Spectrum Estimation

Lim & Oppenheim: Ch. 2

Hayes: Ch. 8

Proakis & Manolakis Ch. 14

Introduction & Issues

Recall Definition of PSD

LO-2.1

Given a WSS random process $x[k]$ the PSD is defined by:

$$S_x(\omega) = \lim_{M\to\infty} E\left\{ \frac{1}{2M+1} \left| \sum_{n=-M}^{M} x[n]e^{-j\omega n} \right|^2 \right\} \qquad (\bigstar)$$

<< Warning: Ch. 2 of L&O uses "ω" for the DT frequency whereas Porat uses "Ω". Also, Hayes expresses DTFTs (and therefore PSDs) in terms of $e^{j\omega}$; it means the same thing – it is just a matter of notation. Hayes' notation is more precise when you consider going from the ZT $H(z)$ to the frequency response $H(e^{j\omega}) = H(z)|_{z=e^{j\omega}}$ >>

Recall the Wiener-Khinchine Theorem:

$$\begin{aligned} S_x(\omega) &= \mathcal{F}\{r_x[k]\} \\ &= \sum_{k=-\infty}^{\infty} r_x[k]e^{-j\omega k} \end{aligned} \qquad (\bigstar\bigstar)$$

$$\boxed{\underbrace{r_x[k] = E\{x[n]x^*[n+k]\}}_{\text{ACF of RP } x[n]}}$$

<u>Problem of PSD Estimation</u>

LO-2.2

1. Both (★) & (★★) involve ensemble averaging **<u>BUT</u>** in practice we <u>get only one realization</u> from the ensemble
2. Both (★) & (★★) use a Fourier transform of <u>infinite length</u> **<u>BUT</u>** in practice we get only a <u>finite number of samples</u>.

   (Note: a finite # of samples allows only a finite # of ACF values)

(★) & (★★) motivate two approaches to PSD estimation:

1. Compute the <u>DFT</u> of the signal and then do <u>some form of averaging</u>
2. Compute and estimate of the ACF using <u>some form of averaging</u> and then compute the <u>DFT</u>

Both of these approaches are called "Classical" Nonparametric Approaches – they strive to do the best with the available data w/o making any assumptions other than that the underlying process is WSS.

The "Modern" Parametric Approach

LO-2.2

There is a so-called "Modern" approach to PSD estimation that tries to deal with the issue of having only a finite # of samples:

- ➔ Assume a recursive model for the ACF
- ➔ Allows recursive extension of ACF using the known values

Example Model

$$r_x[k] = -a_1 r_x[k-1] - a_2 r_x[k-2] - \cdots - a_p r_x[k-p], \quad k \geq p+1$$

We'll see that for this approach all we'll need to do is estimate the model parameters $\{a_i\}$ and then use them to get an estimate of the PSD …. Thus, this approach is called "Parametric"

Review of Statistics

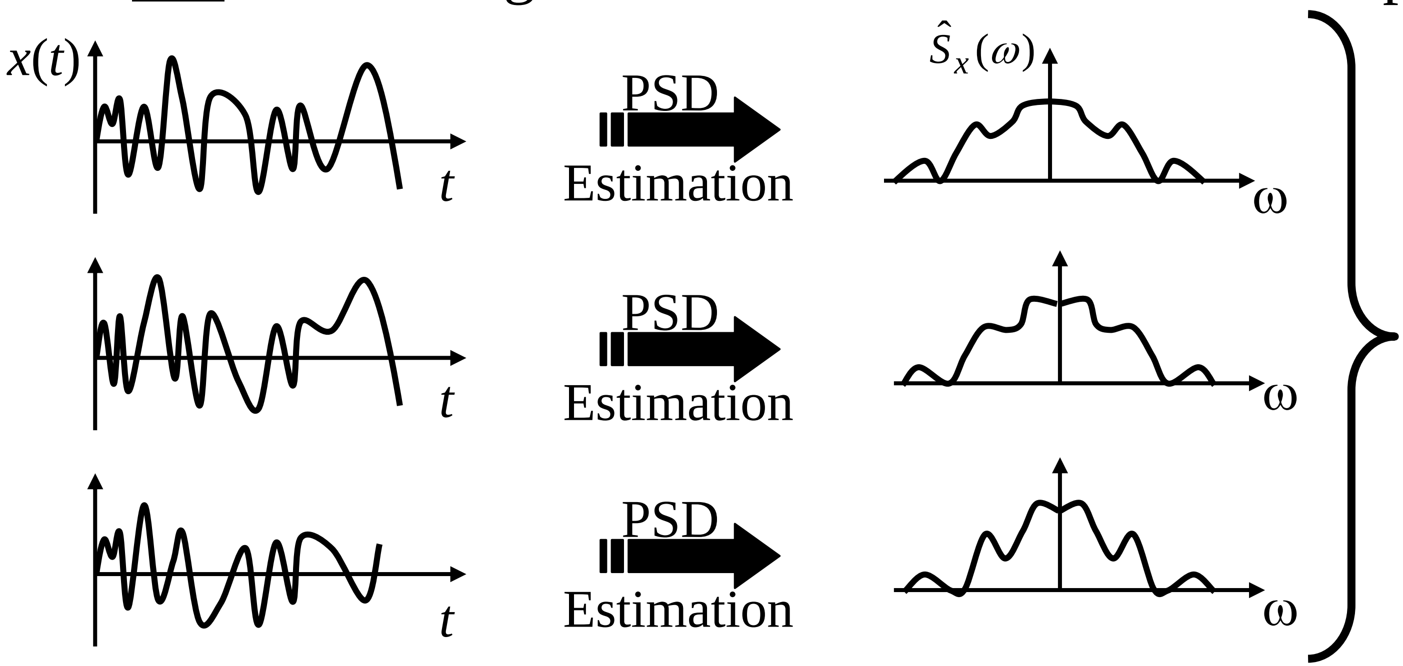
LO-2.3

Before we can really address the issue of estimating a PSD we need to review a few issues from statistics.

What are we doing in PSD Estimation?

Given: Finite # of samples from one realization

Get: Something that "resembles" the PSD of the process



Each Signal Realization gives a different PSD Estimate Realization

Each PSD Estimate is a Realization of a Random Process

Review of Statistics (Cont.)

Thus… must view PSD Estimate as a Random Process

Need to characterize its mean and variance:

- Want Mean of PSD Estimate = true PSD
- Want Var of PSD Estimate= "small"

To make things easier to discuss, we use a slightly different estimation problem to illustrate the ideas… Consider the process

$$x[n] = A + w[n]$$

Constant AWGN, zero-mean, σ^2

Given a finite set of data samples $x[0]$,… $x[N$-1]…. estimate A.

Reasonable estimate is:

$$\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$$

"sample mean"

For each realization of $x[n]$ you get a different value for the estimate of A.

Review of Statistics (Cont.)

We want two things for the estimate:

1. We want our estimate to be "correct on average": $\boxed{E\{\hat{A}\} = A}$

If this is true, we say the estimate is unbiased.

(Ch. 2 of L&O shows that the sample mean is unbiased)

If it is not true then we say the estimate is biased.

If it is not true, but $\boxed{\lim_{N \to \infty} E\{\hat{A}\} = A}$

we say that the estimate is asymptotically unbiased.

2. We want small fluctuations from estimate to estimate:

$$\text{var}\{\hat{A}\} = \textit{small}$$

Also, we'd like $\text{var}\{\hat{A}\} \to 0 \quad as \quad N \to \infty$

(Ch. 2 of L&O shows that this is true for the sample mean)

Review of Statistics (Cont.)

Can capture both mean and variance of an estimate by using Mean-Square-Error (MSE):

$$MSE\{\hat{A}\} = \mathrm{var}\{\hat{A}\} + B^2\{\hat{A}\}$$

$$where \quad B\{\hat{A}\} = A - E\{\hat{A}\}$$

Usual goal of Estimation: Minimize MSE

- Minimize Bias
- Minimize Variance

For PSD Estimation want:

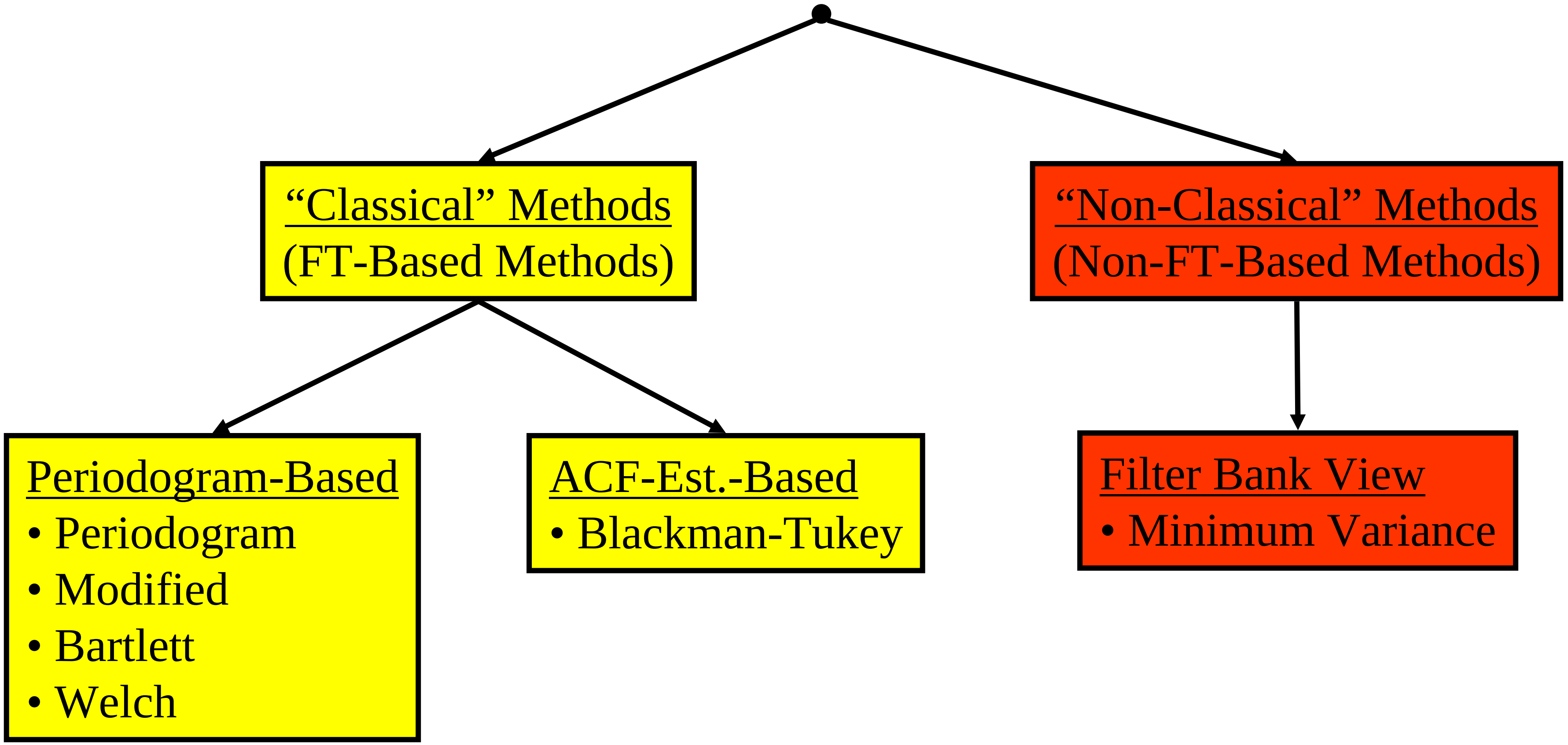
$$E\{\hat{S}_x(\omega)\} = S_x(\omega)$$

$$\mathrm{var}\{\hat{S}_x(\omega)\} = small$$

Non-Parametric Spectral Estimation H-8.2, LO-2.4

- Periodogram
- Windowed Periodogram
- Averaged Periodogram
- Windowed & Averaged Periodogram
- Blackman-Tukey Method
- Minimum Variance Method

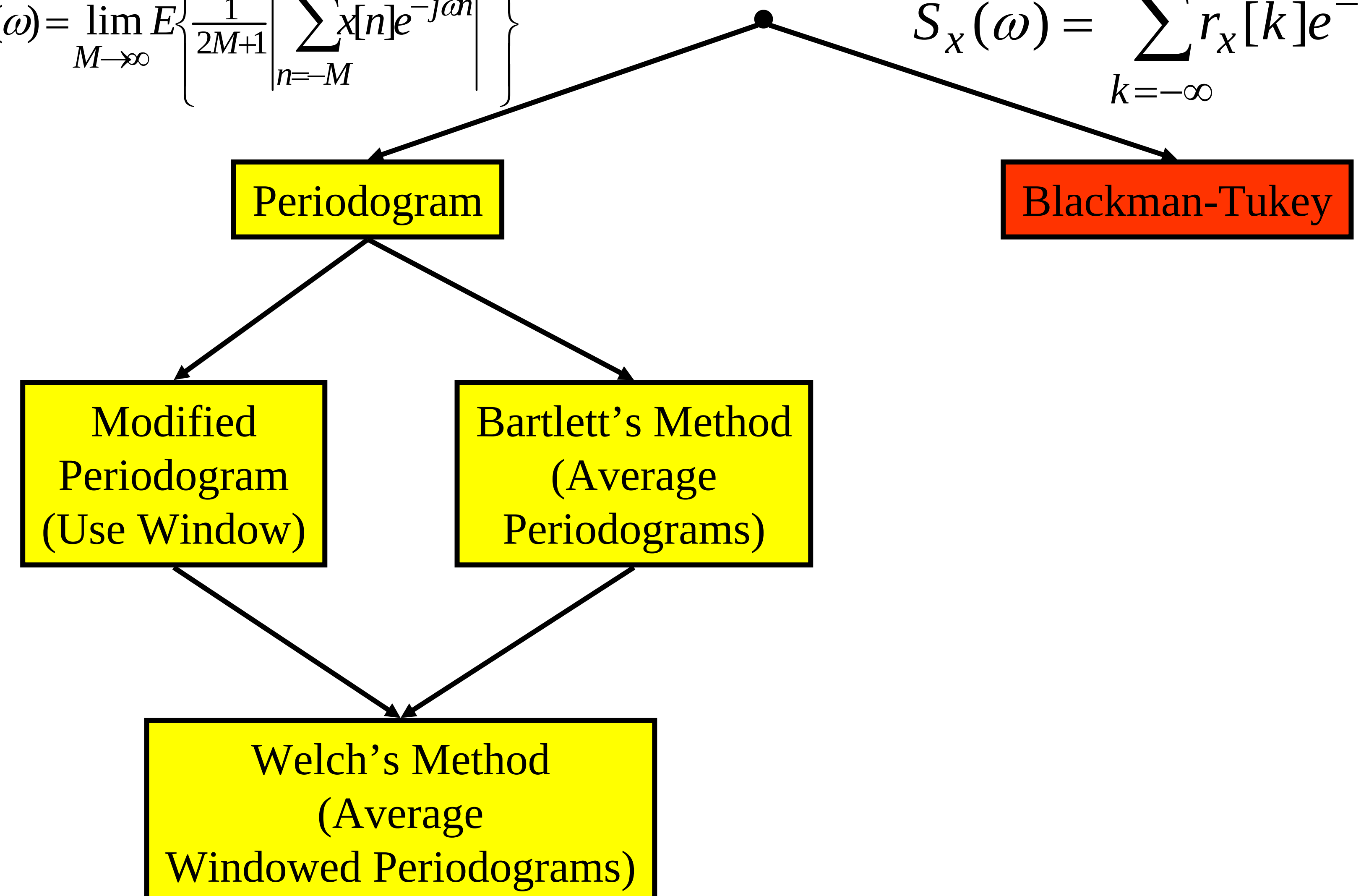
Family of Non-Parametric Methods



Family of "Classical" Methods

$$S_x(\omega) = \lim_{M\to\infty} E\left\{ \frac{1}{2M+1} \left| \sum_{n=-M}^{M} x[n] e^{-j\omega n} \right|^2 \right\}$$

$$S_x(\omega) = \sum_{k=-\infty}^{\infty} r_x[k] e^{-j\omega k}$$



The Periodogram

Periodogram - Definition

Based on:

$$S_X(\omega) = \lim_{M\to\infty} E\left\{ \frac{1}{2M+1} \left| \sum_{n=-M}^{M} x[n] e^{-j\omega n} \right|^2 \right\}$$

In practice we have one set of finite-duration data.

Two Practical Problems:

1. Can't do the expected value
2. Can't do the limit

The periodogram is a method that ignores them both!!!

$$\boxed{\hat{S}_{PER}(\omega) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j\omega n} \right|^2}$$

In practice we compute this using the DFT (possibly using zero-padding) – which computes the DTFT at discrete frequency points ("DFT Bins")

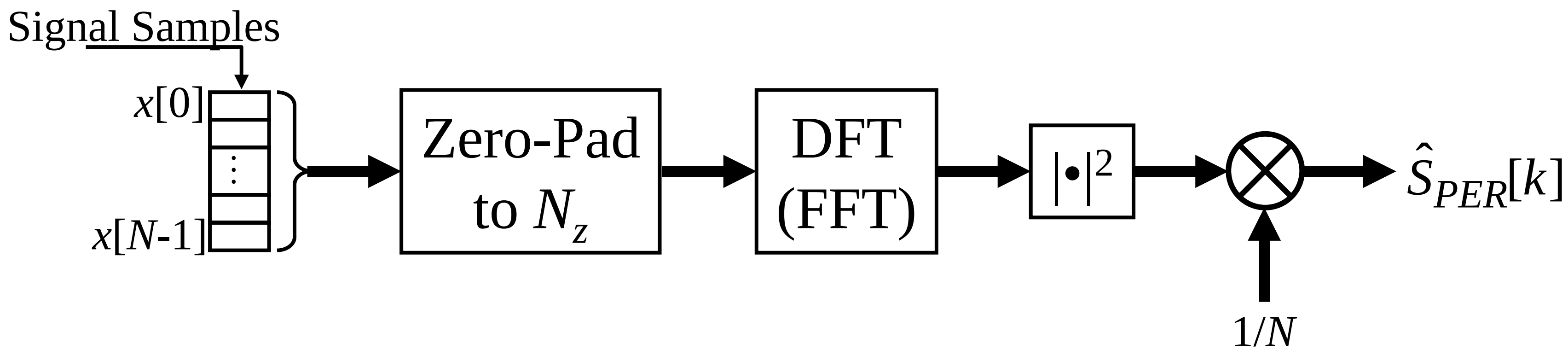
Periodogram - Computation

In practice we compute this using the DFT(FFT) (usually using zero-padding) – which computes the DTFT at discrete frequency points ("DFT Bins"):

$$\boxed{\hat{S}_{PER}[k] = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N_z}\right|^2} \qquad \omega_k = 2\pi k / N_z$$

N = number of signal samples

N_z = DFT size – after zero-padding



Periodogram – Viewed as Filter Bank

Although we ALWAYS implement the periodogram using the DFT, it is helpful to interpret it as a filter bank.

Define the impulse response of an FIR filter as:

$$h_i[n] = \begin{cases} \frac{1}{N} e^{jn\omega_i}, & 0 \le n < N \\ 0, & otherwise \end{cases}$$

Frequency Response of this filter is:

$$H(\omega) = \sum_{n=0}^{N-1} h_i[n] e^{-jn\omega}$$

$$= e^{-jn(\omega-\omega_i)(N-1)/2} \frac{\sin[N(\omega-\omega_i)/2]}{N\sin[(\omega-\omega_i)/2]}$$

<< See Figure 8.3 in Hayes>>

Periodogram – Viewed as Filter Bank (cont.)

Now the output of the ith filter is:

$$y_i[n] = x[n] * h_i[n] = \sum_{k=n-N+1}^{n} x[k] h_i[n-k]$$

$$= \frac{1}{N} \sum_{k=n-N+1}^{n} x[k] e^{j(n-k)\omega_i}$$

Now one estimate of the power at the output of this filter is: $|y_i[n]|^2$ for any value of *n*. Choosing n = *N*-1 gives the periodogram:

$$\left| y_i[N-1] \right|^2 = \left| \frac{1}{N} \sum_{k=0}^{N-1} x[k] e^{j(N-1-k)\omega_i} \right|^2 = \underbrace{\left| e^{j(N-1)\omega_i} \right|^2}_{=1} \left| \frac{1}{N} \sum_{k=0}^{N-1} x[k] e^{j-k\omega_i} \right|^2$$

$$= \left| \frac{1}{N} \sum_{k=0}^{N-1} x[k] e^{j-k\omega_i} \right|^2 = N \hat{S}_{PER}(\omega)$$

<< See Figure 8.4 in Hayes>>

<u>**Periodogram – Performance**</u>

H-8.2.2

<u>For a good PSD estimate</u> we'd like to have (at the very least):

$$\lim_{N\to\infty} E\{\hat{S}_x(\omega)\} = S_x(\omega)$$ "Asymp. UnBiased"

$$\lim_{N\to\infty} \mathrm{var}\{\hat{S}_x(\omega)\} = 0$$

Actually, we would prefer it to be unbiased even for finite *N*

Does the Periodogram have these characteristics????

Let's Find Out!!!

Periodogram – Performance: Bias

Property #1: The Periodogram is Biased.

Property #2: But… The Periodogram is *Asymptotically* Unbiased.

Proof: Taking the EV of the periodogram gives

$$E\left\{\hat{S}_{PER}(\omega)\right\} = \frac{1}{N} E\left\{\left|\sum_{n=0}^{N-1} x[n] e^{-j\omega n}\right|^2\right\}$$

$$= \frac{1}{N} E\left\{\left[\sum_{n=0}^{N-1} x[n] e^{-j\omega n}\right]\left[\sum_{m=0}^{N-1} x^*[m] e^{j\omega m}\right]\right\}$$

$$= \frac{1}{N} \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} r_x[n-m] e^{-j\omega(n-m)}$$

"Sum On Diagonals"

$r_x[n\text{-}m]$ is constant on each diagonal

$$= \sum_{k=-(N-1)}^{N-1} \underbrace{\left(1 - \frac{|k|}{N}\right)}_{w_B[k]} r_x[k] e^{-j\omega k}$$

Bartlett (Triangle) Window

$$= \frac{1}{2\pi} S_x(\omega) \underset{circ}{*} W_B(\omega) \neq S_x(\omega)$$

Periodogram – Performance: Bias (cont.)

Proof (cont.): This shows that the Periodogram is Biased.

The bias comes from the smoothing effect of Bartlett window.
(Smoothing also reduces the resolution of sharp spectral features).

$$\boxed{E\left\{\hat{S}_{PER}(\omega)\right\}=\frac{1}{2\pi}S_x(\omega)\underset{circ}{*}W_B(\omega)}$$

But… as N→∞ the Bartlett Kernel tends to a delta function in the frequency domain, or – equivalently – in the TD the Bartlett window tends 1:

$$\lim_{N\to\infty} E\left\{\hat{S}_{PER}(\omega)\right\}=\lim_{N\to\infty}\sum_{k=-(N-1)}^{N-1}\underbrace{\left(1-\frac{|k|}{N}\right)}_{\to 1}r_x[k]e^{-j\omega k}$$

$$=\sum_{k=-\infty}^{\infty}r_x[k]e^{-j\omega k}=S_x(\omega)$$

Thus, the Periodogram is Asymptotically Unbiased.

Periodogram – Performance: Variance

Property #3: The variance of the Periodogram does NOT (in general) tend to zero as $N\to\infty$.

Proof: Difficult to prove for general case… so this is proved under the assumption: complex-valued white Gaussian process w/ zero mean and variance σ^2.

Under this assumption, the true PSD and ACF are:

$$S(\omega)=\sigma^2, \quad \forall\omega \quad \& \quad r_x[k]=\sigma^2\delta[k]$$

The variance of the periodogram is what we want to analyze and is given by:

$$\operatorname{var}\left\{\hat{S}_{PER}(\omega)\right\}=E\left\{\hat{S}^2_{PER}(\omega)\right\}-\left[E\left\{\hat{S}_{PER}(\omega)\right\}\right]^2$$

Look at this term first: "Bias Term"

Periodogram – Performance: Variance (cont.)

Proof (cont.): So from our previous analysis of bias (and our assumptions on the process) we know that the second term is:

$$\left[E\left\{\hat{S}_{PER}(\omega)\right\}\right]^2 = \left[\sum_{k=-(N-1)}^{N-1}\left(1-\frac{|k|}{N}\right)r_x[k]e^{-j\omega k}\right]^2$$

$$= \left[\sum_{k=-(N-1)}^{N-1}\left(1-\frac{|k|}{N}\right)\left[\sigma^2\delta[k]\right]e^{-j\omega k}\right]^2 = \sigma^4\left[\left(1-\frac{|k|}{N}\right)e^{-j\omega k}\right]^2_{k=0}$$

$$= \sigma^4$$

(Aside: <u>under our assumptions</u> the periodogram is unbiased!)

So the variance of periodogram is now…

$$\mathrm{var}\left\{\hat{S}_{PER}(\omega)\right\} = E\left\{\hat{S}^2_{PER}(\omega)\right\} - \sigma^4 \qquad (\ast)$$

Now.. Look at This Term

Periodogram – Performance: Variance (cont.)

Proof (cont.): As a means of looking at this first term we consider:

$$E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} = E\left\{\left[\underbrace{\frac{1}{N}\left|\sum_{l=0}^{N-1} x[l]e^{-j\omega_1 l}\right|^2}_{\left[\sum_{k=0}^{N-1} x[k]e^{-j\omega_1 k}\right]\left[\sum_{l=0}^{N-1} x[l]e^{-j\omega_1 l}\right]^*}\right]\left[\underbrace{\frac{1}{N}\left|\sum_{n=0}^{N-1} x[n]e^{-j\omega_2 n}\right|^2}_{\left[\sum_{m=0}^{N-1} x[m]e^{-j\omega_2 m}\right]\left[\sum_{n=0}^{N-1} x[n]e^{-j\omega_2 n}\right]^*}\right]\right\}$$

Using these "call-outs" and manipulating gives:

$$E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} \qquad (\bigstar)$$

$$= \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} \underbrace{E\left\{x[k]x^*[l]x[m]x^*[n]\right\}}_{=\ ?????}\exp\{-j[(k-l)\omega_1 + (m-n)\omega_2]\}$$

Periodogram – Performance: Variance (cont.)

Proof (cont.): Now what is this Expected Value???? Well… since we assumed the process is Gaussian we can use a standard result for complex **JOINTLY** Gaussian RVs:

$$E\left\{x[k]x^*[l]x[m]x^*[n]\right\} = E\left\{x[k]x^*[l]\right\}E\left\{x[m]x^*[n]\right\} + E\left\{x[k]x^*[n]\right\}E\left\{x[m]x^*[l]\right\}$$

Now… using this result together with the assumption of whiteness:

$$E\left\{x[l]x[k]x[n]x[m]\right\}$$
$$= \sigma^4\left[\delta[k-l]\delta[m-n] + \delta[k-n]\delta[m-l]\right]$$

Now… using this result in (★) gives:

Periodogram – Performance: Variance (cont.)

$$E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\}$$

$$= \frac{\sigma^4}{N^2}\left[\sum_{l=0}^{N-1}\sum_{k=0}^{N-1}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\delta[k-l]\delta[m-n]\exp\{-j[\omega_1(k-l)+\omega_2(m-n)]\}\right.$$

$$\left. + \sum_{l=0}^{N-1}\sum_{k=0}^{N-1}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\delta[k-n]\delta[m-l]\exp\{-j[\omega_1(k-l)+\omega_2(m-n)]\}\right]$$

$$= \frac{\sigma^4}{N^2}\left[\sum_{l=0}^{N-1}\sum_{n=0}^{N-1}1 + \sum_{l=0}^{N-1}\sum_{k=0}^{N-1}\exp\{-j[\omega_1(k-l)-\omega_2(k-l)]\}\right]$$

$$= \frac{\sigma^4}{N^2}\left[\underbrace{\sum_{l=0}^{N-1}\sum_{n=0}^{N-1}1}_{= N^2} + \underbrace{\sum_{l=0}^{N-1}\sum_{k=0}^{N-1}\exp\{-j[(\omega_1-\omega_2)(k-l)]\}}_{\text{Use "Sum On Diagonals" Trick}}\right]$$

Periodogram – Performance: Variance (cont.)

Proof (cont.):

$$E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\}$$

$$= \frac{\sigma^4}{N^2}\left[N^2 + N\sum_{k=-(N-1)}^{N-1}\underbrace{\left(1-\frac{|k|}{N}\right)}_{w_B[k]}\exp\{-j[(\omega_1-\omega_2)k]\}\right]$$

Now the FT of the Bartlett window is: $\mathfrak{F}\{w_B[k]\} = \left(\frac{\sin(N\omega/2)}{\sin(\omega/2)}\right)^2$
So using it in the above result gives:

$$E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} = \sigma^4\left[1+\left(\frac{\sin[N(\omega_1-\omega_2)/2]}{N\sin[(\omega_1-\omega_2)/2]}\right)^2\right]$$

Periodogram – Performance: Variance (cont.)

Proof (cont.): To find the first term in the variance expression of interest (✸) we must set $\omega = \omega_1 = \omega_2$ in the above expression to get:

$$E\left\{\hat{S}_{PER}^2(\omega)\right\} = 2\sigma^4$$

Now using this in the expression for variance (✸) gives

$$\begin{aligned} \mathrm{var}\left\{\hat{S}_{PER}(\omega)\right\} &= E\left\{\hat{S}_{PER}^2(\omega)\right\} - \sigma^4 \\ &= 2\sigma^4 - \sigma^4 \end{aligned}$$

$$\boxed{\mathrm{var}\left\{\hat{S}_{PER}(\omega)\right\} = \sigma^4}$$

…which DOES NOT go to zero as $N \to \infty$

Periodogram – Performance: Covariance

Property #4: Increasing N leads to rapidly fluctuating periodograms (even where the true PSD is smooth).

"Proof": Use the previous results, the covariance of the periodogram is given by

$$\operatorname{cov}\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} = E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} - E\left\{\hat{S}_{PER}(\omega_1)\right\}E\left\{\hat{S}_{PER}(\omega_2)\right\}$$

$$= \sigma^4\left[1+\left(\frac{\sin[N(\omega_1-\omega_2)/2]}{N\sin[(\omega_1-\omega_2)/2]}\right)^2\right]-\sigma^4$$

$$= \sigma^4\left(\frac{\sin[N(\omega_1-\omega_2)/2]}{N\sin[(\omega_1-\omega_2)/2]}\right)^2$$

Covariance is a measure of how correlated two RVs are. Thus, cov(X,Y)=0 indicates that there is a high probability that X & Y will be very unalike.

Now, the equation above indicates there are (ω_1,ω_2) pairs for which the cov of the periodogram is zero.

➔ **Periodogram Fluctuates Rapidly from freq-to-freq**

Periodogram – Performance for Non-White RP

The above analysis was done for white noise. Hayes p. 407 gives an argument that shows similar results for the non-white case:

$$\boxed{\operatorname{var}\left\{\hat{S}_{PER}(\omega)\right\} \approx S_x^2(\omega)}$$

$$\boxed{E\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} \approx S_x(\omega_1)S_x(\omega_2)\left[1+\left(\frac{\sin[N(\omega_1-\omega_2)/2]}{N\sin[(\omega_1-\omega_2)/2]}\right)^2\right]}$$

$$\boxed{\operatorname{cov}\left\{\hat{S}_{PER}(\omega_1)\hat{S}_{PER}(\omega_2)\right\} \approx S_x(\omega_1)S_x(\omega_2)\left(\frac{\sin[N(\omega_1-\omega_2)/2]}{N\sin[(\omega_1-\omega_2)/2]}\right)^2}$$

Periodogram – Examples

1. **Bias – Effect of Window**
   Periodogram of Sinusoid: **See Hayes Fig. 8.5**

2. **Variance and Covariance**
   Periodogram of White Noise: **See L&O Fig. 2.4**

   Periodogram of Sinusoid: **See Hayes Fig. 8.6**

3. **Resolution – Effect of Window**
   Periodogram of 2 Sinusoids: **See Hayes Fig. 8.8**