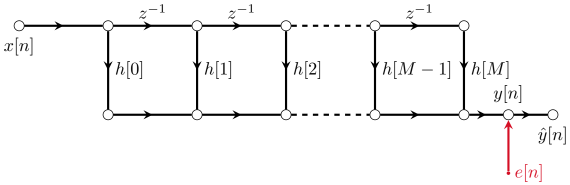
- There are different forms of realizing IIR and FIR rational systems
- All forms are based on the difference equation of a rational LTI system
- Pipelining and parallel processing solve the problem of using a slow hardware to process a fast signal in two complementary ways.
- Pipelining adds memory (delays) to minimize the critical path. Consequently, pipelining increases latency
- In parallel processing the hardware is replicated to allow processing of multiple input samples simultaneously
- Pipelining and parallel processing can be realized together
- Pipelining and parallel processing are more difficult in IIR systems due to their inherent feedback

Quantization in digital filter structures

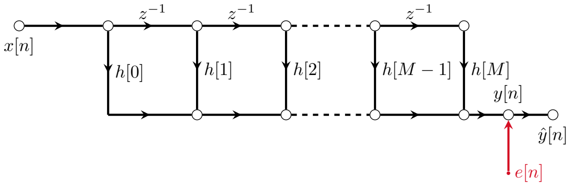
- Two's complement is a fixed-point representation that allows fractions to be represented as integers
- There's an inherent trade-off between roundoff noise and overflow/clipping
- FIR systems remain stable after coefficient quantization
- Linear phase FIR systems remain linear phase after coefficient quantization, since the impulse response remains symmetric
- Coefficient quantization may lead to instability in IIR systems, as poles may move outside the unit circle
- Similarly to quantization noise, roundoff noise is modeled by an additive white noise that is independent of the input signal (the linear noise model).
- Roundoff noise is minimized by performing quantization only after accumulation, but this requires $(2B+1)$-bit adders
- In FIR structures the equivalent roundoff noise at the output is white
- IIR structures lead to roundoff noise shaping
- Least noisy IIR structure depends on the system
- Cascade and parallel forms are used to mitigate total roundoff noise

Roundoff noise in FIR systems

1. Quantization immediately after each multiplication: $M+1$ noise sources are lumped into one: $\mathbb{E}(e^2[n]) = (M+1)\frac{\Delta^2}{12}$, where $\Delta = \frac{X_m}{2^B}$.
   Note: in linear phase FIR filter, only $\lfloor M/2+1 \rfloor$ multiplications are required.



2. Quantization immediately after accumulation: Only one noise source with average power $\mathbb{E}(e^2[n]) = \frac{\Delta^2}{12}$.



Roundoff noise in IIR systems

PSD of roundoff noise at the output

	I	II
Direct	$(M+N+1)\sigma_B^2 \frac{1}{\lvert A(e^{j\omega})\rvert^2}$	$N\sigma_B^2\lvert H(e^{j\omega})\rvert^2 + (M+1)\sigma_B^2$
Transposed	$N\sigma_B^2\lvert H(e^{j\omega})\rvert^2 + (M+1)\sigma_B^2$	$(M+N+1)\sigma_B^2 \frac{1}{\lvert A(e^{j\omega})\rvert^2}$

$M+1$ coefficients $\{b_0, b_1, \dots, b_M\}$, assumed different from zero and one ($b_i \neq 0, b_i \neq 1$).
N coefficients $\{a_1, \dots, a_N\}$, assumed different from zero and one ($a_i \neq 0, a_i \neq 1$).
$\sigma_B^2 = 2^{-2B}/12$ for a $(B+1)$-bit two's complement representation.

Desigining digital filters from analog filter

Impulse invariance:
Design $h[n]$ by sampling $h_{eq}(t)$ with period T.

$$h[n] = Th_c(nT) \qquad \text{(impulse invariance)}$$

Bilinear transformation:
Bilinear transformation maps the left-hand side of the s-plane into the unit circle in the z-plane

$$s = \frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right) \qquad \text{(Bilinear transformation)}$$

This nonlinear mapping causes frequency warping:

$$\omega = 2\arctan(\Omega T/2) \qquad \text{(frequency warping)}$$

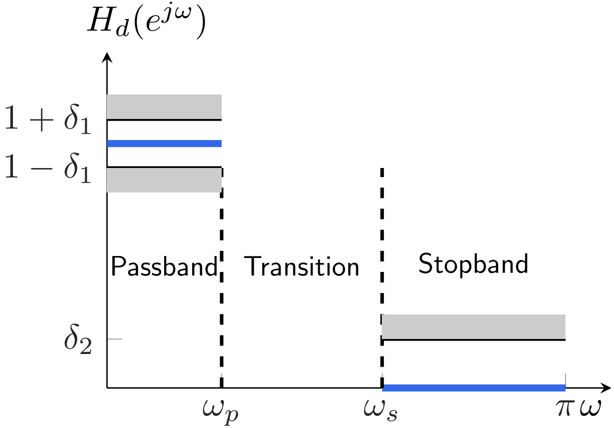
Frequency warping can be mitigated by performing frequency pre-warping, so that $H(e^{j\Omega_p T}) = H_{eq}(j\Omega_p)$ at a specified frequency Ω_p.

$$H(z) = H_{eq}(s)\Big|_{s = \frac{\Omega_p}{\tan(\Omega_p T/2)}\frac{1-z^{-1}}{1+z^{-1}}}$$

(bilinear transformation with frequency pre-warping)

Digital FIR filter design from specifications

How to find FIR $H(z)$ such that $H(e^{j\omega})$ best approximates a desired frequency response $H_d(e^{j\omega})$? Essentially a polynomial curve fitting problem.



Design techniques:

- Window method
- Optimal filter design
  - Parks-McClellan algorithm
  - Least-squares algorithm

Summary on FIR filter design by the window method

1. From the desired frequency response $H_d(e^{j\omega})$ calculate the desired impulse response $h_d[n]$.
2. Choose the filter order M and the window $w[n]$. Then,

$$h[n] = \begin{cases} h_d[n]w[n], & n = 0, \dots, M \\ 0, & \text{otherwise} \end{cases}$$

   Kaiser window depends on parameters β and M. Other windows only depend on M.
3. Linear phase is guaranteed by selecting causal and symmetric window $w[n]$ and desired impulse response $h_d[n]$

In Matlab:
`fir1` uses Hamming window by default. Other windows can be passed as parameters:

```
>> fir1(M, wc/pi, 'lowpass', kaiser(M+1, beta))
```

designs a lowpass FIR filter of order M and cutoff frequency ω_c using the window method with Kaiser window with parameter β

Optimal FIR filter design

Two algorithms:

1. Parks-McClellan algorithm: minimizes the maximum weighted error

$$\min_{h[n]} \max_{\omega} E(\omega) \qquad \text{(min-max problem)}$$

$$\min_{h} \max_{i} |e_i| \qquad \text{(in matrix notation)}$$

   `firpm` in Matlab.

2. Least squares: minimizes the mean-square weighted error

$$\min_{h[n]} \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(\omega)|^2 d\omega \qquad \text{(least squares)}$$

$$\min_{h} ||e||_2^2 \qquad \text{(in matrix notation)}$$

   `firls` in Matlab.

Non-linear phase FIR filter design using least squares

$$h = A^{\dagger} b \qquad \text{(least-squares solution)}$$

where $A = WQ$ and $b = Wd$.

$$Q = \begin{bmatrix} 1 & e^{-j\omega_1} & e^{-j2\omega_1} & \dots & e^{-jM\omega_1} \\ 1 & e^{-j\omega_2} & e^{-j2\omega_2} & \dots & e^{-jM\omega_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\omega_N} & e^{-j2\omega_N} & \dots & e^{-jM\omega_N} \end{bmatrix}_{N \times M+1}$$

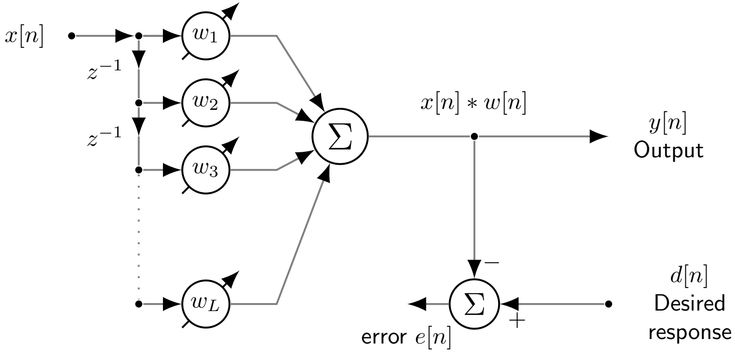
Important: d and W have to be defined for the same frequencies used in calculating Q.
If $H_d(e^{j\omega})$ is **Hermitian symmetric** i.e., $H_d(e^{j\omega}) = H_d^*(e^{-j\omega})$, then h is purely real.

Adaptive signal processing

- The linear combiner is the basis of adaptive systems and adaptive filtering
- We use the mean square error (MSE) as the performance metric
- The Wiener solution is the optimal set of weights that minimizes the MSE
- The LMS algorithm is a simple way to train the adaptive filter to approximate the Wiener solution
- The LMS algorithm uses the instantaneous error to obtain an estimate of the gradient
- This estimate is very noisy, but on average it converges to the Wiener solution
- We adjust the adaption constant to control how fast the LMS algorithm converges and how noisy the solutions near the Wiener solution (excess noise and misadjustment)

The adaptive FIR filter

At a time n, the input to the adaptive linear combiner is $X = [x[n], x[n-1], \dots, x[n-L+1]]^T$



The resulting FIR filter has L coefficients $\{w_1, \dots, w_L\}$.
Our goal is to adapt the coefficients $W = [w_1, \dots, w_L]^T$ to achieve the Wiener solution $W^\star = R^{-1}P$, which is the set of weights $W^\star = [w_1^\star, \dots, w_L^\star]^T$ that minimize the MSE $\mathbb{E}(e[n]^2)$

Adaptation algorithms: the LMS algorithm

The **least mean squares (LMS)** algorithm is a form of steepest descent where the gradient is estimated from the **instantaneous error**

$$\varepsilon_n^2 = (d_n - X_n^T W)^2 \qquad \text{(instantaneous error)}$$

$$\hat{\nabla} = \frac{\partial \varepsilon_n}{\partial W} = 2(d_n - X_n^T W)X_n = 2\varepsilon_n X_n \qquad \text{(gradient estimate)}$$

$$W \leftarrow W + 2\mu e_n X_n \qquad \text{(LMS weight update)}$$

Gradient estimate is very noisy, but on average the weights *generally* move to the Wiener solution.

The LMS algorithm

- **Learning curve:** The learning curve is a plot of the average MSE $\mathbb{E}(e^2[n])$ over time. In practice we use time averages of $e^2[n]$ to estimate $\mathbb{E}(e^2[n])$.
- **Stability or convergence:** The LMS algorithm is stable if $0 < \mu < 1/\mathrm{trace}(R)$. In the case of L-coefficient adaptive FIR filter, $\mathrm{trace}(R) = L\phi_{xx}[0] = L\,\mathbb{E}(x^2[n])$
- **Time constants:** The learning curve is a sum of decaying exponentials with time constants

$$(\tau_{MSE})_n \approx \frac{1}{4\mu\lambda_n} \text{ iterations} \qquad \text{(Steepest descent \& LMS)}$$

  where λ_n is the nth eigenvalue of matrix R.
- **Excess MSE:** The LMS may oscillate around the Wiener solution $W^\star = R^{-1}P$. This leads to an excess MSE.

$$\text{Total MSE} = \text{Minimum MSE} + \text{Excess MSE}$$

  The excess MSE is related to the misadjustment.

$$\text{Misadjustment} = \frac{\text{excess noise}}{\text{minimum MSE}} = \mu\mathrm{trace}(R) \qquad \text{(definition)}$$

The discrete Fourier transform

- Sampling the DTFT in frequency domain results in signal replicas in time domain
- The N-point DFT of $x[n]$ is equal to the DTFT of $x[n]$ sampled with period $2\pi/N$, only if $x[n]$ is time-limited with duration $\leq N$
- For sequences longer than N, the N-point DFT is equal to the samples of the windowed DTFT
- Fast Fourier transform (FFT) algorithms compute the DFT with complexity $\mathcal{O}(N \log_2 N)$
- We can use the DFT/FFT to perform linear convolution (filtering) efficiently using block convolution
- In the overlap-add method, blocks are non-overlapping and the result of circular convolution of each block is added to produce the output signal
- In the overlap-save method, blocks do overlap and we have to discard samples that are unusable due to the circular convolution not being equal to the linear convolution at all points

Spectrum analysis using the DFT

- Leakage and resolution are important considerations in spectrum analysis
- By choosing proper windows we can minimize these issues
- Kaiser window is a nearly optimal choice. Must choose correct β and window length L
- β controls the ratio between the amplitudes of the main-lobe and the largest side-lobe i.e., β controls the amount of leakage.
- The larger the main-lobe width, the smaller the resolution
- By increasing the window length we reduce the main-lobe width and consequently improve the resolution
- Time-dependent Fourier transform or short-time Fourier transform allows us to keep track of frequency variation in time
- Spectrogram is a commonly used way to display the TDFT
- In the spectrogram the TDFT is sampled both in time and in frequency
- The window length determines the resolution of the spectrogram

Power spectrum density estimation

- The PSD is the DTFT of the autocorrelation function
- The PSD may be two-sided or one-sided. Careful with conventions!
- The periodogram method estimates the PSD directly from the magnitude squared of the DFT of the windowed signal
- The periodogram is an biased estimator of the PSD, and it has large variance. Hence, the periodogram must be averaged to produce useful estimates
- The Welch method is a time-averaged periodogram
- The Welch method breaks the data into overalapping segments, each of length L. Usually, the segments overlap by $L/2$.
- Differently from the periodogram and Welch method, the Blackman-Tukey method estimates the PSD by computing the DFT of the estimated autocorrelation function
- Although the estimated autocorrelation function may be an unbiased estimator, the PSD estimate is biased. Windows with non-negative frequency response are typically preferred e.g., Bartlett
- Increasing the sequence length Q improves accuracy. Reducing the window length L improves accuracy at the expense of poorer frequency resolution.

Hope you enjoyed it.

Thank you.