Adaptive Signal Processing & Machine Intelligence

Lecture 3 - Spectrum Estimation

Danilo Mandic

room 813, ext: 46271

Department of Electrical and Electronic Engineering
Imperial College London, UK

d.mandic@imperial.ac.uk, URL: www.commsp.ee.ic.ac.uk/∼mandic

Outline

Part 1: Background

- Some intuition and history
- The Discrete Fourier Transform (DFT)
- Practical issues with DFT
  - Aliasing
  - Frequency resolution
  - Incoherent sampling
  - Leakage
  - Time-bandwidth product

Part 2: The Periodogram and its modifications

- Periodogram
- The role of autocorrelation estimation
- Windowing
- Averaging
- Blackman-Tukey Method
- Statistical properties of these methods (bias, variance)

Problem Statement

From a **finite** record of stationary data sequence, **estimate** how the total power is distributed over frequency.

Has found a tremendous number of applications:-

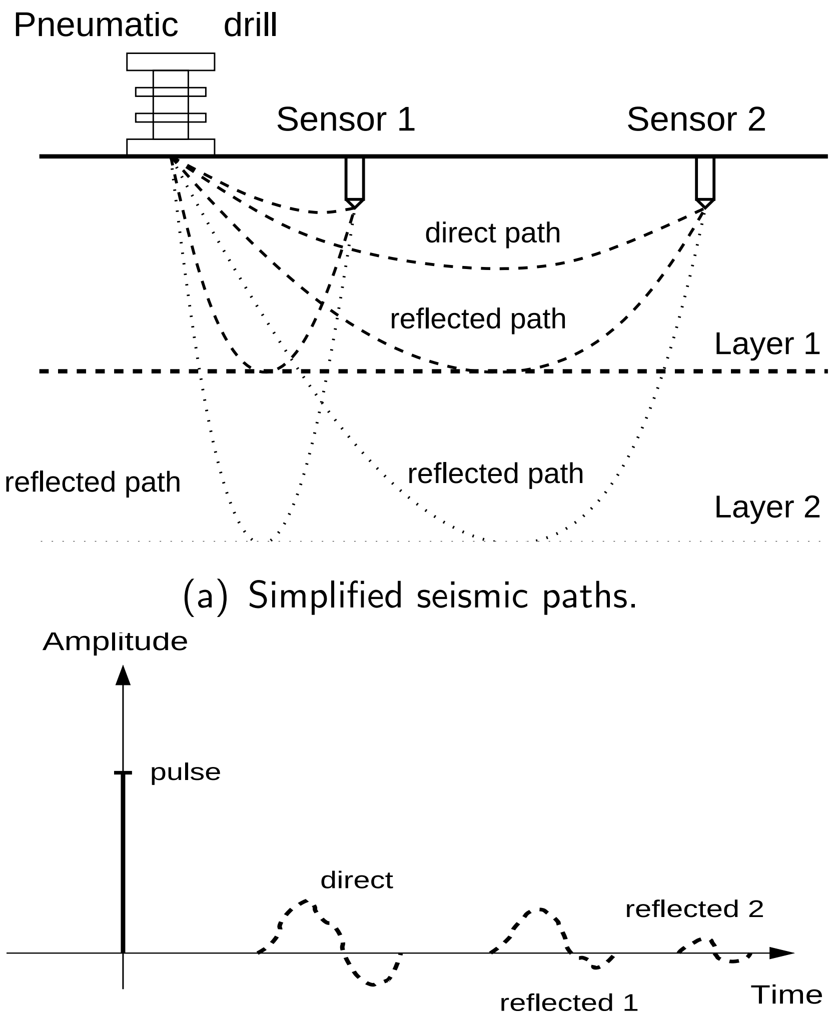
- Seismology $\rightarrow$ oil exploration, earthquake
- Radar and sonar $\rightarrow$ location of sources
- Speech and audio $\rightarrow$ recognition
- Astronomy $\rightarrow$ periodicities
- Economy $\rightarrow$ seasonal and periodic components
- Medicine $\rightarrow$ EEG, ECG, fMRI
- Circuit theory, control systems

Some examples

Seismic estimation

Speech processing

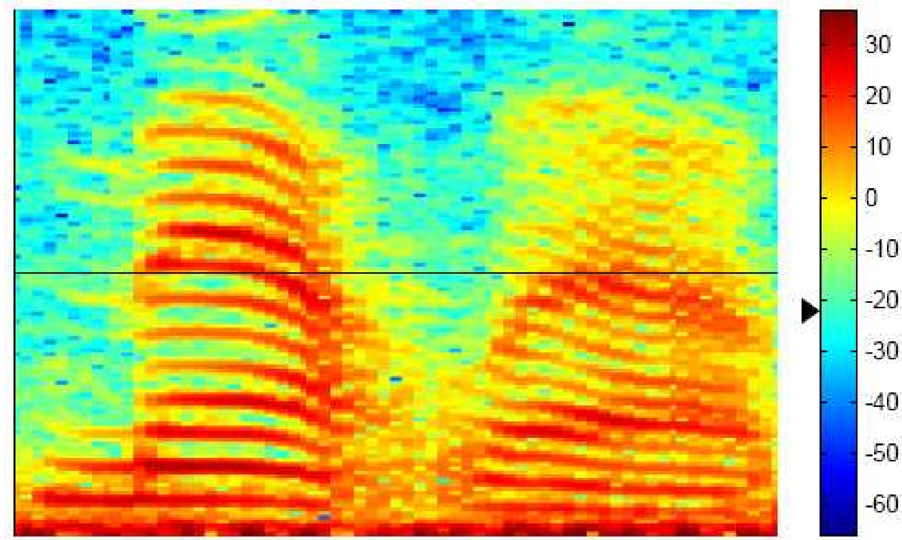
periodic pulse excitation



(a) Simplified seismic paths.

(b) Seismic impulse response.

frequency



time

M aaa t l aaa b

For every time segment 'Δt', the PSD is plotted along the vertical axis. Observe the harmonics in 'a'

Darker areas: higher magnitude of PSD (magnitude encoded in color)

Use Matlab function 'specgram'

Fourier transform & the DTFT

Fourier transform (continuous case):

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-\jmath\omega t}dt$$

Not really convenient for real–world signals $\Rightarrow$ **need for a signal model.**

More natural: Can we estimate the spectrum from N samples of $x(t)$, that is $[x[0], x[1], \ldots, x[N-1]]$ where the spacing in time is T?

One solution $\Rightarrow$ perform a rectangular approximation of the above integral.

$$X(\omega) = \sum_{k=0}^{N-1} x[k]e^{-\jmath\omega k}$$

Since ω is a continuous variable, there are an infinite number of possible values of ω from $-\pi$ to $\pi \Rightarrow$ DTFT.

We have two problems with this approach:-

i) due to the sampling of $x(t)$, aliasing for non–bandlimited signals;

ii) only N samples retained $\Rightarrow$ resolution?

Discrete Fourier Transform

Special case: we can use N uniformly spaced frequencies around the unit circle $\omega_m = \frac{2\pi m}{N}$, such that

$$X(\omega_m) = \sum_{k=0}^{N-1} x[k] e^{-\jmath \omega_m k} \qquad m \in [0, N-1]$$

Alternatively, this can be expressed as the inner product of the signal $\mathbf{x}$ and a complex sinusoidal basis $\mathbf{f}_m$ sampled at frequency m

$$X(\omega_m) = \mathbf{f}_m^H \mathbf{x} = \sum_{k=0}^{N-1} x[k] e^{-\jmath \omega_m k} \in \mathbb{C}$$

where

$$\mathbf{f}_m = \begin{bmatrix} 1 & e^{\jmath \omega_m} & e^{\jmath 2\omega_m} & e^{\jmath 3\omega_m} & \dots & e^{\jmath (N-1)\omega_m} \end{bmatrix}^T \in \mathbb{C}^N$$

Intuition: We multiply an oscillatory signal $Ae^{\jmath \omega t}$ by $e^{-\jmath \omega t}$, to obtain $Ae^{\jmath \omega t} e^{-\jmath \omega t} = A$ which is effectively a Fourier coefficient.

$\rightsquigarrow$ **Fourier transform performs demodulation.**

Discrete Fourier Transform

Matrix formulation

The DFT at N uniformly spaced frequencies can be expressed as

$$\begin{bmatrix} X(\omega_0) \\ X(\omega_1) \\ X(\omega_2) \\ \vdots \\ X(\omega_{N-1}) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{N-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{N-1} & \alpha^{2(N-1)} & \cdots & \alpha^{(N-1)^2} \end{bmatrix}^H}_{\mathbf{F}} \begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ \vdots \\ x[N-1] \end{bmatrix}$$

where $\alpha = e^{\jmath \frac{2\pi}{N}}$, or

$$[X(\omega_0), X(\omega_1), \cdots, X(\omega_{N-1})]^T = \mathbf{F}^H \mathbf{x} \in \mathbb{C}^N$$

where

$$\mathbf{F} = [\ \mathbf{f}_0\ ,\quad \mathbf{f}_1\ ,\quad \cdots\ ,\quad \mathbf{f}_{N-1}\ ] \in \mathbb{C}^{N \times N}$$

Each column of $\mathbf{F}$ represents a sinusoid with a different frequency.

Properties of the Fourier matrix:

$$\mathbf{F}^H\mathbf{F} = N\mathbf{I} \qquad \mathbf{F}^H = N\mathbf{F}^{-1}$$

Hence Fourier matrix **rotates** and **amplifies** $\mathbf{x}$.

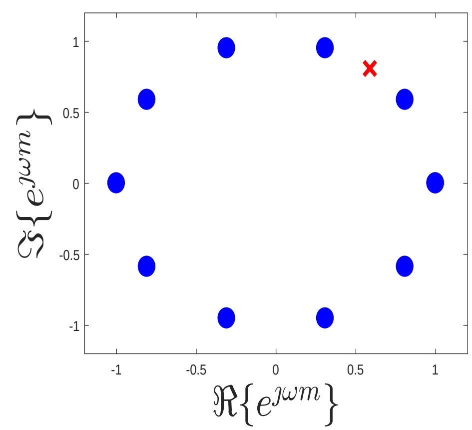
Alternatively, a **normalised** Fourier matrix with $\alpha = \frac{1}{\sqrt{N}} e^{\jmath\frac{2\pi}{N}}$, given by

$$\mathbf{F}^H\mathbf{F} = \mathbf{I} \qquad \mathbf{F}^H = \mathbf{F}^{-1} \qquad \text{(unitary)}$$

would **purely rotate** data $\mathbf{x}$.

What happens if your signal $\mathbf{x}$ cannot be represented as a sum of the uniformly spaced sinusoids?

Incoherent sampling $\Longrightarrow$ **A limitation of the DFT for a small N.**



Discrete frequencies at $2\pi/10$ (blue)

Actual frequency (red cross)

Dirichlet kernel

Issues with finite duration measurements

To analyse the effects of a finite signal duration, consider a rectangular window

$$\underbrace{\Big|\;\;\Big|\;\;\cdots\;\;\Big|}_{0,\dots,N-1} \quad \xrightarrow{\mathcal{F}} \quad \sum_{k=0}^{N-1} e^{-\jmath\omega k}$$

$$W(\omega) = \sum_{k=0}^{N-1} e^{-\jmath\omega k} = \frac{1-e^{-\jmath\omega N}}{1-e^{-\jmath\omega}} = \frac{e^{-\jmath\frac{\omega N}{2}}}{e^{-\jmath\frac{\omega}{2}}}\,\frac{2\jmath\sin(\frac{\omega N}{2})}{2\jmath sin(\frac{\omega}{2})} =$$

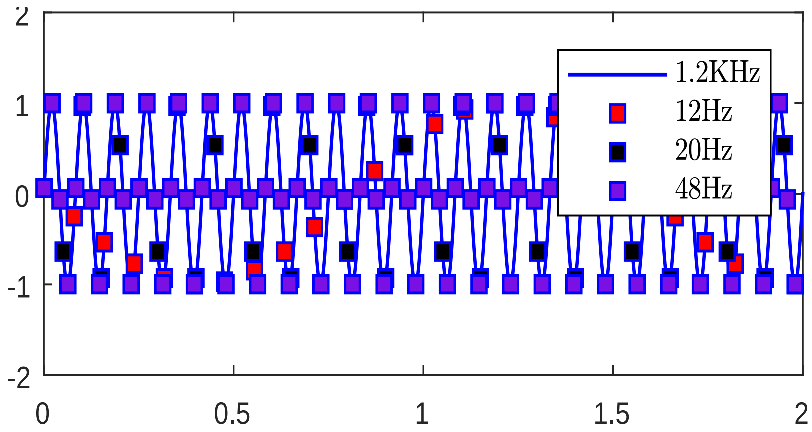
$$e^{-\jmath\frac{\omega(N-1)}{2}}\,\frac{\sin(\frac{\omega N}{2})}{\frac{\omega N}{2}} \times \frac{\frac{\omega N}{2}}{\sin(\frac{\omega}{2})} = e^{-\jmath\frac{\omega(N-1)}{2}}\,\frac{sinc(\frac{\omega N}{2})}{sinc(\frac{\omega}{2})} \times N$$

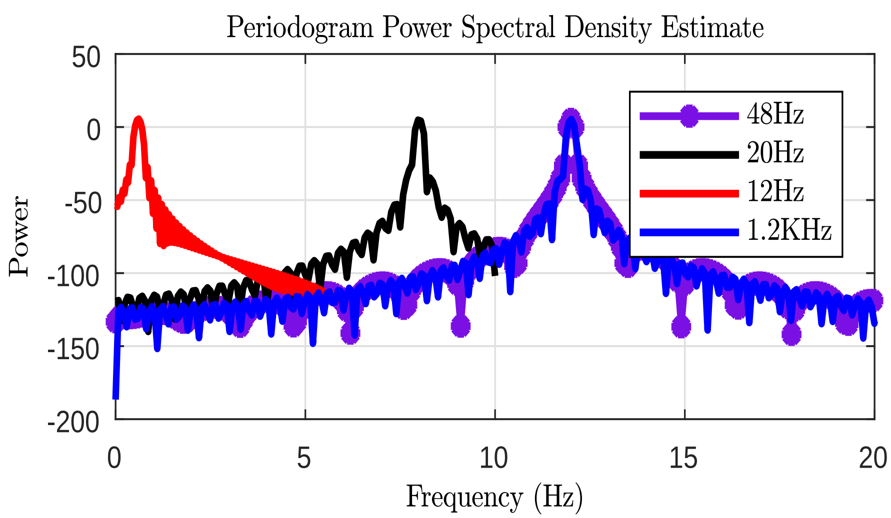
If the sampling is **coherent**, zeroes of the sinc functions all lie at multiples of $\omega = \frac{2\pi}{N}$, and hence the outputs of DFT are all zero except at $\omega = \pm\frac{2\pi}{N}$.

Practical Issue #1: Aliasing

Sampling theorem: An example

For sampling period T and sampling frequency $f_s = 1/T \Rightarrow f_s \geq 2f_h$





- Sub-Nyquist sampling causes aliasing
- This distorts physical meaning of information
- In signal processing, we require faithful data representation
- Problem: the noise model is always all-pass
- The easiest and most logical remedy is to low-pass filter the data so that the Nyquist criterion is satisfied.

Practical Issue #2: Frequency resolution

Time-bandwidth product $\hookrightarrow$ freq. bins resolvable if separated by $\frac{2\pi}{NT}$

- Suppose we know the **maximum frequency** in the signal ω_{max}, and the required resolution $\Delta\omega$. Then

$$\Delta\omega > 2\frac{2\pi}{NT} = 2\frac{\omega_s}{N} \qquad \Rightarrow \qquad N > \frac{4\omega_{max}}{\Delta\omega}$$

- For both the **prescribed resolution and bandwidth**, then $\omega_s = \frac{2\pi}{T} > 2\omega_{max}$ and $2\omega_s < \Delta\omega N$, hence $\frac{f_s}{2} = \frac{\pi}{T} > \omega_{max}$, that is

$$T < \frac{\pi}{\omega_{max}} \;\; \Leftrightarrow \;\; N > \frac{4\omega_{max}}{\Delta\omega}$$

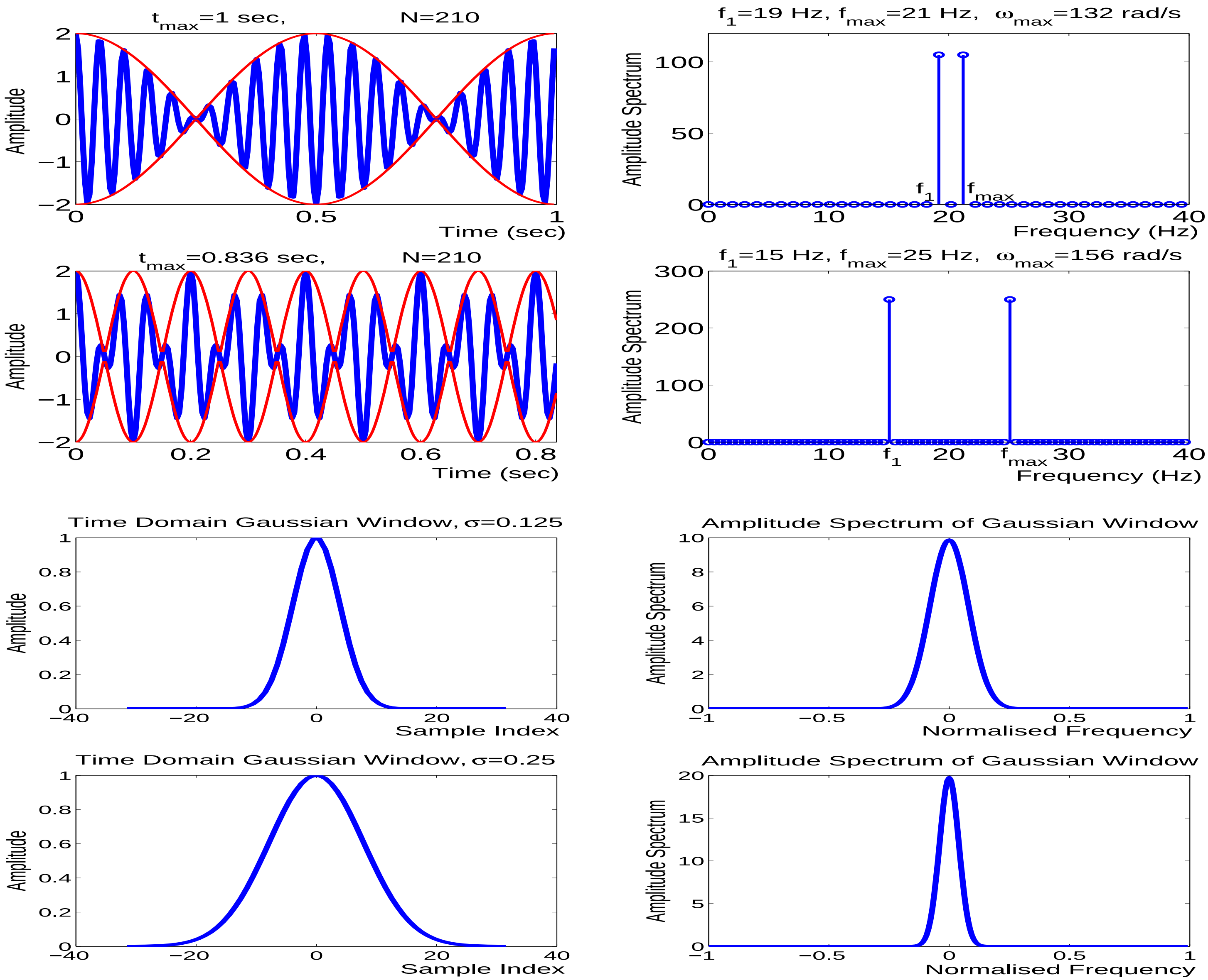
- For known **signal duration** $\left(f_s \geq 2f_{max} \Rightarrow \frac{2\pi}{T} \geq 2\omega_{max} \Rightarrow T < \frac{\pi}{\omega_{max}}\right)$

$$N > \frac{2t_{max}}{T} \qquad \Rightarrow \qquad N > \frac{2t_{max}\omega_{max}}{\pi}$$

$t_{max} \times \omega_{max} \rightarrow$ **time–bandwidth product of a signal.**

Example: the time–bandwidth product

Top: AM signals Bottom: Gaussian signals



Practical Issue #3: Spectral leakage

Two sines with close frequencies

Top: A 32-point DFT of an $N = 32$ long sampled ($f_s = 64\text{Hz}$) mixed sinewave

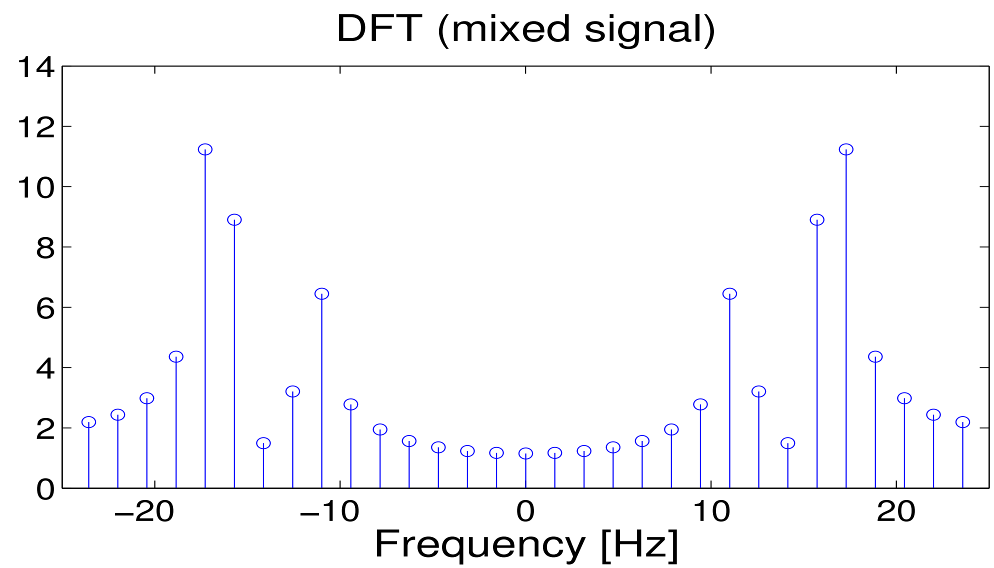
$$x[k] = \sin(2\pi 11k) + \sin(2\pi 17k)$$

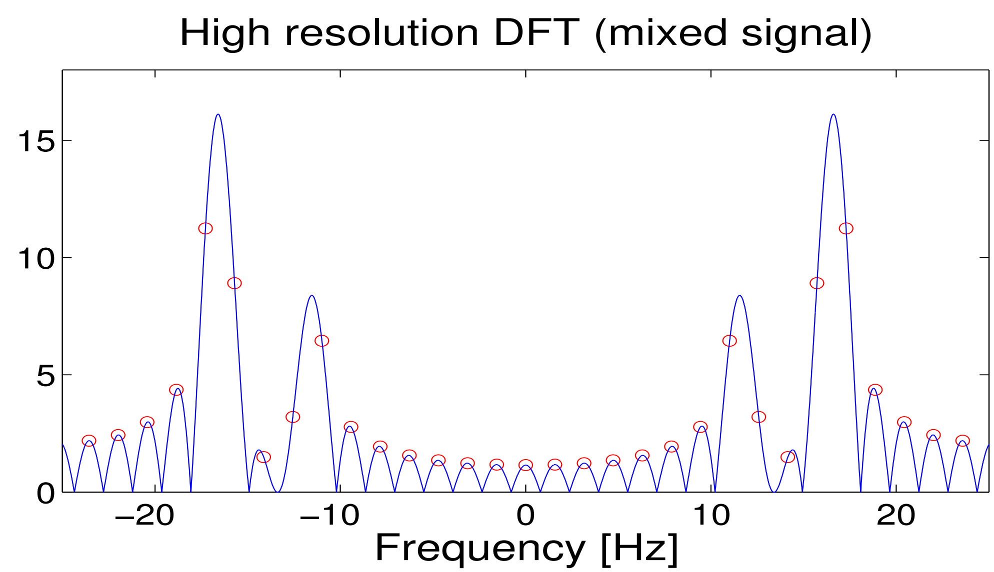
It is difficult to determine how many distinct sinewawes we have.

Bottom: A 3200-point DFT of an $N = 32$ long sampled ($f_s = 64\text{Hz}$) sine

$$x[k] = \sin(2\pi 11k) + \sin(2\pi 17k)$$

- Both the $f = 11\text{Hz}$ and $f = 17\text{Hz}$ sinewaves appear quite sharp
- This is a consequence of a high-resolution ($N = 3200$) DFT
- The overlay plot compares it with the top diagram





Practical Issue #4: Incoherent sampling

Are the frequencies in a signal exactly at $f = \frac{k}{N} f_s$?

Top: A 32-point DFT of an $N = 32$ long sampled ($f_s = 64$Hz) sinewave of $f = 10$Hz

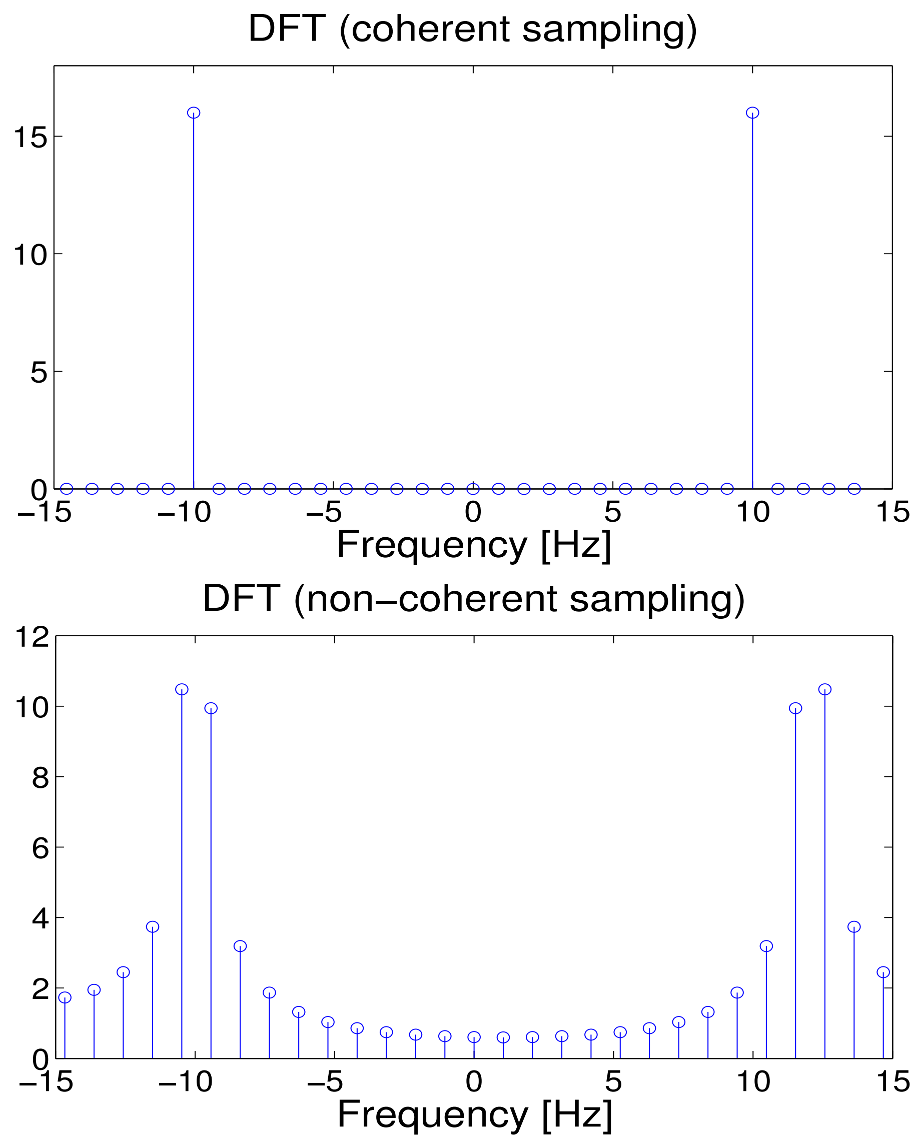
- For $f_s = 64$ Hz, the DFT bins will be located in Hz at $k/NT = 2k, \; k = 0, 1, 2, \ldots, 63$
- One of these points is at given signal frequency of 10 Hz

Bottom: A 32-point DFT of an $N = 32$ long sampled ($f_s = 64$Hz) sine of $f = 11$Hz

- Since

$$\frac{f_R}{f_s} = \frac{f \times N}{f_s} = \frac{11 \times 32}{64} = 5.5$$

  the impulse at $f = 11$ Hz appears between the DFT bins $k = 5$ and $k = 6$
- The impulse at $f = -11$ Hz appears between DFT bins $k = 26$ and $k = 27$ (10 and 11 Hz)



Power Spectrum estimation: Problem statement

Estimate Power Spectral Density (PSD) of a wide-sense stationary signal

Recall that PSD $= \mathcal{F}(ACF)$.

Therefore, estimating the power spectrum is equivalent to estimating the autocorrelation.

Recall that for an autocorrelation ergodic process,

$$\mathbf{r}_{xx} = \lim_{N\to\infty} \left\{ \frac{1}{2N+1} \sum_{n=-N}^{N} x[n+k]x[n] \right\} [k]$$

If $x[k]$ is known for all n, estimating the power spectrum is straightforward

- **Difficulty 1:** the amount of data is **always limited**, and may be very small (genomics, biomedical)
- **Difficulty 2:** real world data is **almost invariably corrupted by noise**, or contaminated with an interfering signal

PSD properties

i) $P_{xx}(\omega)$ is a **real** function $(P_{xx}(\omega) = P^*_{xx}(\omega))$.

Proof: Since $\mathbf{r}_{xx}[k] = \mathbf{r}_{xx}[-k]$, we have

$$P_{xx}(\omega) = \mathcal{F}\{\mathbf{r}_{xx}\}(\omega) = \sum_{k=-\infty}^{\infty} \mathbf{r}_{xx}[k]e^{-\jmath\omega k} = \sum_{k=-\infty}^{\infty} \mathbf{r}_{xx}[-k]e^{\jmath\omega k}$$

and hence it has ***no notion of the phase information in data***

$$P_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m)\cos(2\pi m f) = r_{xx}(0) + 2\sum_{m=1}^{\infty} r_{xx}(m)\cos(2\pi m f)$$

ii) $P_{xx}(\omega)$ is a **symmetric** function $P_{xx}(-\omega) = P_{xx}(\omega)$. This follows from the last expression.

iii) $\mathbf{r}_{xx}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\omega)d\omega = E\{x^2[n]\} \geq 0.$

$\Rightarrow$ **the area below the PSD (power spectral density) curve = Signal Power**

Periodogram based estimation of power spectrum

more intuition $\looparrowright$ connection with DFT $(|\Sigma_k|^2 = \Sigma_k \Sigma_l)$

$$P_{xx}(\omega) = \sum_{k=-\infty}^{+\infty} \mathbf{r}_{xx}[k] e^{-\jmath \omega k}$$

$$P_{xx}(\omega) = \lim_{N \to +\infty} \frac{1}{2N+1} E\left\{ \sum_{k=-N}^{N} \sum_{l=-N}^{N} x[k] x[l] e^{-\jmath \omega (k-l)} \right\}$$

$$P_{xx}(\omega) = \lim_{N \to +\infty} \frac{1}{2N+1} E\left\{ \left| \sum_{k=-N}^{N} x[k] e^{-\jmath \omega k} \right|^2 \right\}$$

In practice, we only have access to $[x[0], \ldots, x[N-1]]$ data points (we drop the expectation), and provided the autocorrelation function decays fast enough, we have

$$\hat{P}_{per}(\omega_m) = \frac{1}{N} \left| \sum_{k=0}^{N-1} x[k] e^{-\jmath \omega_m k} \right|^2 = \frac{1}{N} \left| X(\omega_m) \right|^2 = \frac{1}{N} \mathbf{f}_m^H \mathbf{x} \mathbf{x}^\mathbf{T} \mathbf{f}_m$$

Symbol $\hat{(\cdot)}$ denotes an estimate, since due to the finite N the ACF is imperfect

What to look for next?

- We must examine the statistical properties of the periodogram estimator
- For the general case, the statistical analysis of the periodogram is intractable
- We can, however, derive the mean of the periodogram estimator for any real process
- The variance can only be derived for the special case of a real zero–mean WGN process with $P_{xx}(\omega) = \sigma_x^2$
- Can this can be used as indication of the variance of the periodogram estimator for other random signals
- Can we use our knowledge about the analysis of various estimators, to treat the periodogram in the same light (is it an MVU estimator, does it attain the CRLB)
- Can we make a compromise between the bias and variance in order to obtain a mean squared error (MSE) estimator of power spectrum?

Physical intuition: Connecting PSD and ACF
positive (semi)-definiteness

Recall: $\mathbf{R}_{xx} = E\{\mathbf{x}\mathbf{x}^T\} = \begin{bmatrix} \mathbf{r}_{xx}[0] & \mathbf{r}_{xx}[1] & \cdots & \mathbf{r}_{xx}[N-1] \\ \mathbf{r}_{xx}[1] & \mathbf{r}_{xx}[0] & \cdots & \mathbf{r}_{xx}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_{xx}[N-1] & \mathbf{r}_{xx}[N-2] & \cdots & \mathbf{r}_{xx}[0] \end{bmatrix}$

Then, for a linear system with input sequence $\{x\}$, output $\{y\}$, and the vector of coefficients $\mathbf{a}$, the output has the form

$$y[n] = \sum_{k=0}^{N-1} a[k]x[n-k] = \mathbf{x}^T\mathbf{a} = \mathbf{a}^T\mathbf{x} \quad \text{where} \quad \mathbf{a} = [a[0], \ldots, a[N-1]]^T$$

The power $P_y = E\{y^2\}$ is **always** positive, and thus $((\mathrm{a}^T\mathrm{b})^T = \mathrm{b}^T\mathrm{a}^T)$

$$E\big\{y^2[n]\big\} = E\big\{y[n]y^T[n]\big\} = E\big\{\mathbf{a}^T\mathbf{x}\mathbf{x}^T\mathbf{a}\big\} = \mathbf{a}^T E\big\{\mathbf{x}\mathbf{x}^T\big\}\mathbf{a} = \mathbf{a}^T\mathbf{R}_{xx}\mathbf{a}$$

$\Rightarrow$ **to maintain positive power, the autocorrelation matrix R_{xx} must be positive semidefinite**

In other words: a positive semidefinite R_{xx} will alway produce positive power spectrum!

But, is our estimate of ACF always positive definite?

Periodogram and Matlab

```
Px=abs(fft(x(n1:n2))).^2/(n2-n1-1)
```

or the direct command '**periodogram**'

- `Pxx = PERIODOGRAM(X)`
  returns the PSD estimate of the signal specified by vector X in the vector Pxx. By default, the signal X is windowed with a BOXCAR window of the same length as X;

- `PERIODOGRAM(X,WINDOW)`
  specifies a window to be applied to X. WINDOW must be a vector of the same length as X;

- `[Pxx,W] = PERIODOGRAM(X,WINDOW,NFFT)`
  specifies the number of FFT points used to calculate the PSD estimate.

Performance of the periodogram
(we desire a minimum variance unbiased (MVU) est.)

Its performance is analysed in the same was as the performance of any other estimator:

- **Bias**, that is, whether

$$\lim_{N\to\infty} E\left\{\hat{P}_{per}(\omega_m)\right\} = P_{xx}(\omega_m)$$

- **Variance**

$$\lim_{N\to\infty} Var\left\{\hat{P}_{per}(\omega_m)\right\} \qquad \text{(ideally this goes to 0)}$$

- **Mean square convergence**

$$\text{MSE} = \text{bias}^2 + \text{variance} = E\left\{\left[\hat{P}_{per}(\omega_m) - P_{xx}(\omega_m)\right]^2\right\}$$

$$\text{we desire} \quad \lim_{N\to\infty} E\left\{\left[\hat{P}_{per}(\omega_m) - P_{xx}(\omega_m)\right]^2\right\} = 0$$

☞ we need to check $\hat{P}_{per}(\omega_m)$ is a **consistent** estimator of the true PSD.

Bias of the periodogram as an estimator

We can calculate this by finding the expected value of

$$\hat{\mathbf{r}}_{xx}[k] = \frac{1}{N} \sum_{n=0}^{N-1-|k|} x[n]x[n+|k|]$$

Thus (**biased estimate**)

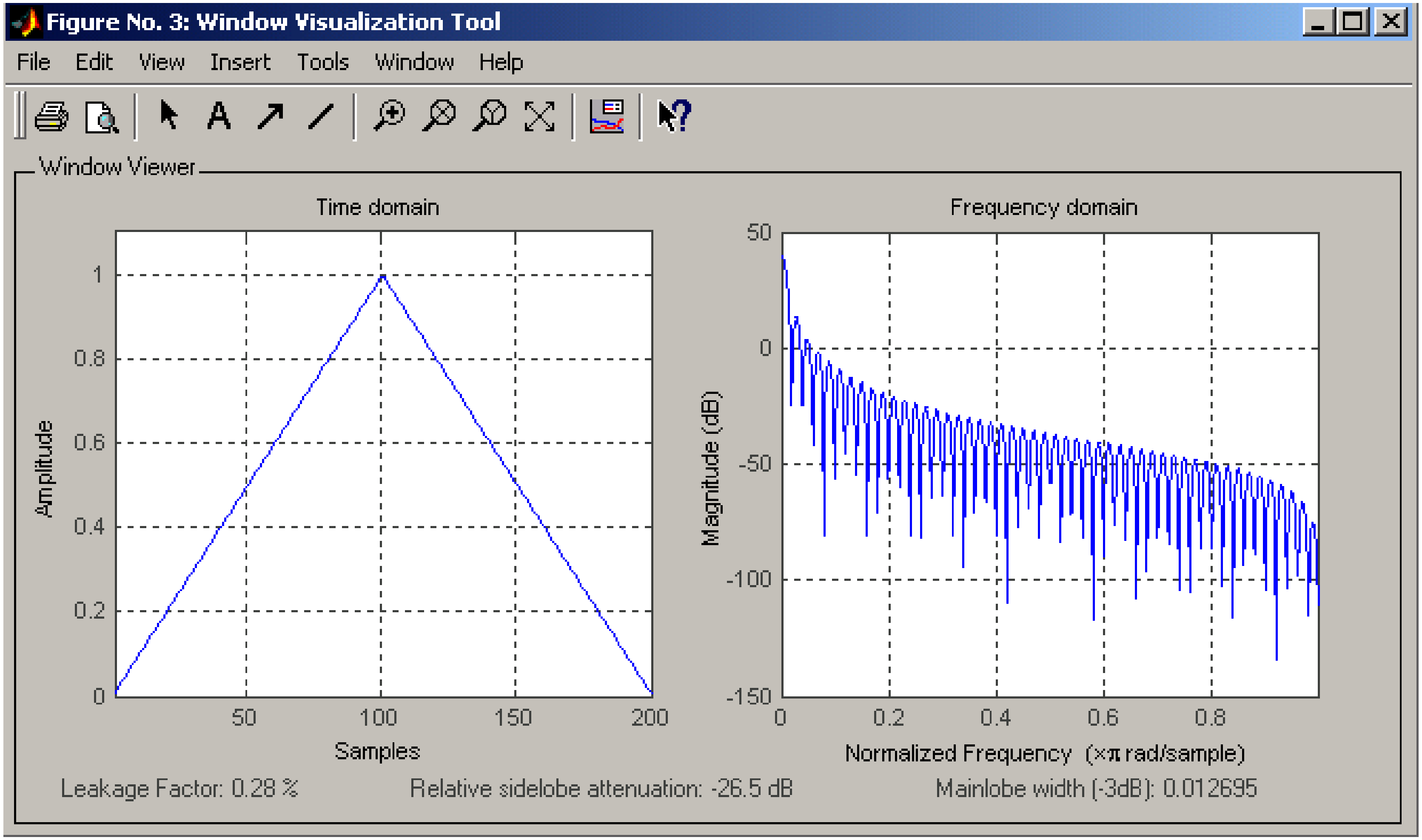
$$\begin{aligned} E\left\{P_{per}(\omega_m)\right\} &= \sum_{k=-(N-1)}^{N-1} E\{\hat{\mathbf{r}}_{xx}[k]\}e^{-\jmath\omega_m k} \\ &= \sum_{k=-(N-1)}^{N-1} \frac{N-|k|}{N}\, \mathbf{r}_{xx}[k]e^{-\jmath\omega_m k} = \text{``}\mathbf{w}_B[k] \;\times\; \mathbf{r}_{xx}[k]'' \end{aligned}$$

where $\mathbf{r}_{xx}$ is the **true ACF and the Bartlett window**, $\mathbf{w}_B[k]$, is defined as

$$\mathbf{w}_B[k] = \begin{cases} 1 - \frac{|k|}{N}; & |k| \leq N \\ 0; & |k| > N-1 \end{cases}$$

Notice the maximum at n=0, and a slow decay towards the end of the sequence

Effects of the Bartlett window on resolution



Behaves as $sinc^2$

Periodogram bias – continued

From the previous observation, we have

$$E\left\{\hat{P}_{per}(\omega_m)\right\} \;=\; \sum_{k=-N-1}^{N-1} \mathbf{r}_{xx}[k]\mathbf{w}_B[k]e^{-\jmath\omega_m k} \Leftrightarrow W_B(\omega_m) \;*\; P_{xx}(\omega_m)$$

where $W_B(\omega_m) = \frac{1}{N}\left[\frac{\sin\frac{\omega_m N}{2}}{\sin\frac{\omega_m}{2}}\right]^2$.

In words, the expected value of the periodogram is the **convolution** of the power spectrum $P_{xx}(\omega_m)$ with the Fourier transform of the Bartlett window, and therefore, the periodogram is a **biased** estimate.

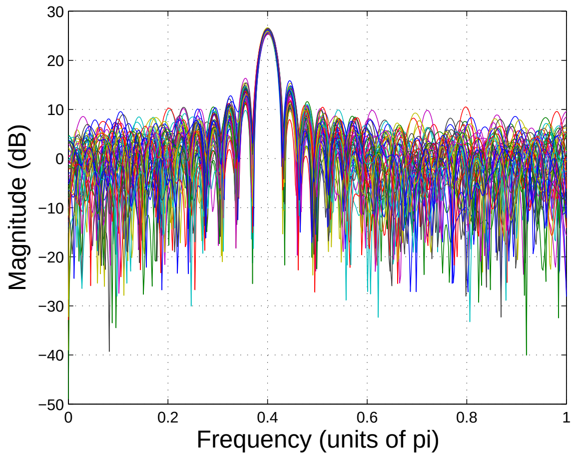
Since when $N \to \infty$, $W_B(\omega_m) \to \delta(0)$, the periodogram is **asymptotically unbiased**

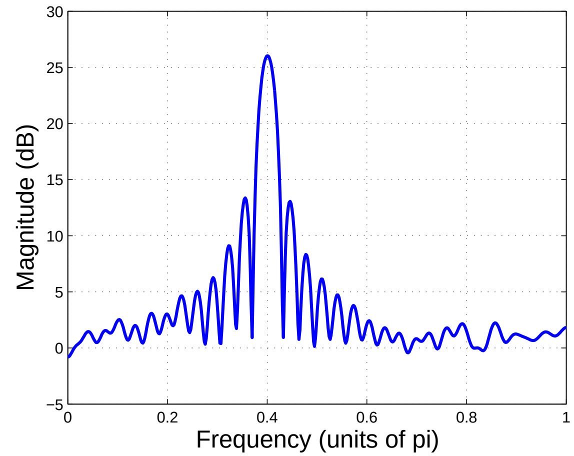
$$\lim_{N\to\infty} E\left\{\hat{P}_{per}(\omega_m)\right\} = P_{xx}(\omega_m)$$

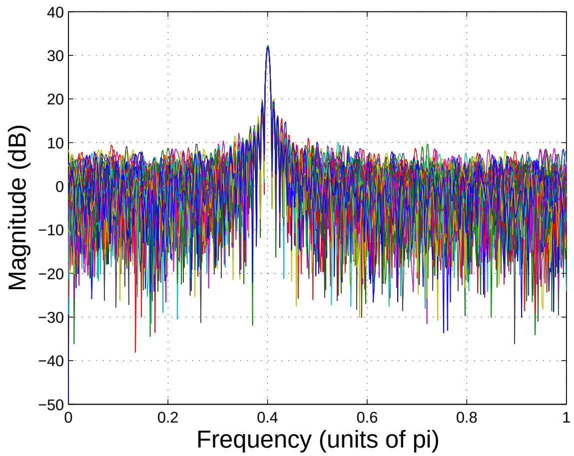
Example: Sinusoid in WGN

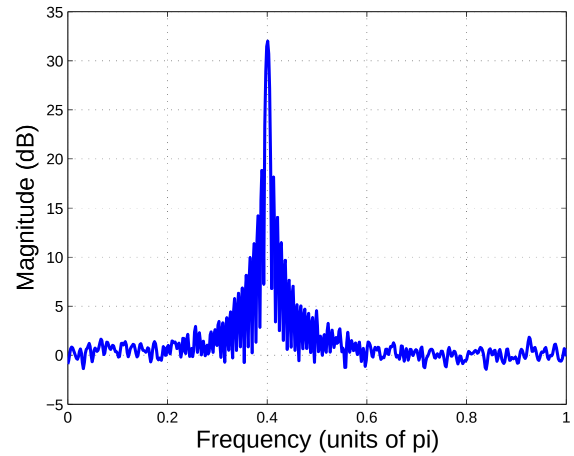
$x[n] = A\sin(\omega_0 n + \Phi) + w[n]$, $A = 5, \omega_0 = 0.4\pi$

N=64: Overlay of 50 periodograms periodogram average









N=256: Overlay of 50 periodograms periodogram average

Periodogram resolution: Two sinusoids in white noise

This is a random process ($\Phi_1 \perp \Phi_2$, $w[n] \sim \mathcal{U}(0, \sigma_w^2)$) described by :

$$x[n] = A_1 \sin(\omega_1 n + \Phi_1) + A_2 \sin(\omega_2 n + \Phi_2) + w[n]$$

The true PSD is

$$P_{xx}(\omega_m) = \ \sigma_w^2 + \frac{1}{2}\pi A_1^2 \left[\delta(\omega - \omega_1) + \delta(\omega + \omega_1)\right] + \frac{1}{2}\pi A_2^2 \left[\delta(\omega - \omega_2) + \delta(\omega + \omega_2)\right]$$

The expected PSD $E\left\{\hat{P}_{per}(\omega_m)\right\}$ $(P_{xx}(\omega_m) * W_B(\omega_m))$ becomes

$$\sigma_w^2 + \frac{1}{4}A_1^2 \left[W_B(\omega - \omega_1) + W_B(\omega + \omega_1)\right] + \frac{1}{4}A_2^2 \left[W_B(\omega - \omega_2) + W_B(\omega + \omega_2)\right]$$

☞ **there is a limit on how closely two sinusoids or two narrowband processes may be located before they can no longer be resolved.**

Example: Estimation of two sinusoids in WGN

Based on previous example, try to generate these yourselves

$$x[n] = A_1 \sin(n\omega_1 + \Phi_1) + A_2 \sin(n\omega_2 + \Phi_2) + w[n]$$

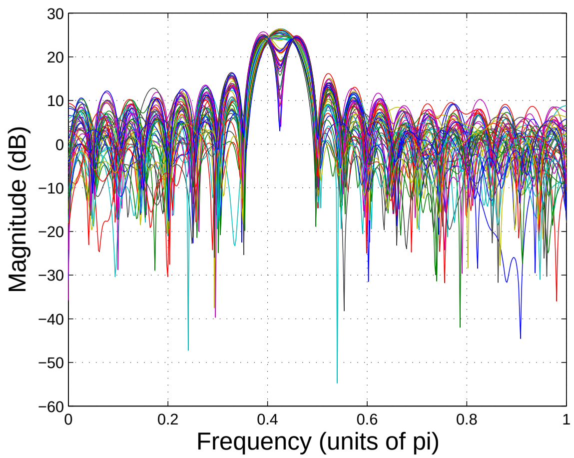
where

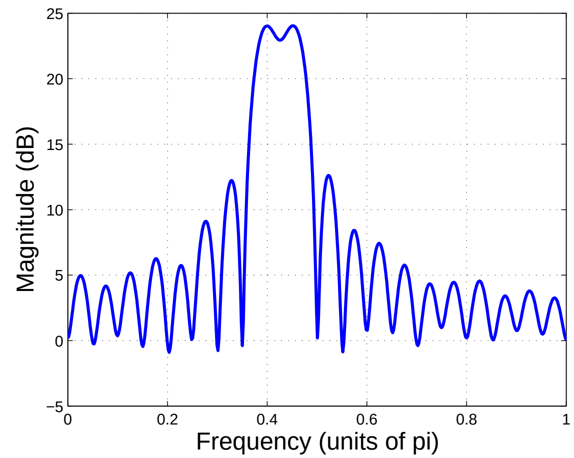
- datalength $N = 40$, $N = 64$, $N = 256$
- $A_1 = A_2$, $\omega_1 = 0.4\pi$, $\omega_2 = 0.45\pi$
- $A_1 \neq A_2$, $\omega_1 = 0.4\pi$, $\omega_2 = 0.45\pi$
- produce overlay plots of 50 periodograms and also averaged periodograms

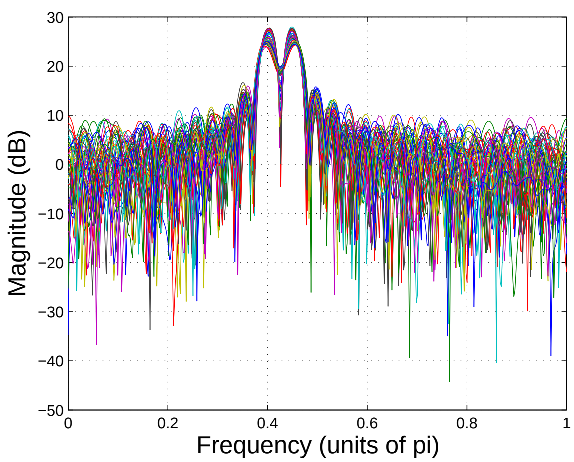
Example: Periodogram resolution $\looparrowright$ two sinusoids

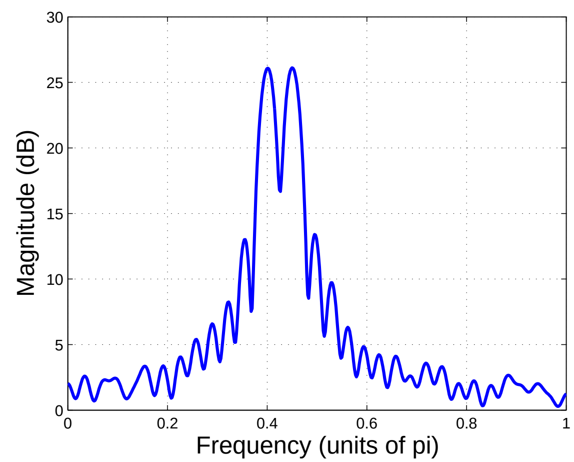
see also Problem 4.6 in your Problem/Answer set

N=40: Overlay of 50 periodograms   periodogram average









N=64: Overlay of 50 periodograms   periodogram average

Effects of the Window Choice

Recall: The spectrum of the (rectangular) window is a ***sinc*** which has a main lobe and sidelobes

All the other window functions (addressed later) also have the mainlobe and sidelobes.

- The effect of the main lobe (its width) is to **smear** or **smooth** the estimated spectrum shape
- From the previous slide: the width of the mainlobe causes the next peak in the spectrum to be masked if the two peaks are not separated by $2\pi/N \looparrowright$ the spectral resolution
- The sidelobes cause **spectral leakage** $\looparrowright$ transferring power from the correct frequency bin into the frequency bins which contain no signal power

These effects are dangerous, e.g. when estimating peaky spectra

Some observations

- The Bartlett window **biases** the periodogram;
- It also introduces **smoothing**, which **limits** the ability of the periodogram to resolve closely–spaced narrowband components in $x[n]$;
- This is due to the width of the main lobe of $W_B(\omega_m)$;
- Periodogram **averaging would reduce the variance** (remember MVU estimators!)
- **Resolution of the periodogram**
  - set $\Delta\omega =$ width of the main lobe of spectral window, at its "half power"
  - for Bartlett window $\Delta\omega \sim 0.89(2\pi/N) =$ periodogram resolution!
  - notice that the resolution is inversely proportional to the amount of data N

Variance of the periodogram

☹ it is difficult to evaluate the variance of the periodogram of an arbitrary process $x[n]$ since the variance depends on the fourth–order moments.

☺ the variance may be evaluated in the special case of WGN $\longrightarrow$

$$E\left\{\hat{P}_{per}(\omega_m)\hat{P}_{per}(\omega_m)\right\} = E\left\{X^2(\omega_m)X^{*2}(\omega_m)\right\}$$
$$= E\left\{X^2(\omega_m)\right\}E\left\{X^{*2}(\omega_m)\right\} + \left(E\left\{X(\omega_m)X^*(\omega_m)\right\}\right)^2$$

For WGN, these fourth–order moments become $E\left\{X^2(\omega_m)\right\} = 0$ and $E\left\{X(\omega_m)X^*(\omega_m)\right\} = \sigma^2$, then we have $E\left\{\hat{P}_{per}(\omega_m)\hat{P}_{per}(\omega_m)\right\} = \sigma_x^4$, and the variance of the periodogram for a given frequency becomes:

$$var\left\{\hat{P}_{per}(\omega_m)\right\} = P_{xx}^2(\omega_m)\left[1 + \left(\frac{\sin\frac{\omega_m}{N}}{N\sin\omega_m}\right)^2\right]$$

For the periodogram to be consistent, $var(P_{per}) \rightarrow 0 \; as \; N \rightarrow \infty$.
From the above, this is **not** the case $\Rightarrow$ the **periodogram estimator is inconsistent.** In fact, $var(P_{per}(\omega_m)) = P_{xx}^2(\omega_m) \quad \nrightarrow \quad$ quite large

Properties of the standard periodogram

Functional relationship:

$$\hat{P}_{per}(\omega_m) = \frac{1}{N}\left|\sum_{k=0}^{N-1} x[k]e^{-\jmath\omega_m k}\right|^2$$

- **Bias**

$$E\left\{\hat{P}_{per}(\omega_m)\right\} = \frac{1}{2\pi}P_{xx}(\omega_m) \ * \ W_B(\omega_m)$$

- **Resolution**

$$\Delta\omega = 0.89\frac{2\pi}{N}$$

- **Variance**

$$Var\left\{\hat{P}_{per}(\omega_m)\right\} \approx P_{xx}^2(\omega_m)$$

Bias vs variance

Recall that for any estimator, its mean square error (MSE) is given by:

$$\mathrm{MSE} = \mathrm{bias}^2 + \mathrm{variance}$$

A way to overcome periodogram limitations:

- bias performance must be traded for variance performance
- the dataset is divided up into independent blocks
- the periodograms for every block may be averaged
- the resultant estimator is termed the **averaged periodogram**

$$\hat{P}_{aver,per}(\omega_m) = \frac{1}{L}\sum_{l=0}^{L-1} \hat{P}_{per}^{(l)}(\omega_m)$$

From Estimation Theory: averaging of random trials reduces noise power!

Periodogram modifications $\looparrowright$ some intuition

Clearly, we need to reduce the variance of the periodogram, since in general they are not adequate for precise estimation of PSD.

We can think of several modifications:

1) **averaging over a set of periodograms** (we have already seen the effect of this in some simulations).

   Recall that from the general estimation theory, by averaging M times we have the effect of $var \rightarrow var/M$.

2) **applying different windows** $\looparrowright$ it is possible to choose or design a window which will have a narrow mainlobe

3) **overlapping windowed segments** for additional variance reduction $\looparrowright$ averaging periodograms along one realisation of a random process (instead of across the ensemble)

Overview of Periodogram Modifications

Periodogram

$$\hat{P}_{per}(\omega_m) = \frac{1}{N}\left|\sum_{k=0}^{N-1} x[k]e^{-\jmath\omega_m k}\right|^2$$

Windowing
Modified Periodogram

$$\hat{P}_{mod}(\omega_m) = \frac{1}{NU}\left|\sum_{k=0}^{N-1} w[k]x[k]e^{-\jmath\omega_m k}\right|^2$$

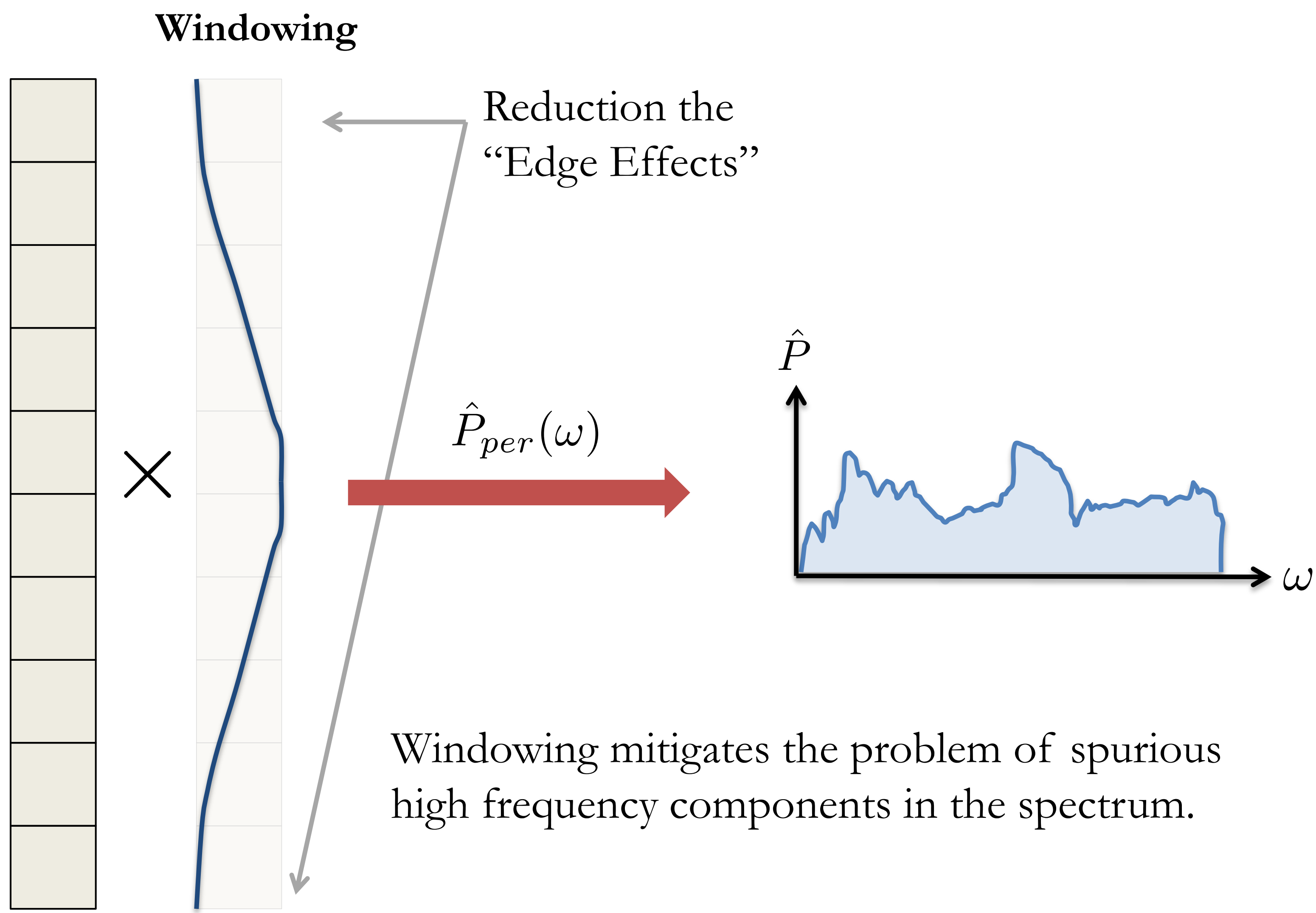
Averaging
Bartlett's Method

$$\hat{P}_{B}(\omega_m) = \frac{1}{N}\sum_{i=0}^{K-1}\left|\sum_{k=0}^{N-1} x[k+iL]e^{-\jmath\omega_m k}\right|^2$$

+ Overlapping windows
Welch's Method

$$\hat{P}_{W}(\omega_m) = \frac{1}{KLU}\sum_{i=0}^{K-1}\left|\sum_{k=0}^{N-1} w[k]x[k+iD]e^{-\jmath\omega_m k}\right|^2$$

Windowing: The Modified Periodogram



Windowing mitigates the problem of spurious high frequency components in the spectrum.

The Modified Periodogram

The periodogram of a process that is windowed with a suitable general window $w[n]$ is called a **modified periodogram** and is given by:

$$\hat{P}_M(\omega_m) = \frac{1}{NU} \left| \sum_{k=0}^{N-1} x[k] w[k] e^{-\jmath \omega_m k} \right|^2$$

where N is the window length and $U = \frac{1}{N} \sum_{n=0}^{N-1} |w[n]|^2$ is a constant, **and is defined so that $\hat{P}_M(\omega_m)$ is asymptotically unbiased.**

In Matlab:

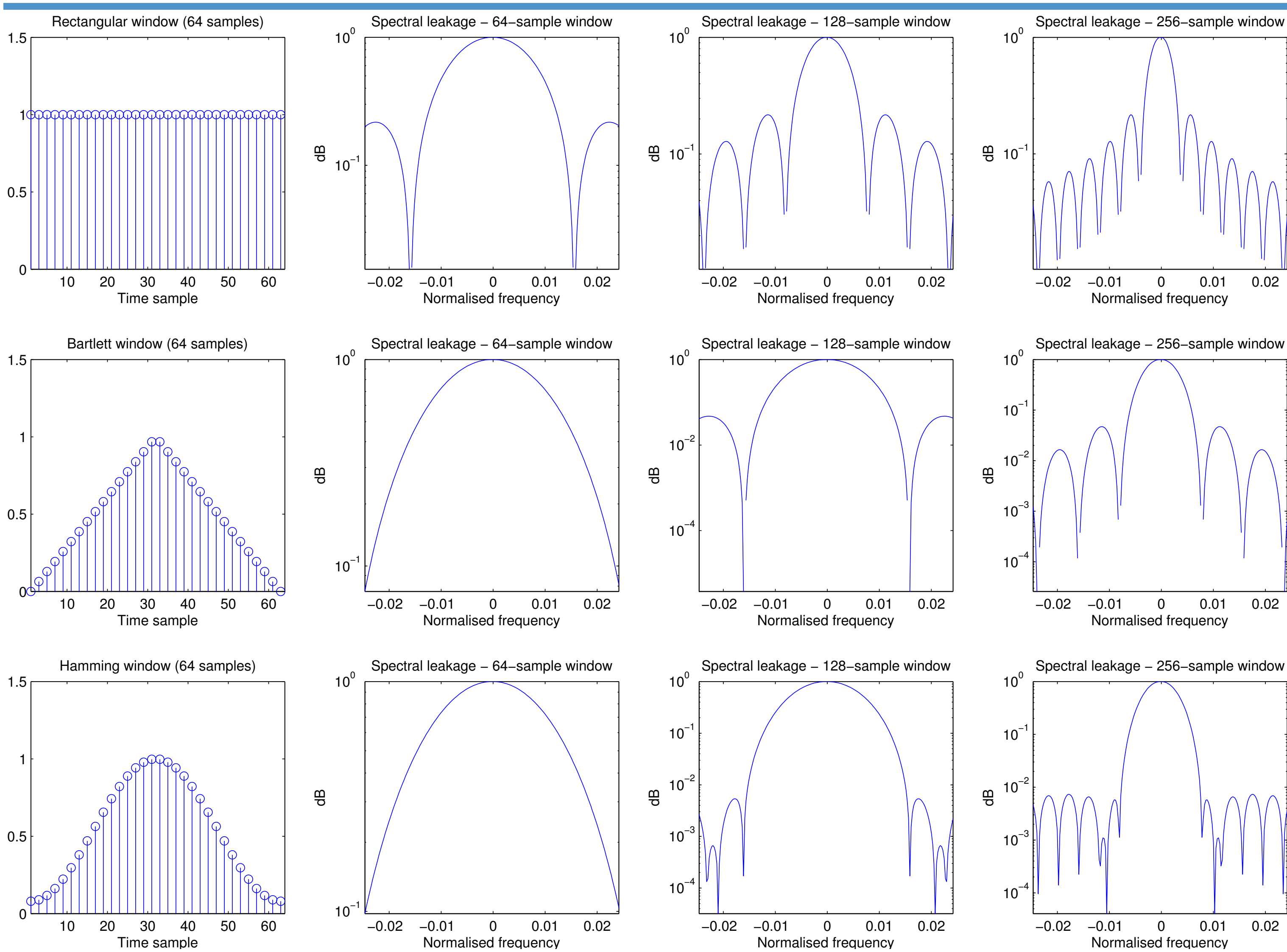
```
xw=x(n1:n2).*w/norm(w);
Pm=N * periodogram(xw);
```

where, for different windows

```
w=hanning(N); w=bartlett(N);w=blackman[n];
```

Some common windows for different window lengths:

Time domain | **Spectrum N=64** | **Spectrum N=128** | **Spectrum N=256**



The Modified Periodogram – "Windowing"

Recall that

$$Periodogram \quad \sim \quad \mathcal{F}\left(|x[n]w_r[n]|^2\right)$$

Therefore: The amount of smoothing in the periodogram is determined by the window that is applied to the data. For instance, a rectangular window has a narrow main lobe (and hence least amount of spectral smoothing), but its relatively large sidelobes may lead to masking of weak narrowband components.

Question: Would there be any benefit of using a different data window on the bias and resolution of the periodogram.

Example: can we differentiate between the following two sinusoids for $\omega_1 = 0.2\pi, \omega_2 = 0.3\pi, N = 128$

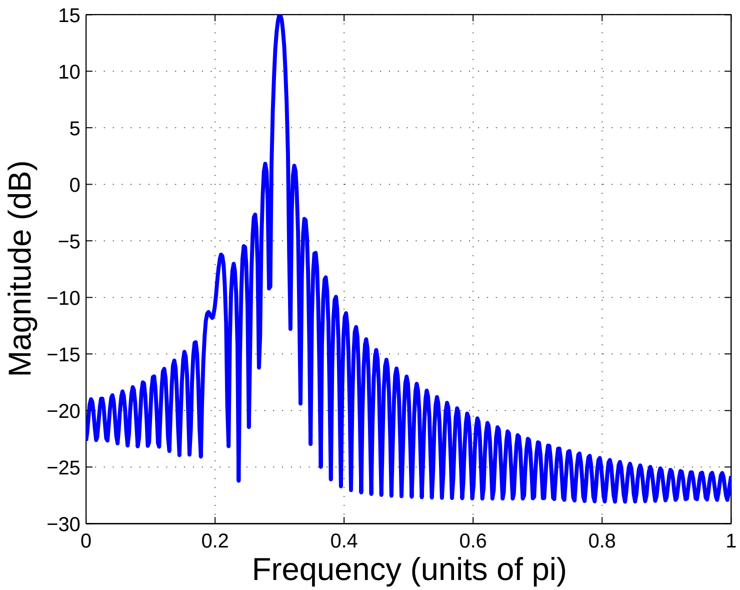
$$x[n] = 0.1\sin(n\omega_1 + \Phi_1) + \sin(n\omega_2 + \Phi_2) + w[n]$$

Example: Estimation of two sinusoids in WGN
Modified periodogram using Hamming window

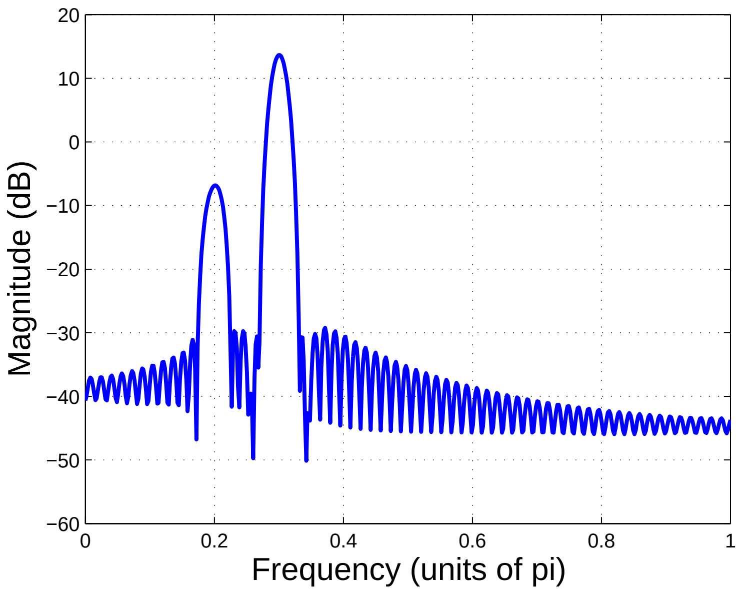
Problem: Estimate spectra of the following two sinusoids using: (a) The standard periodogram; (b) Hamming-windowed periodogram

$$x[n] = 0.1\sin(n * 0.2\pi + \Phi_1) + \sin(n * 0.3\pi + \Phi_2) + w[n] \qquad N = 128$$

Hamming window $$w[n] = 0.54 - 0.46\cos\left(2\pi\frac{n}{N}\right)$$



Expected value of periodogram



Periodogram using Hamming window

Performance of the modified periodogram

- **Bias:** Since

$$U = \frac{1}{N}\sum_{n=0}^{N-1}|w[n]|^2 = \frac{1}{N}\int_{-\pi}^{\pi}|W(\omega)|^2 d\omega \quad \Rightarrow \quad \frac{1}{2\pi NU}\int_{-\pi}^{\pi}|W(\omega)|^2 d\omega = 1$$

for $N \to \infty$ the modified periodogram is asymptotically unbiased.

- **Variance:** Since $\hat{P}_M$ is simply $\hat{P}_{per}$ of a windowed data sequence

$$Var\left\{\hat{P}_M(\omega_m)\right\} \approx P_{xx}^2(\omega_m)$$

$\Rightarrow$ **not a consistent estimate** of the power spectrum, and the data window offers no benefit in terms of reducing the variance

- **Resolution:** Data window provides a trade–off between spectral resolution (main lobe width) and spectral masking (sidelobe amplitude).

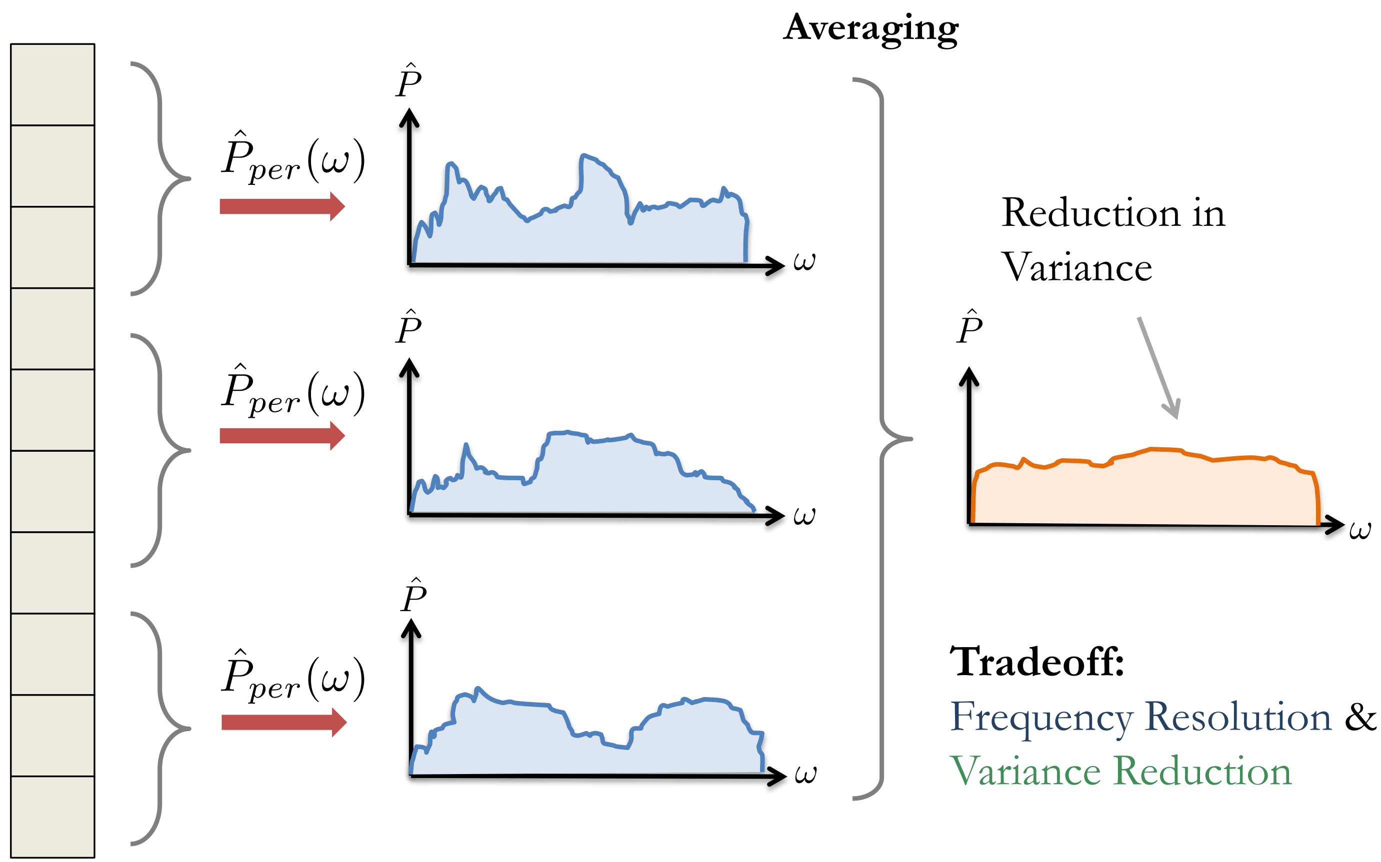
Periodogram modifications: Effects of different windows

Properties of several commonly used windows with length N:

- **Rectangular** – Sidelobe level = -13 [dB], $3\ dB$ BW $\rightarrow 0.89(2\pi/N)$
- **Bartlett** – Sidelobe level = -27 [dB], $3\ dB$ BW $\rightarrow 1.28(2\pi/N)$
- **Hanning** – Sidelobe level = -32 [dB], $3\ dB$ BW $\rightarrow 1.44(2\pi/N)$
- **Hamming** – Sidelobe level = -43 [dB], $3\ dB$ BW $\rightarrow 1.30(2\pi/N)$
- **Blackman** – Sidelobe level = -58 [dB], $3\ dB$ BW $\rightarrow 1.68(2\pi/N)$

Notice the relationship between the sidelobe level and bandwidth!

Bartlett's Method



Partitioning $x[n]$ into K non–overlapping segments

This way, the total length $N = K \times L$

Bartlett's method: Averaging periodograms

The **averaged** periodogram can be expressed as:

$$\hat{P}_{aver,per}(\omega_m) = \frac{1}{K}\sum_{l=1}^{K}\hat{P}_{per}^{(l)}(\omega_m)$$

where for each of the K segments, the segment-wise PSD estimate $P_{per}^{(i)},\ i = 1,\ldots,K$ is given by

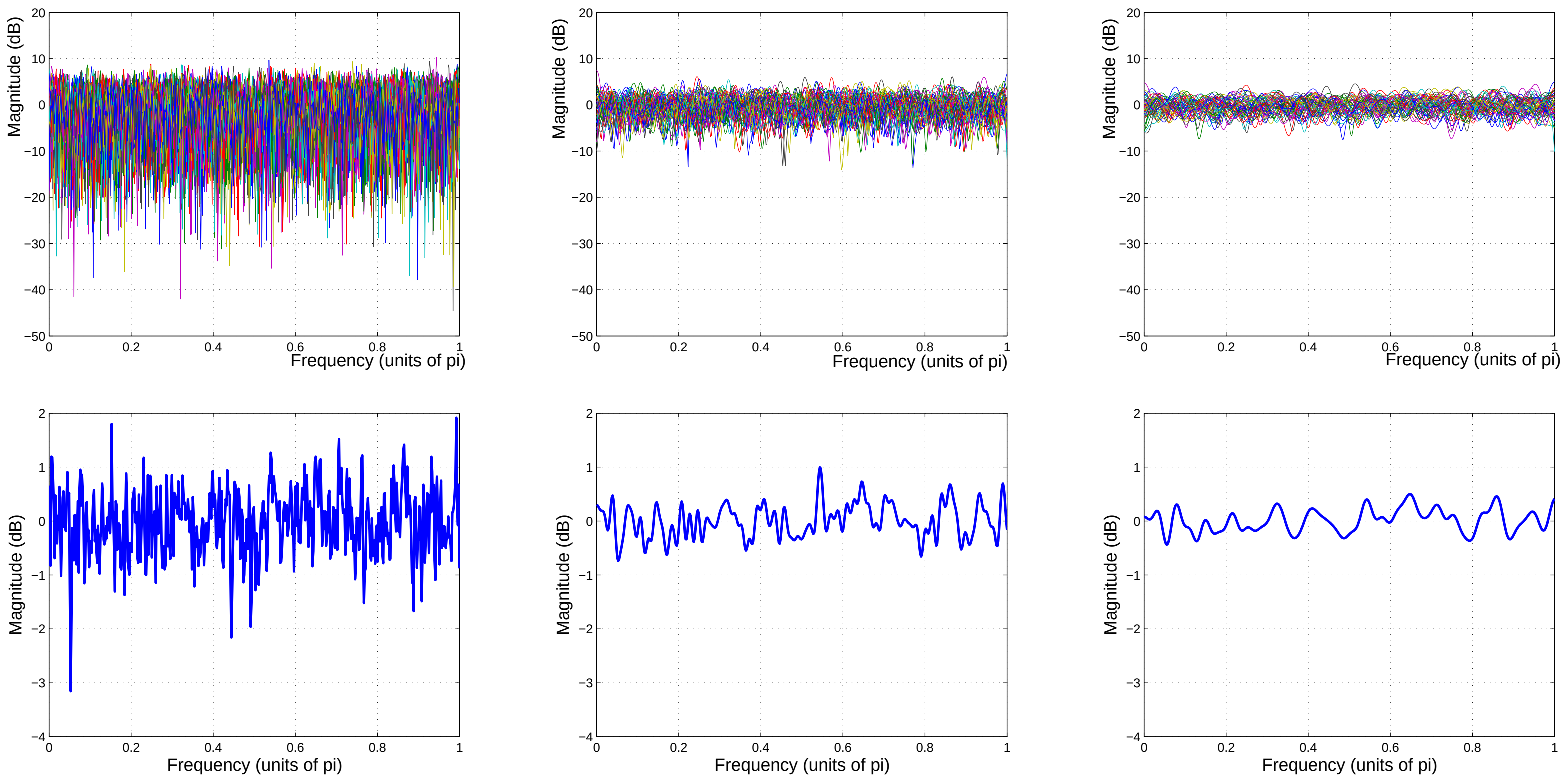
$$P_{per}^{(i)}(\omega_m) = \frac{1}{L}\left|\sum_{k=0}^{L-1} x_i[n]e^{-\jmath\omega_m k}\right|^2$$

- Idea: to reduce the variance by the factor of "K" $=$ total number of blocks
- Therefore: provided that the blocks are statistically independent (not often the case in practice) we desire to have

$$var\left\{\hat{P}_{aver,per}(\omega_m)\right\} = \frac{1}{K}var\left\{\hat{P}_{per}(\omega_m)\right\}$$

Example: Estimation of WGN spectrum using Bartlett's method

50 periodograms with $N = 512$	50 Bartlett estimates $K = 4,\ L = 128$	50 Bartlett estimates $K = 8,\ L = 64$



Ensemble averages

Performance evaluation of Bartlett's method

- **Bias:** The expected value of Bartlett's estimate

$$E\left\{\hat{P}_B(\omega_m)\right\} = \frac{1}{2\pi} P_{xx}(\omega_m) \ * \ W_B(\omega_m)$$

$\Rightarrow$ **asymptotically unbiased**.

- **Resolution:** Due to K segments of length L, as a consequence we have that Res$(P_B) <$ Res(P_{per}), that is

$$Res\left[\hat{P}_B(\omega_m)\right] = 0.89\ \frac{2\pi}{L} = 0.89\ K\ \frac{2\pi}{N}$$

- **Variance:**

$$Var\left\{\hat{P}_B(\omega_m)\right\} \approx \frac{1}{K}\, Var\left\{\hat{P}_{per}^{(i)}(\omega_m)\right\} \approx \frac{1}{K}\, P_{xx}^2(\omega_m)$$

For non–white data, variance reduction is not as large as K times!

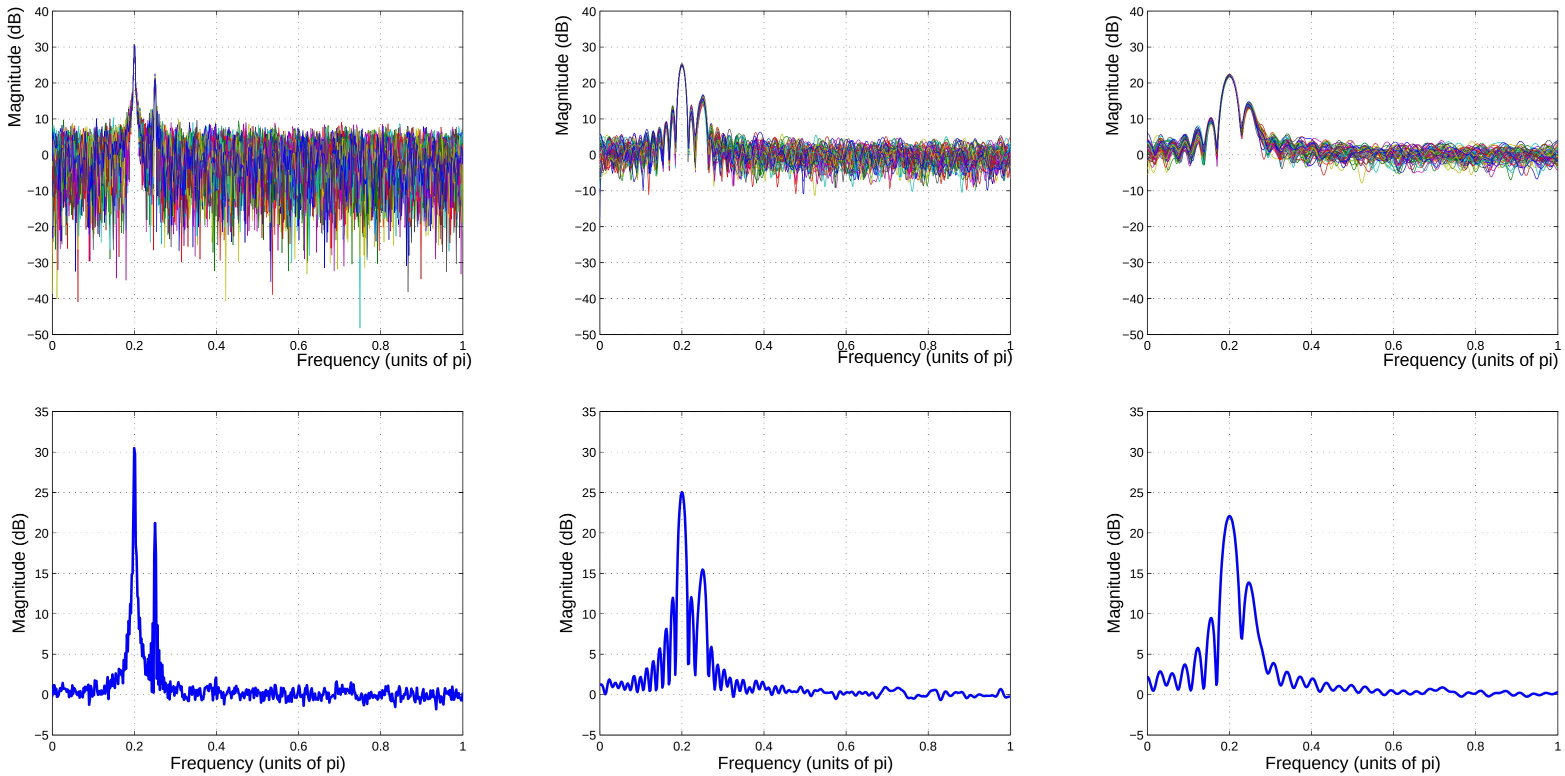
By changing the values of L and K, Bartlett's method allows us to:

trade a reduction in spectral resolution for a reduction in variance

Example: Estimation of two sinewaves in white noise

$x[n] = \sqrt{10} sin(n * 0.2\pi + \Phi_1) + sin(n * 0.25\pi + \Phi_2) + w[n]$

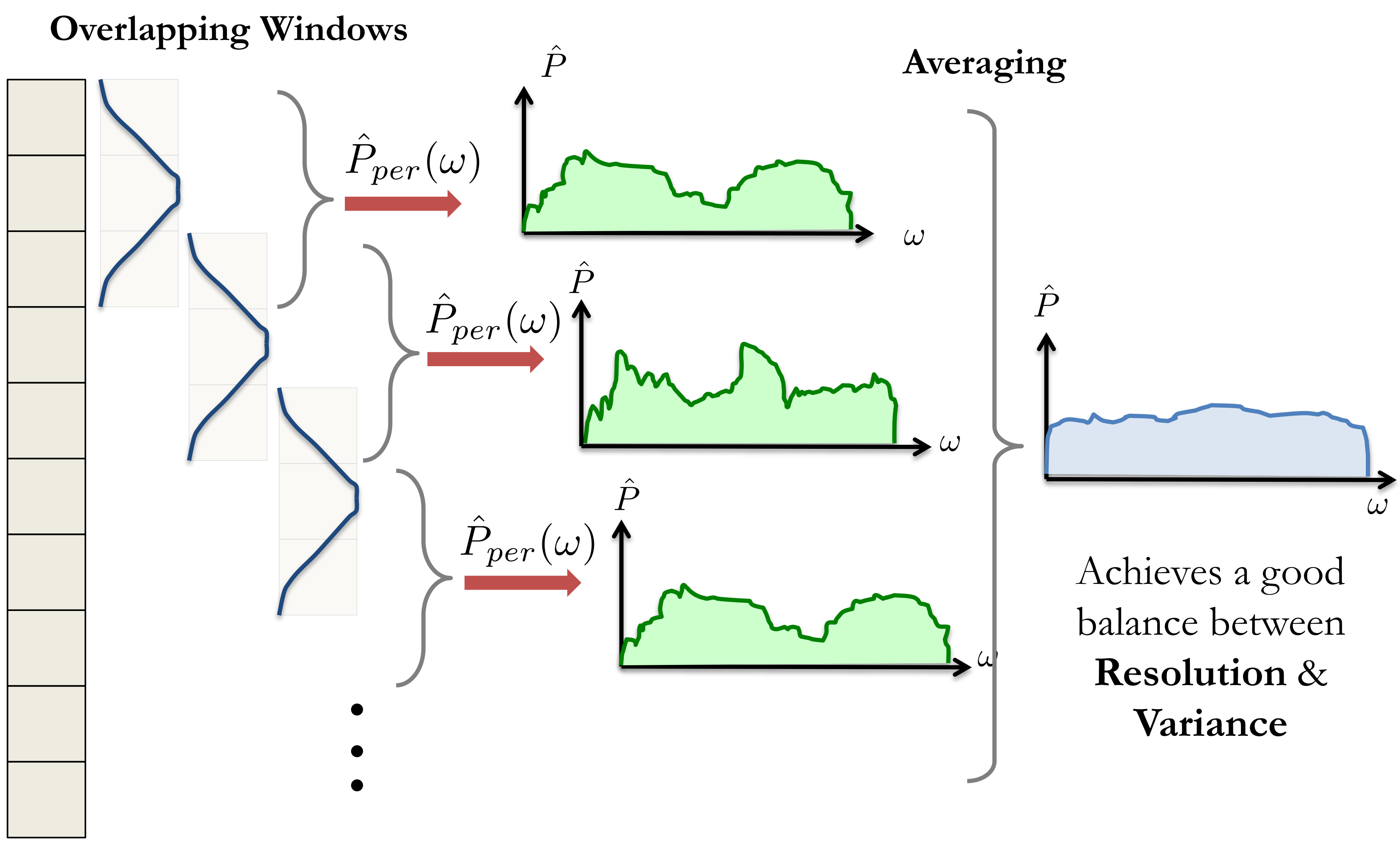
50 periodograms with $N = 512$ | **50 Bartlett estimates $K = 4, \; L = 128$** | **50 Bartlett estimates $K = 8, \; L = 64$**



Ensemble averages

Notice the variance – resolution trade–off!

Welch Method



Achieves a good balance between **Resolution** & **Variance**

Welch's method: Averaging modified periodograms

Welch proposed two modifications to Bartlett's method:

- allow the sequences $x_i[n]$ to overlap

- to allow data window $w[n]$ to be applied to each sequence $\Rightarrow$ averaging modified periodograms

This way, successive segments are offset by D points and each segment is L points long

$$x_i[n] = x[n + iD] \qquad n = 0, 1, \ldots, L - 1$$

The amount of overlap between $x_i[n]$ and $x_{i+1}[n]$ is $L - D$ points and

$$N = L + D(K - 1)$$

N- total number of points, L- length of segments, D- amount of overlap, K- number of sequences

Variations on the theme

We may vary between **no overlap D=L** and say 50 % overlap **D = L/2** or anything else.

☺ we can trade a reduction in the variance for a reduction in the resolution, since

$$\hat{P}_W(\omega_m) = \frac{1}{KLU} \sum_{i=0}^{K-1} \left| \sum_{k=0}^{L-1} w[k]x[k+iD]e^{-\jmath \omega_m k} \right|^2$$

or in terms of modified periodograms

$$\hat{P}_W(\omega_m) = \frac{1}{K} \sum_{i=0}^{K-1} \hat{P}_M^{(i)}(\omega_m)$$

$\rightsquigarrow$ **asymptotically unbiased** (follows from the bias of the modified periodogram)

Properties of Welch's method

- **Functional relationship:**

$$\hat{P}_W(\omega_m) = \frac{1}{KLU} \sum_{i=0}^{K-1} \left| \sum_{k=0}^{L-1} w[k]x[k+iD]e^{-\jmath \omega_m k} \right|^2 \qquad U = \frac{1}{L} \sum_{n=0}^{L-1} |w[n]|^2$$

- **Bias**

$$E\left\{ \hat{P}_W(\omega_m) \right\} = \frac{1}{2\pi LU} P_{xx}(\omega_m) \; * \; |W(\omega_m)|^2$$

- **Resolution** $\looparrowright$ window dependent

- **Variance** (assuming 50 % overlap and Bartlett window)

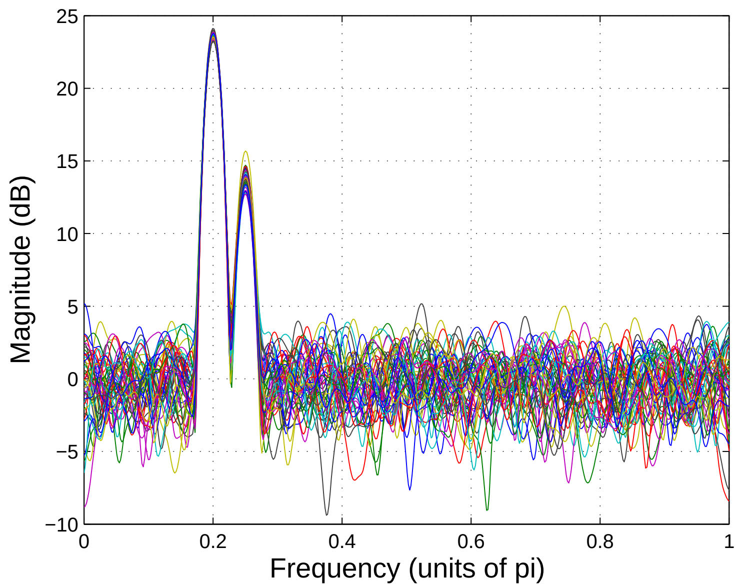
$$Var\left\{ \hat{P}_W(\omega_m) \right\} \approx \frac{9}{16} \frac{L}{N} P_{xx}^2(\omega_m)$$

Example: Two sinusoids in noise $\looparrowright$ Welch estimates

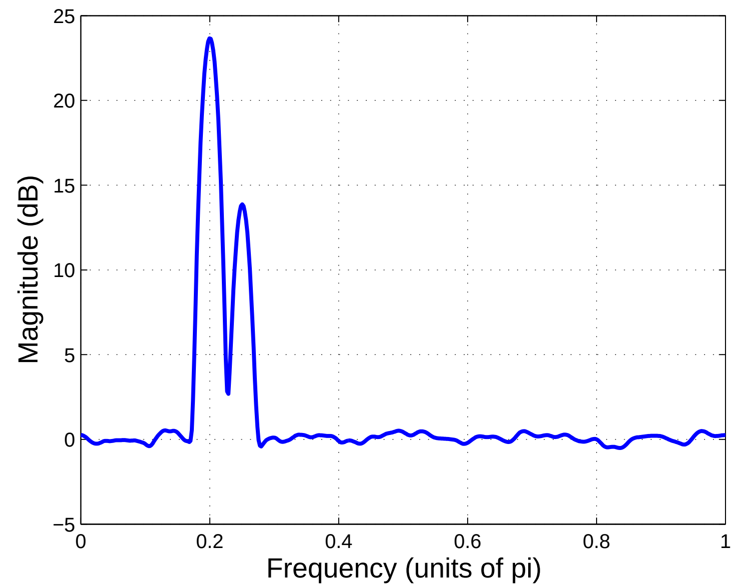
Problem: Estimate the spectra of the following two sinewaves using Welch's method

$$x[n] = \sqrt{10}\sin(n * 0.2\pi + \Phi_1) + \sin(n * 0.3\pi + \Phi_2) + w[n]$$

Unit noise variance, $N = 512$, $L = 128$, 50 % overlap (7 sections)



Overlay of 50 estimates



Periodogram using Welch's method

Blackman-Tukey Method

The Periodogram can also be expressed as:

$$\hat{P}_{per}(\omega_m) = \sum_{k=-N+1}^{N-1} \hat{\mathbf{r}}_{xx}[k] e^{-\jmath \omega_m k}$$

Autocorrelation Estimates at large lags are **unreliable**

$$\hat{P}_{BT}(\omega_m) = \sum_{k=-M}^{M} w[k] \hat{\mathbf{r}}_{xx}[k] e^{-\jmath \omega_m k}$$

Lags: $M < N - 1$

Windowing

Next: Can we **extrapolate the autocorrelation** estimates for lags $k > M$?

Blackman–Tukey method: Periodogram smoothing

Recall that the methods by Bartlett and Welch are designed to reduce the variance of the periodogram by averaging periodograms and modified periodograms, respectively.

Another possibility is "periodogram smoothing" often called the Blackman–Tukey method.

Let us identify the problem ☹

$$\hat{\mathbf{r}}_{xx}[N-1] = \frac{1}{N} x[N-1]x[0]$$

$\Rightarrow$ there is little averaging when calculating the estimates of $\hat{\mathbf{r}}_{xx}[k]$ for $|k| \approx N$.

These estimates will be **unreliable** no matter how large N. We have two choices:

- reduce the variance of those unreliable estimates
- reduce the contribution these unreliable estimates make to the periodogram

Blackman–Tukey Method: Resolution vs. Variance

The variance of the periodogram is decreased by reducing the variance of the ACF estimate by calculating more robust ACF estimates over fewer data points ($M < N$).

$\Rightarrow$ Apply a window to $\hat{\mathbf{r}}_{xx}[k]$ to decrease the contribution of unreliable estimates and obtain the Blackman–Tukey estimate:

$$\hat{P}_{BT}(\omega_m) = \sum_{k=-M}^{M} \hat{\mathbf{r}}_{xx}[k] w[k] e^{-\jmath \omega_m k}$$

where $w[k]$ is a **lag window** applied to the ACF estimate.

$$\hat{P}_{BT}(\omega_m) = \frac{1}{2\pi} \hat{P}_{per}(\omega_m) \; * \; W(\omega_m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{P}_{per}(\jmath u) W(\jmath(\omega_m - u)) du$$

that is, **we trade the reduction in the variance for a reduction in the resolution** (smaller number of ACF estimates used to calculate the PSD)

Properties of the Blackman–Tukey method

- **Functional relationship:**

$$\hat{P}_{BT}(\omega_m) = \sum_{k=-M}^{M} \hat{\mathbf{r}}_{xx}[k]w[k]e^{-\jmath k\omega}$$

- **Bias**

$$E\left\{\hat{P}_{BT}(\omega_m)\right\} \approx \frac{1}{2\pi}P_{xx}(\omega) \ * \ W(\omega)$$

- **Resolution**– window dependent (window – conjugate symmetric and with non–negative FT)

- **Variance:** Generally, it is recommended $M < N/5$.

$$Var\left\{\hat{P}_{BT}(\omega_m)\right\} \approx P_{xx}^2(\omega)\frac{1}{N}\sum_{k=-M}^{M} w^2[k]$$

Trade–off: for a small bias M needs to be large to minimize the width of the mainlobe of $W(\omega_m)$, whereas M should be small in order to minimize the variance.

Performance comparison of periodogram–based methods

Let us introduce criteria for performance comparison:

- **Variability of the estimate**

$$\nu = \frac{var\left\{\hat{P}_{xx}(\omega_m)\right\}}{E^2\left\{\hat{P}_{xx}(\omega_m)\right\}}$$

  which is effectively **normalised variance**

- **Figure of merit**

$$\mathcal{M} = \nu \times \Delta\omega$$

  that is, **product of variability and resolution.**

$\mathcal{M}$ should be as small as possible.

Performance measures for the Nonparametric methods of Spectrum Estimation

Method	Variability ν	Resolution $\Delta\omega$	Figure of merit $\mathcal{M}$
Periodogram	1	$0.89\frac{2\pi}{N}$	$0.89\frac{2\pi}{N}$
Bartlett	$\frac{1}{K}$	$0.89K\frac{2\pi}{N}$	$0.89\frac{2\pi}{N}$
Welch	$\frac{9}{8}\frac{1}{K}$	$1.28\frac{2\pi}{L}$	$0.72\frac{2\pi}{N}$
Blackman–Tukey	$\frac{2}{3}\frac{M}{N}$	$0.64\frac{2\pi}{M}$	$0.43\frac{2\pi}{N}$

- Observe that each method has a Figure of Merit which is approximately the same

- Figure of merit are inversely proportional to N

- Although each method differs in its resolution and variance, **the overall performance is fundamentally limited by the amount of data that is available.**

Conclusions

FFT based spectral estimation is limited by:

- correlation assumed to be zero beyond N - biased/unbiased estimates
- resolution limited by the DFT "baggage"
- if two frequencies are separated by $\Delta\omega$, then we need $N \geq \frac{2\pi}{\Delta\omega}$ data points to separate them
- limitations for spectra with narrow peaks (resonances, speech, sonar)
- limit on the resolution imposed by N also causes bias
- variance of the periodogram is almost independent of data length
- the derived variance formulae are only illustrative for real–world signals

But also many opportunities: spectral coherency, spectral entropy, TF, ...

Next time: model based spectral estimation for discrete spectral lines

Appendix: FT basics

Periodic signal	$\leftrightsquigarrow$	**Discrete** FT
Discrete signal	$\leftrightsquigarrow$	**Periodic** FT
Periodic and **Discrete** signal	$\leftrightsquigarrow$	**Discrete** and **Periodic** FT
Discrete and **Periodic** signal	$\leftrightsquigarrow$	**Periodic** and **Discrete** FT

- **Sampling** yields a new signal ($\omega_s = \frac{2\pi}{T}$) (poor approximation)

$$g[n] = f[nT] \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad G(\omega) = \sum_{k=-\infty}^{\infty} F(\omega + k\Omega_0)$$

- **Limiting** the length to N samples effectively introduces rectangular windowing (Leakage)

$$W(\omega) = \frac{\sin(N\omega T/2)}{\sin(\omega T/2)} e^{-\jmath \frac{N-1}{2}\omega T}$$

$\Rrightarrow$ **Estimated Spectrum $=$ True spectrum $*$ Dirichlet kernel**

Appendix: Spectral Coherence and LS Periodogram see also Problem 4.7 in your P/A sets

The **spectral coherence** shows similarity between two spectra

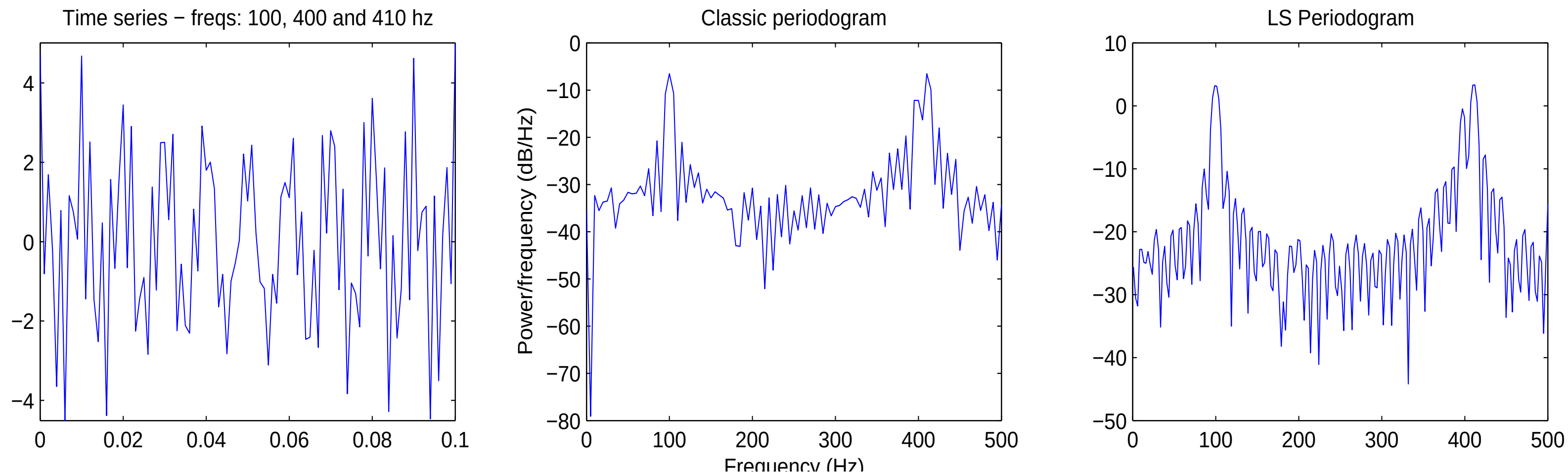
$$C_{xy}(\omega) = \frac{P_{xy}(\omega)}{\left[P_{xx}(\omega)P_{yy}(\omega)\right]^{1/2}}$$

It is invariant to linear filtering of x and y (even with different filters)

The periodogram $P_{per}(\omega_m)$ can be seen as a **Least Squares** solution to

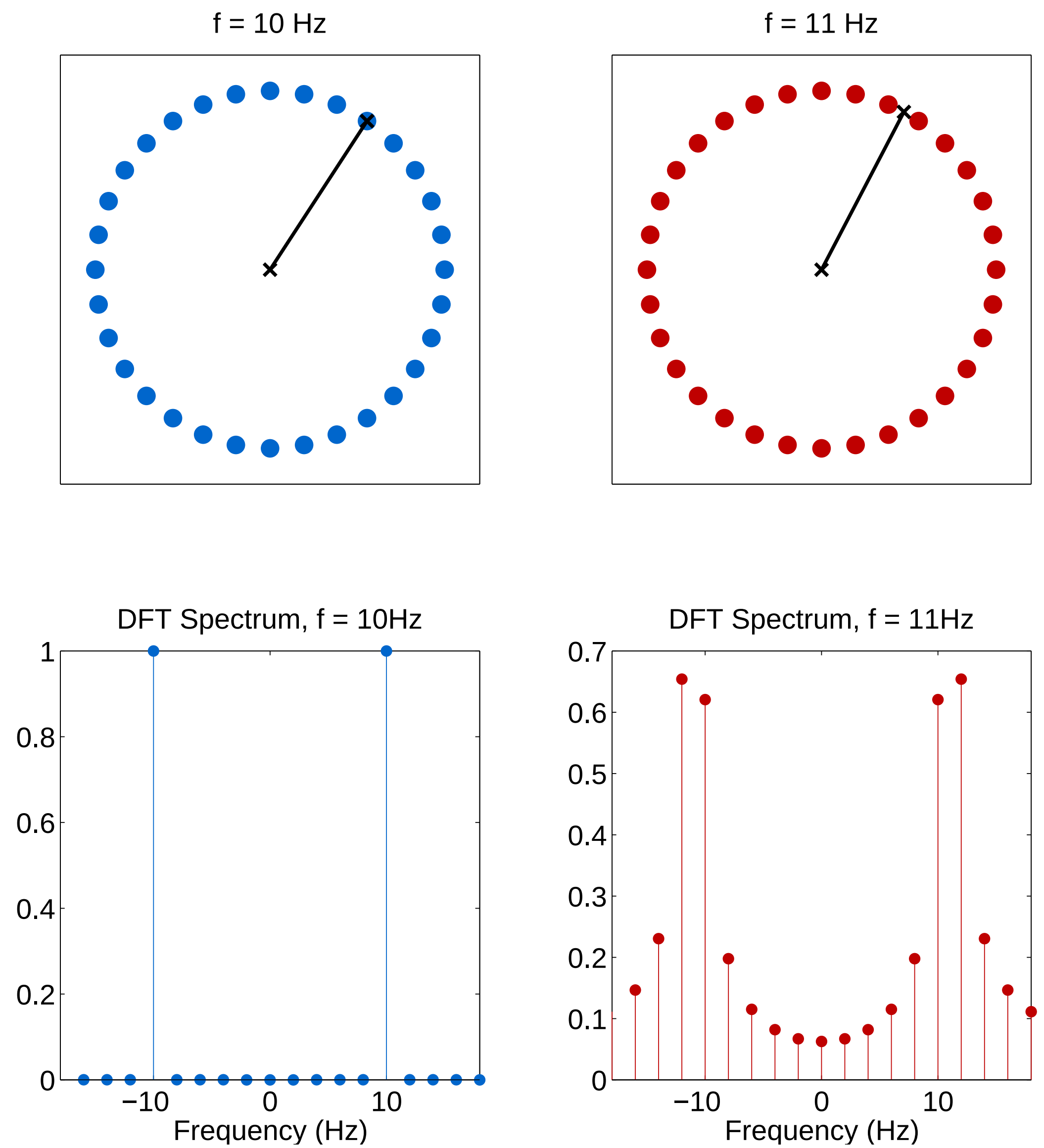
$$P_{per}(\omega_m) = \|\hat{\beta}(\omega_m)\|^2, \;\; \hat{\beta} = \underset{\beta(\omega_m)}{\operatorname{argmin}} \sum_{n=1}^{N} \|y[n] - \beta e^{j\omega_m n}\|^2,$$

Periodogram and LS periodog. for a sinewave mixture (100, 400, 410) Hz



Practical Issue #4: Incoherent sampling

Visual representation $\looparrowright$ dots denote angles of $\rightsquigarrow \frac{2\pi}{N} \cdot m$



Rank of the covariance matrix for sinusoidal data

The difference between $\mathbb{R}^2$ and $\mathbb{C}$

Consider a single complex sinusoid with no noise

$$z_k = Ae^{\jmath\omega k} = A\cos(\omega k + \phi) + \jmath A\sin(\omega k + \phi)$$

There are two possible representations of the signal: A univariate complex-valued vector *or* bivariate real-valued matrix:

1. $\mathbf{z} = [z_0, z_1, \ldots, z_{N-1}]^T = A[1, e^{j\omega}, \ldots, e^{j(N-1)\omega}]^T \overset{\text{def}}{=} A\mathbf{f}$

2. $\mathbf{Z} = \begin{bmatrix} \text{Re}\{\mathbf{z}\} \\ \text{Im}\{\mathbf{z}\} \end{bmatrix} = A \begin{bmatrix} 1 & \cos(\omega + \phi) & \ldots & \cos(\omega(N-1) + \phi) \\ 0 & \sin(\omega + \phi) & \ldots & \sin(\omega(N-1) + \phi) \end{bmatrix}^T$

The corresponding covariance matrices exhibit a very interesting property:

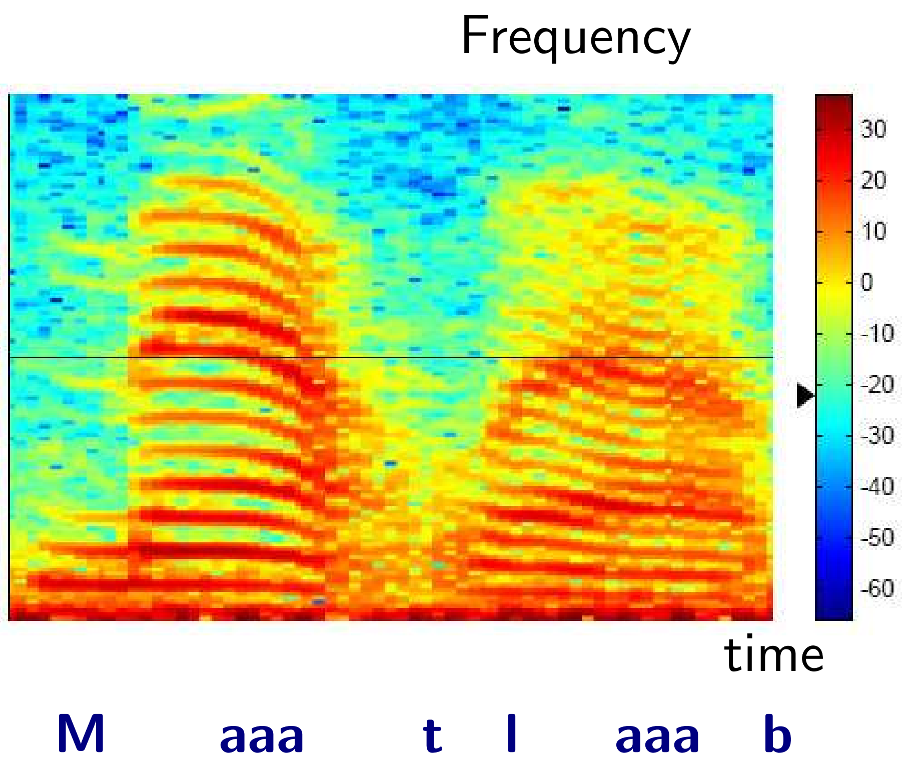
- $\mathbf{C}_{zz} = E\{\mathbf{z}\mathbf{z}^H\} = |A|^2\mathbf{f}\mathbf{f}^H \rightarrow$ Rank $= 1$.

- $\mathbf{C}_{ZZ} = E\{\mathbf{Z}\mathbf{Z}^T\} \rightarrow$ Rank $= 2$.

What would happen with p sinusoids?

Appendix: Time-Frequency estimation

time–frequency spectrogram of "Matlab" ↬ 'specgramdemo'



For every time instant "t", the PSD is plotted along the vertical axis

Darker areas: higher magnitude of PSD

Appendix: Time-Frequency (TF) analysis - Principles

Assume $x[n]$ has a Fourier transform $X(\omega)$ and power spectrum $|X(\omega)|^2$.

The function $TF(n,\omega)$ determines how the energy is distributed in time-frequency, and it satisfies the following **marginal properties:**

$$\sum_{n=-\infty}^{\infty} TF(n,\omega) = |X(\omega)|^2 \quad \text{energy in the signal at frequency } \omega$$

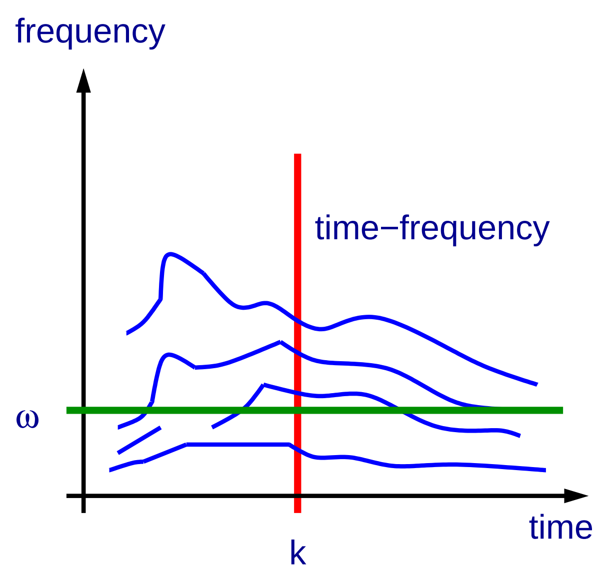
$$\frac{1}{2\pi}\int_{-\pi}^{\pi} TF(n,\omega)d\omega = |x[n]|^2 \quad \text{energy at time instant k due to all } \omega$$

Then

$$\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\int_{\infty}^{\infty} TF(n,\omega)d\omega = \sum_{n=-\infty}^{\infty}|x[n]|^2$$

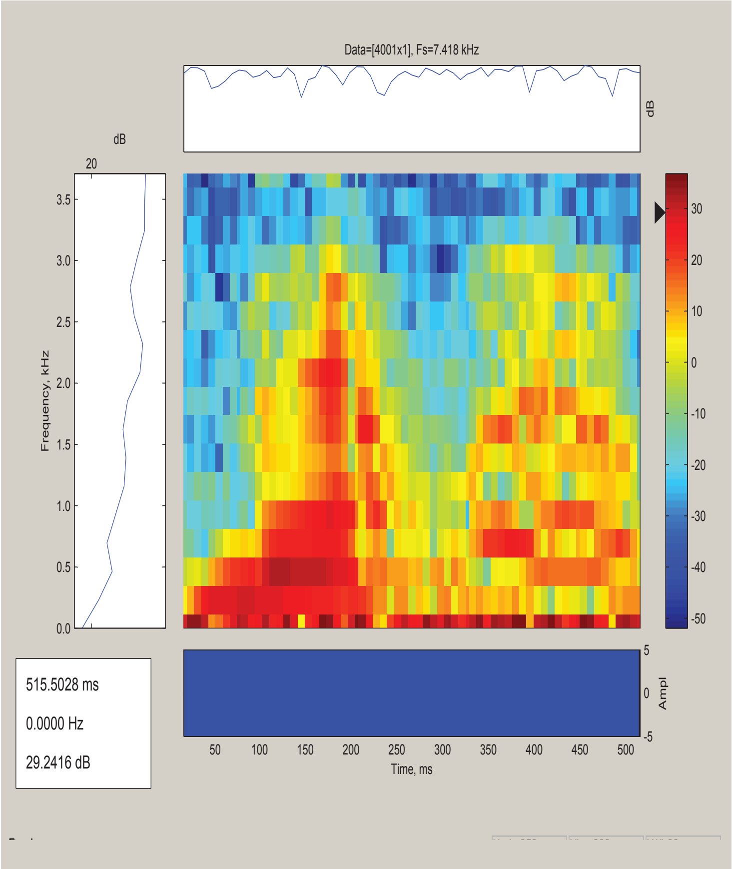
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}|X(\omega)|^2 d\omega$$

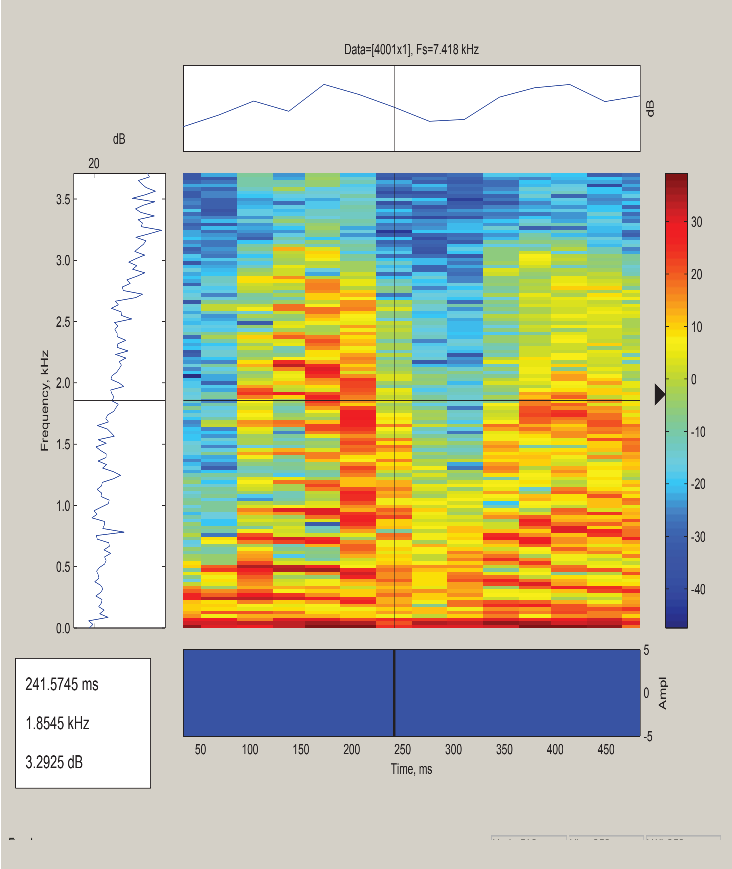
giving the **total energy** (all frequencies and samples) of a signal.



Time–frequency spectrogram of a speech signal

(wide band spectrogram) **(narrow band spectrogram)**





(win-len=256, overlap=200, ftt-len=32) **(win-len=512, overlap=200, ftt-len=256)**

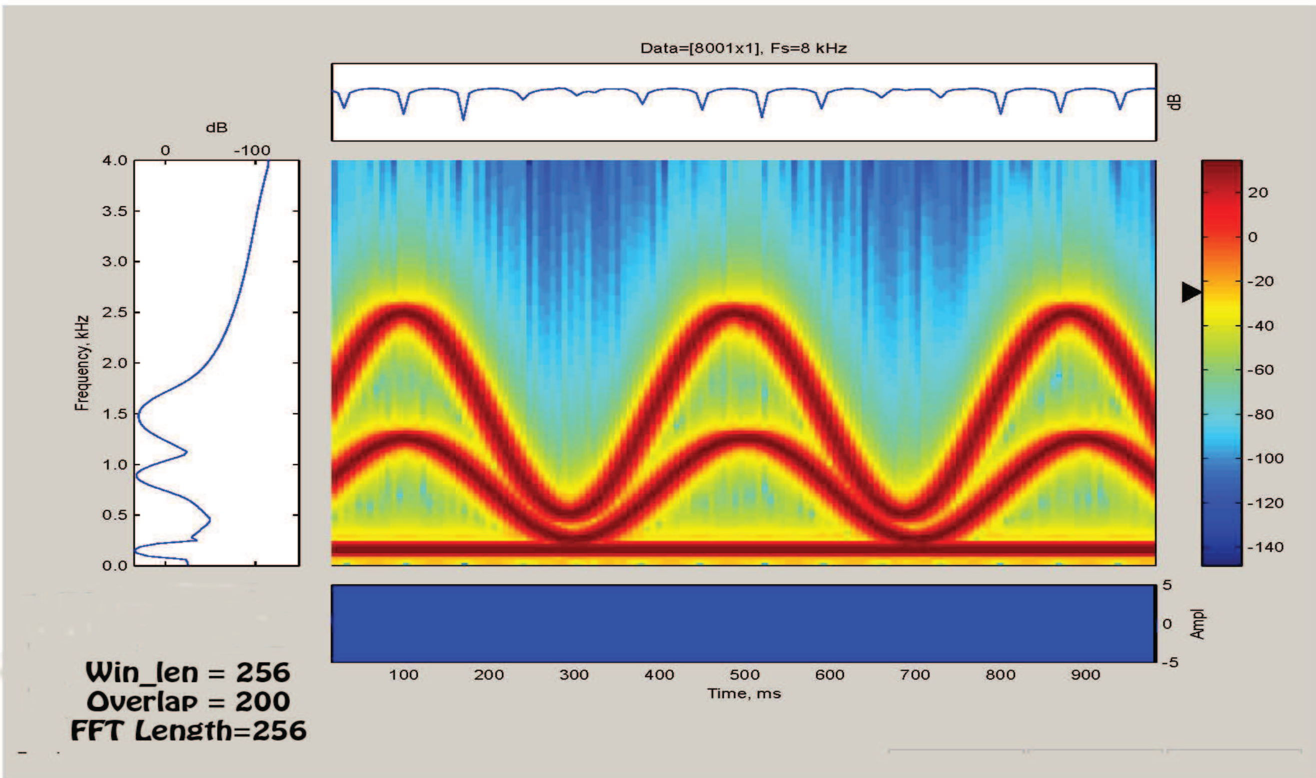
Homework: evaluate all the methods from the lecture for this T-F spectrogram

TF spectrogram of a frequency-modulated signal
(check also your coursework)

The time-frequency spectrogram of a frequency modulated (FM) signal

$$y(t) = A\cos\left[\omega_0 t + k_f \int_{-\infty}^{t} x(\alpha)d\alpha\right]$$

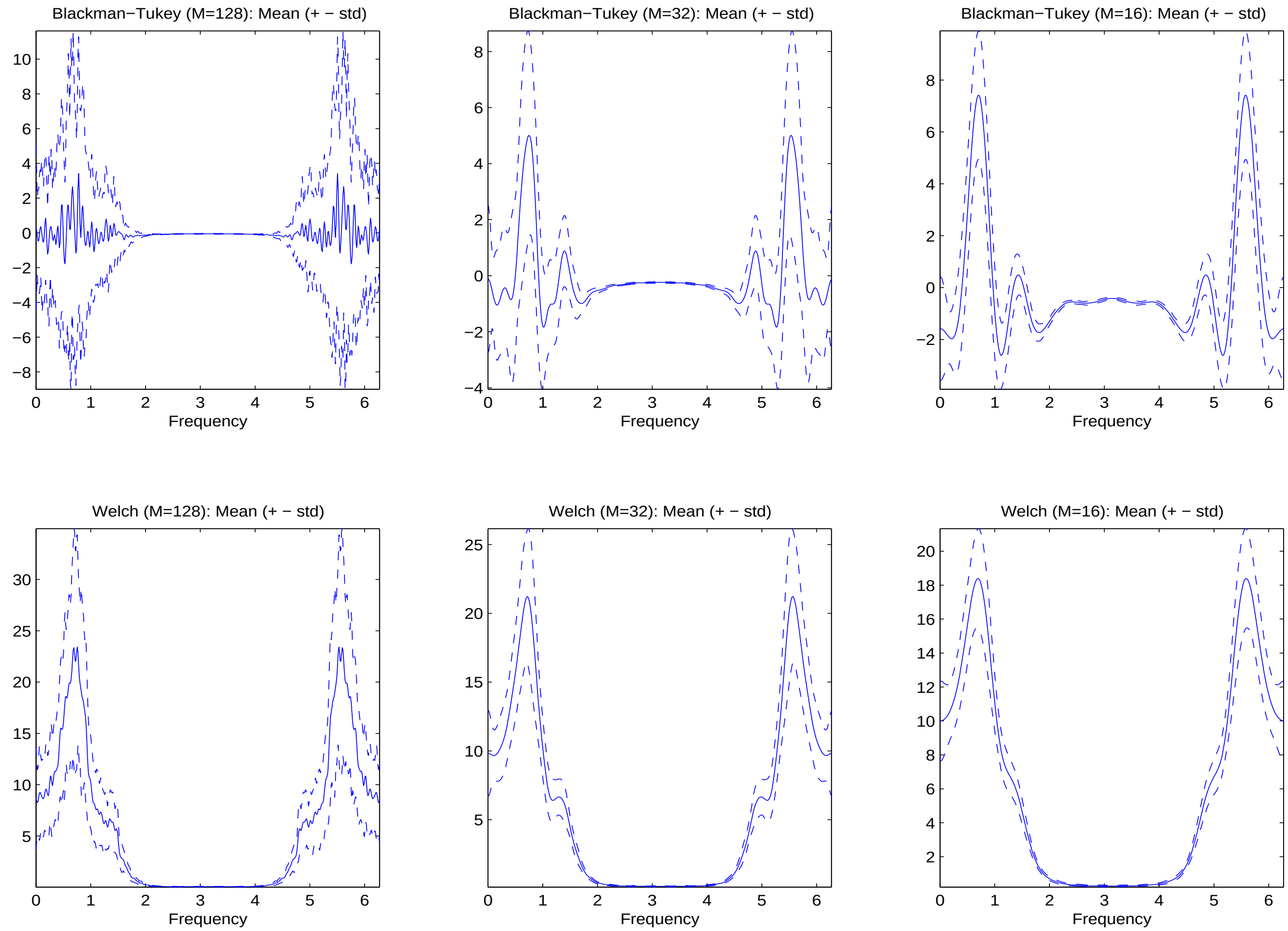
frequency



time

Opportunities: ARMA spectrum

N=512 samples, freq. res=1/500



Signal: ARMA(4,4), b=[1, 0.3544, 0.3508, 0.1736, 0.2401] a=[1, -1.3817, 1.5632, -0.8843, 0.4096]

Sometimes we only desire the correct position of the peaks $\looparrowright$ **ARMA Spectrum Estimation**

A note on positive-semidefiniteness of the $\mathbf{R}_{xx}$

The autocorrelation matrix $\mathbf{R}_{xx} = E\big[\mathbf{x}\mathbf{x}^T\big]$
where $\mathbf{x} = \big[x[0], \ldots, x[N-1]\big]^T$. It is symmetric and of size $N \times N$.

There are four ways to define positive semidefiniteness: (see also your Problem-Answer sets)

1. All the eigenvalues of the autocorrelation matrix $\mathbf{R}$ are such that $\lambda_i \geq 0$, for i=1,. . . ,N
2. For any nonzero vector $\mathbf{a} \in \mathbb{R}^{N\times 1}$ we have $\mathbf{a}^T\mathbf{R}\mathbf{a} \geq 0$. For complex valued matrices, the condition becomes $\mathbf{a}^H\mathbf{R}\mathbf{a}$
3. There exists a matrix $\mathbf{U}$ such that $\mathbf{R} = \mathbf{U}\mathbf{U}^T$, where the matrix $\mathbf{U}$ is called a root of $\mathbf{R}$
4. All the principal submatrices of $\mathbf{R}$ are positive semidefinite. A principal submatrix is formed by removing $i = 1, \ldots, N$ rows and columns of $\mathbf{R}$

For positive definiteness conditions, replace $\geq$ with $>$

Two ways to estimate the ACF

For an **autocorrelation ergodic** process with an unlimited amount of data, the ACF may be determined:

1) Using the time–average

$$\mathbf{r}_{xx}[k] = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x[n+k]x[k]$$

If $x[n]$ is measured over a finite time interval, $n = 0, 1, \ldots, N-1$ then we need to *estimate* the ACF from a finite sum

$$\hat{\mathbf{r}}_{xx}[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n+k]x[n]$$

2) In order to ensure that the values of $x[n]$ that fall outside interval $[0, N-1]$ are excluded from the sum, we have (**biased estimator**)

$$\hat{\mathbf{r}}_{xx}[k] = \frac{1}{N} \sum_{n=0}^{N-1-k} x[n+k]x[n], \quad k = 0, 1, \ldots, N-1$$

Cases 1) and 2) are equivalent for small lags and a fast decaying ACF

Case 1) gives positive semidefinite ACF, this is not guaranteed for Case 2)

Opportunities: Spectral Entropy

Spectral entropy can be used to measure the peakiness of the spectrum.

This is achieved via the probability mass function (PMF) (normalised PSD) given by

$$\eta(\omega_m) = \frac{P_{per}(\omega_m)}{\sum_{l=0}^{N-1} P_{per}(\omega_l)} \quad \rightarrow \quad H_{sp} = -\sum_{m=0}^{N-1} \eta(\omega_m) \log_2 \eta(\omega_m)$$

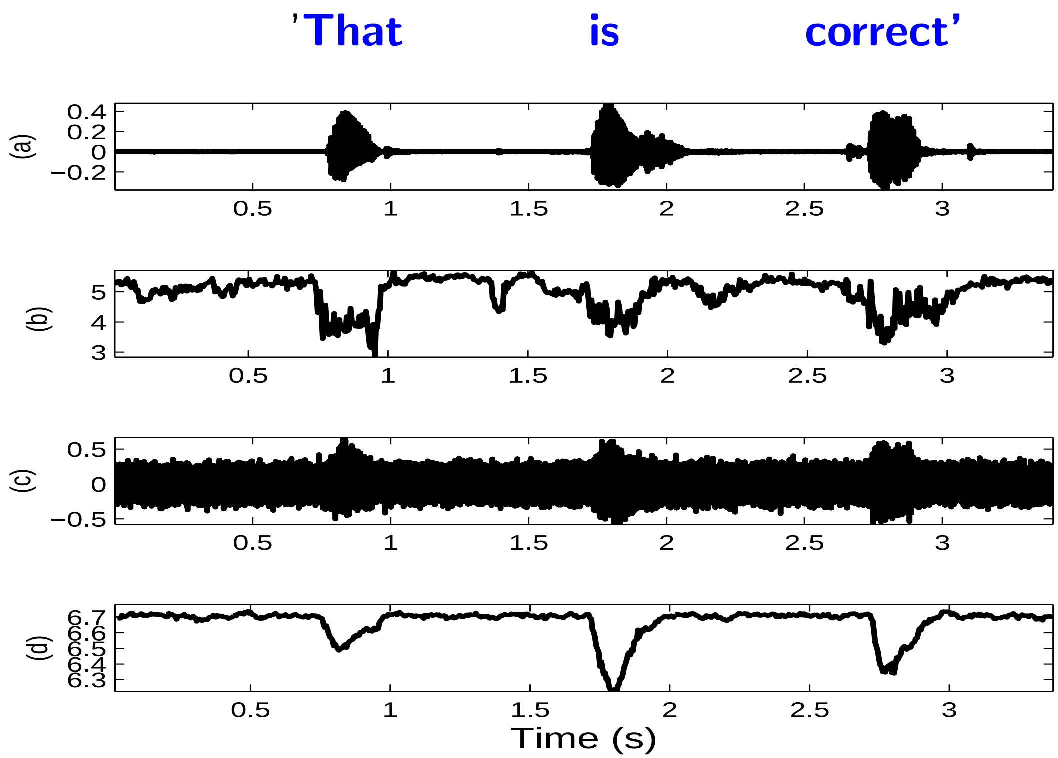
Intuition:

- peaky spectrum (e.g. $\sin(x)$) $\looparrowright$ low spectral entropy

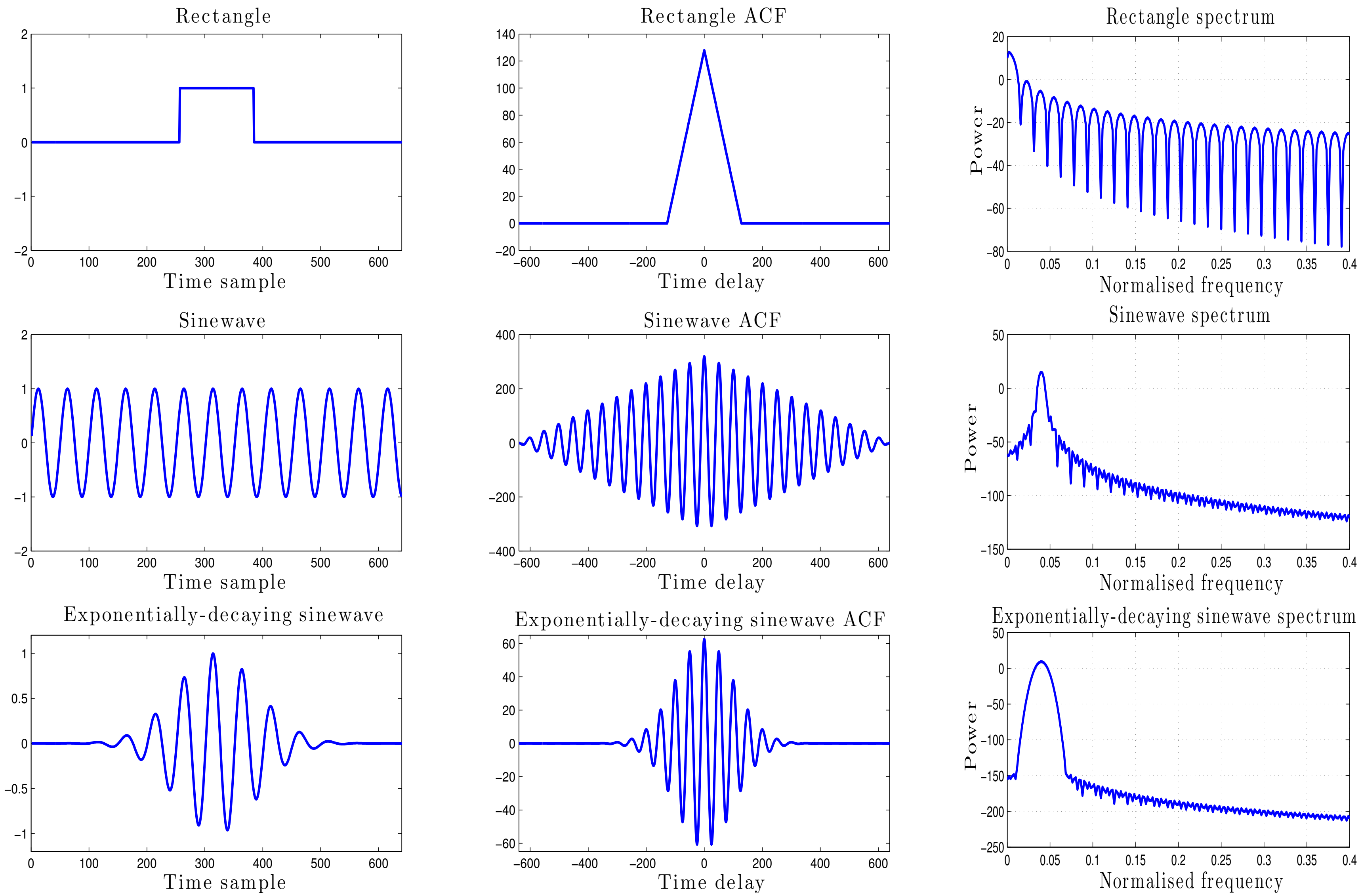
- flat spectrum (e.g. WGN) $\looparrowright$ high spectral entropy

Figure on the right:

From top to bottom: a) clean speech, b) spectral entropy, c) speech + noise, d)spectral entropy of (speech+noise)



Appendix: Practical issues in correlation and spectrum estimation



Properties of an ideal window function

Consider a window sequence $w[n]$ whose DFT is a **squared magnitude of another sequence** $v[n]$, that is

$$V(\omega) = \sum_{k=0}^{M-1} v[k]e^{-\jmath\omega k} \quad \looparrowright \quad W(\omega) = |V(\omega)|^2 \quad (\textbf{positive definite})$$

Then

$$\sum_{k=-(M-1)}^{M-1} w[k]e^{-\jmath\omega k} = \sum_{n=0}^{M-1}\sum_{p=0}^{M-1} v[n]v[p]e^{-\jmath(n-p)}$$

$$= \sum_{k=-(M-1)}^{M-1}\Big[\sum_{n=0}^{M-1} v[n]v[n-k]\Big]e^{-\jmath k}, \quad \text{for } v[k]=0, \ k \notin [0, M-1]$$

This gives

$$w[k] = \sum_{n=0}^{M-1} v[n]v[n-k] = v[k]*v[k] \quad \Leftrightarrow \quad W(\omega) \geq 0 \quad \textbf{pos. semidefinit.}$$

A window design should trade-off between smearing and leakage

For instance: weak sinewave + strong narrowband interference $\rightarrow$ leakage more detrimental than smearing

Homework: can we use optimisation to balance between smearing and leakage

Appendix: Trade-off in window design

window length $\looparrowright$ trade-off between spectral resolution and statistical variance

- most windows take non-negative values in both time and frequency
  - They also peak at origin in both domains

For this type of window we can define:

- An **equivalent time width** N_x $\quad$ ($N_x \approx 2M$ for rectangular and $N_x \approx M$ for triangular window)
- An **equivalent bandwidth** B_x ($\approx$ determined by window's length), as

$$N_w = \frac{\sum_{k=-(M-1)}^{M-1} w[k]}{w[0]} \qquad B_w = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega)d\omega}{W(0)}$$

We also know that

$$W(0) = \sum_{k=-\infty}^{\infty} w[k] = \sum_{k=-(M-1)}^{M-1} w[k] \qquad \text{and} \qquad w[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega)d\omega$$

It then follows that $\qquad N_w \times B_w = 1$

A window cannot be both time-limited and band-limited, usually $M \leq N/10$

Appendix: More on time–bandwidth products

The previous slide assumes that both $w[n]$ and $W(\omega)$ peak at the origin $\hookrightarrow$ most energy concentrated in the main lobe, whose width is $\sim 1/M$.

For a general signal: $x[n]$ and $X(\omega_m)$ can be negative or complex

If x[n] peaks at n_0 $(cf.\ X(\omega)\ at\ \omega_0) \hookrightarrow\ N_x = \dfrac{\sum_{n=-\infty}^{\infty}|x[k]|}{|x[n_0]|},\ \ B_x = \dfrac{\frac{1}{2\pi}\int_{-\pi}^{\pi}|X(\omega_0)|d\omega}{|X(\omega_0)|}$

Because $x[n]$ and $X(\omega)$ are Fourier transform pairs:

$$
\begin{aligned}
|X(\omega_0)| &= \left|\sum_{n=-\infty}^{\infty} x[k]e^{-\jmath\omega_0 n}\right| \leq \sum_{n=-\infty}^{\infty}|x[k]| \\
|x[n_0]| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{\jmath\omega n_0}d\omega\right| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}|X(\omega)|d\omega
\end{aligned}
$$

This implies

$N_x \times B_x \geq 1$ (a sequence cannot be narrow in both time and frequency)

More precisely: if the sequence is narrow in one domain then it must be wide in the other domain (uncertainty principle)

Some intuition: Fourier transform as a digital filter

We can see FT as a convolution of a complex exponential and the data (under a mild assumption of a one-sided h sequence, ranging from 0 to ∞)

1) Continuous FT. For a continuous FT $\quad F(\omega) = \int_{-\infty}^{\infty} x(t) e^{-\jmath\omega t} dt$

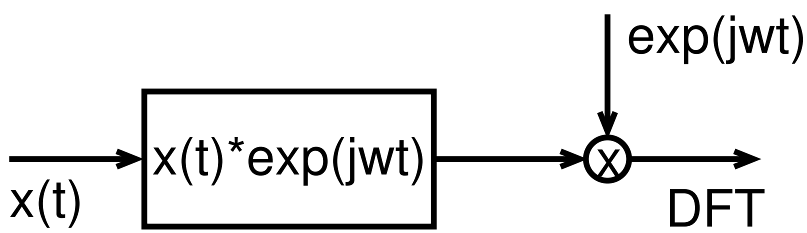
Let us now swap variables $t \rightarrow \tau$ and multiply by $e^{\jmath\omega t}$, to give

$$e^{\jmath\omega t} \int x(\tau) e^{-\jmath\omega\tau} d\tau = \int x(\tau) \underbrace{e^{\jmath\omega(t-\tau)}}_{h(t-\tau)} d\tau = x(t) * e^{\jmath\omega t} \qquad (= x(t) * h(t))$$

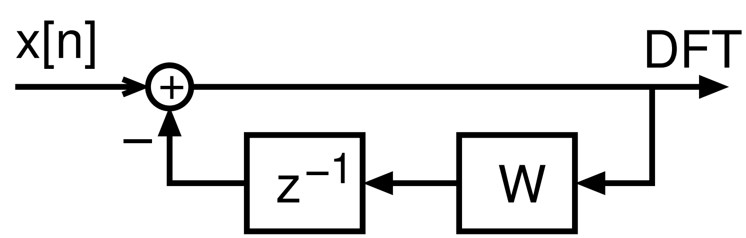
2) Discrete Fourier transform. For DFT, we have a filtering operation

$$X[k] = \sum_{n=0}^{N-1} x(n) e^{-\jmath\frac{2\pi}{N}nk} = \underbrace{x(0) + W\Big[x(1) + W\big[x(2) + \cdots\Big]}_{\text{cumulative add and multiply}} \qquad W = e^{-\jmath\frac{2\pi}{N}k}$$

with the transfer function (large N) $\quad H(z) = \frac{1}{1 - z^{-1}W} = \frac{1 - z^{-1}W^*}{1 - 2\cos\theta_k z^{-1} + z^{-2}}$



continuous time case

discrete time case


Appendix: Historical perspective

1772 **Lagrange** proposes use of rational functions to identify multiple periodic components;

1840 **Buys–Ballot**, tabular method;

1860 **Thomson**, harmonic analyser;

1897 **Schuster**, periodogram, periods not necessarily known;

1914 **Einstein**, smoothed periodogram;

1920-1940 Probabilistic theory of time series, Concept of spectrum;

1946 **Daniell**, smoothed periodogram;

1949 **Hamming & Tukey** transformed ACF;

1959 **Blackman & Tukey**, B–T method;

1965 **Cooley & Tukey**, FFT;

1976 **Lomb**, periodogram of unevenly spaced data;

1970– Modern spectrum estimation!

Why do not you think a little about ...

- ⊛ The resolution for zero-padded spectra is higher, what can we tell about the variance of such a periodogram?
- ⊛ If the samples at the start and end of a finite-length data sequence have significantly different amplitudes, how does this affect the spectrum?
- ⊛ What uncertainties are associated with the concept of "frequency bin"?
- ⊛ What happens with high frequencies in tapered periodograms?
- ⊛ What would be the ideal properties of a "data window"?
- ⊛ How frequently do we experience incoherent sampling in real life applications and what is a most pragmatic way to deal with the frequency resolution when calculating spectra of such signals?
- ⊛ How can we use the time–bandwidth product to ensure physical meaning of spectral estimates?
- ⊛ The "double summation" formula that uses progressively fewer samples to estimate the ACF is very elegant, but does it come with some problems too, especially for larger lags? (See Appendix)

Bias vs variance – recap

- **Bias** pertains to the question: "**Does the estimate approach the correct value as $N \to \infty$**".
  - ⊛ If yes then the estimator is unbiased, else it is biased
  - ⊛ Notice that the main lobe of the window has a width of $2\pi/N$ and hence when $N \to \infty$ we have $\lim_{N\to\infty} \hat{P}_{per}(\omega_m) = P_{xx}(\omega_m) \Rightarrow$ periodogram is an **asymptotically unbiased** estimator of true PSD.
  - ⊛ **For the window to yield an unbiased estimator:**

$$\sum_{n=0}^{N-1} w^2[n] = N \quad \& \quad \text{the mainlobe width} \sim \quad \frac{1}{N}$$

- **Variance** refers to the "goodness" of the estimate, that is, whether the power of the estimation error tend to zero when $N \to \infty$.
  - ⊛ We have shown that even for a very large window the variance of the estimate is as large as the true PSD
  - ⊛ This means that the periodogram **is not a consistent** estimator of true PSD.

Welch vs. Bartlett

- the amount of overlap between $x_i[n]$ and $x_{i+1}[n]$ is $L - D$ points, and if K sequences cover the entire N data points, then

$$N = L + D(K + 1)$$

- If there is no overlap, $(D = L)$ we have $K = \frac{N}{L}$ sections of length L as in Bartlett's method
- Of the sequences are overlapping by 50 % $D = \frac{L}{2}$ then we may form $K = 2\frac{N}{L} - 1$ sections of length L. thus maintaining the same resolution as Bartlett's method while doubling the number of modified periodograms that are averaged, thereby reducing the variance.
- With 50% overlap we could also form $K = \frac{N}{L} - 1$ sequences of length $2L$, thus increasing the resolution while maintaining the same variance as Bartlett's method.

Therefore, by allowing sequences to overlap, it is possible to increase the number and/or length of the sequences that are averaged, thereby trading a reduction in variance for a reduction in resolution.

Notes

Notes

Notes

Notes

Notes