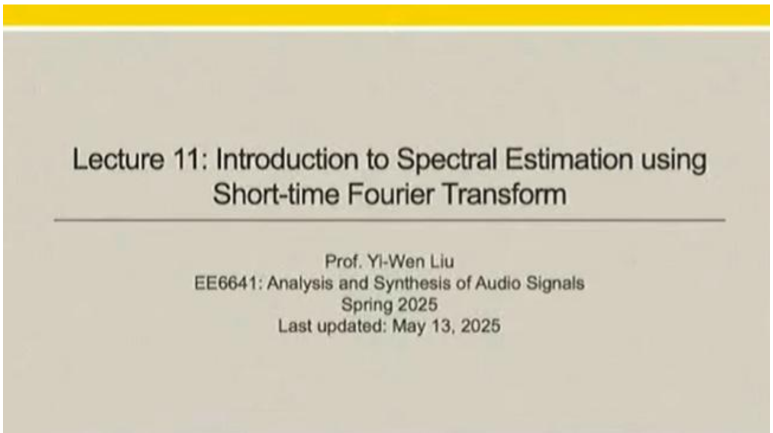
Lecture11 Introduction to Spectral Estimation using Short-time



A question and a story

- Why was it necessary for the buttons to sound?
- A surprising hack that used to work

NTHU campus
Dec. 2016

A question and a story

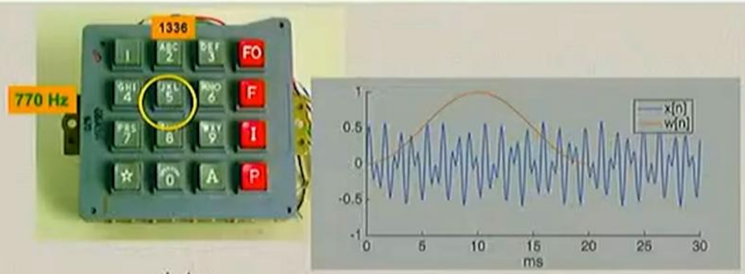
- Why was it necessary for the buttons to sound?
- A surprising hack that used to work

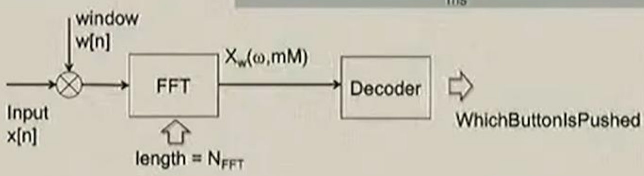
NTHU campus
Dec. 2016

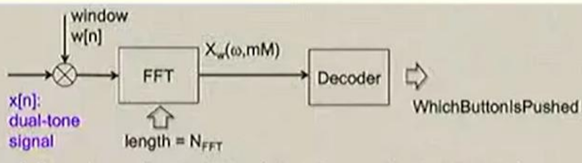
Outlines

- The classic *dual tone multiple frequency* technology (Part A)
  - Choice of window
  - Resolution requirement
- Resolution vs. accuracy (Part B)
  - Interference vs. noise
- The *Fisher information* and the *Cramer-Rao bound* (Optional)*

Part A: *Let's begin spectral estimation*!
Dual Tone Multi-Frequency Decoding



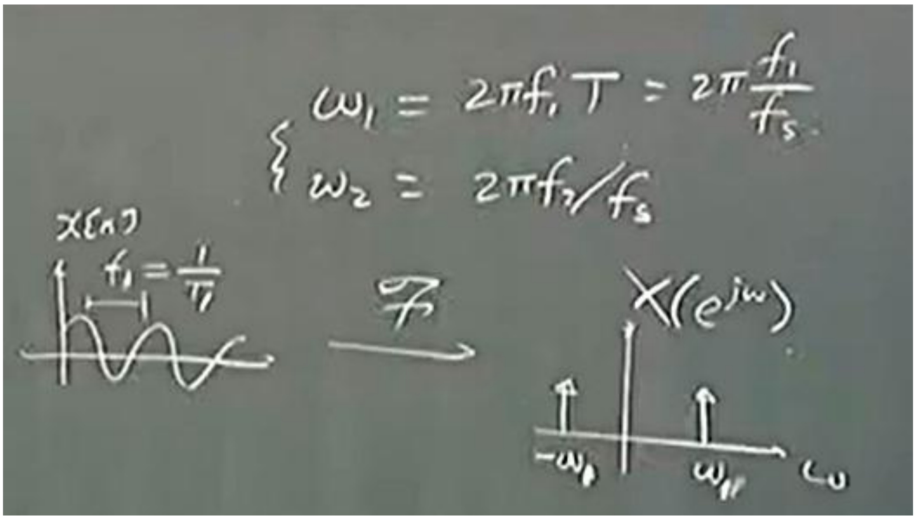


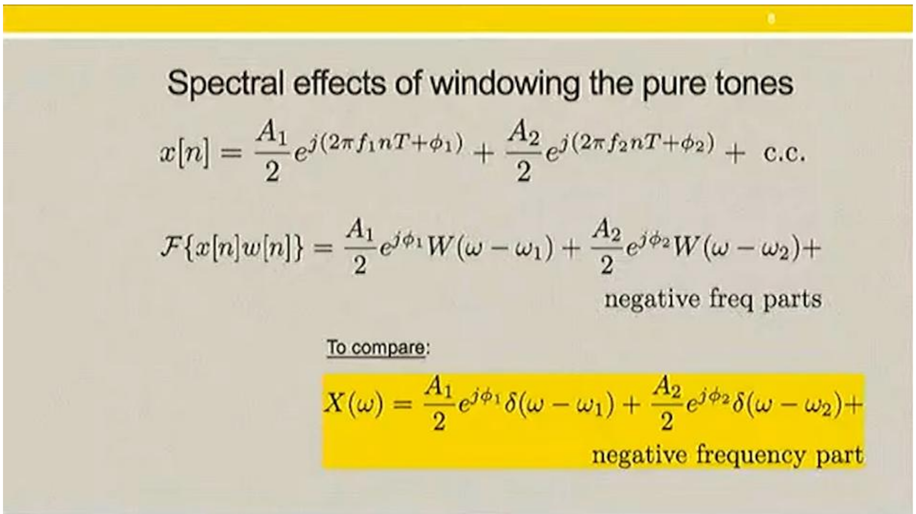


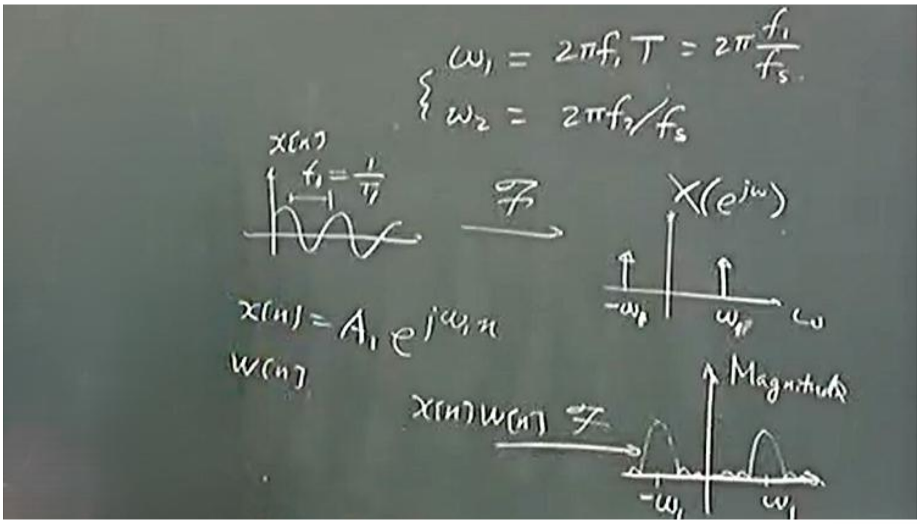
$$\begin{aligned} x[n] &= A_1\cos(2\pi f_1 nT + \phi_1) + A_2\cos(2\pi f_2 nT + \phi_2) \\ &= \mathrm{Re}\{A_1 e^{j(2\pi f_1 nT + \phi_1)} + A_2 e^{j(2\pi f_2 nT + \phi_2)}\} \\ &= \frac{A_1}{2} e^{j(2\pi f_1 nT + \phi_1)} + \frac{A_2}{2} e^{j(2\pi f_2 nT + \phi_2)} + \text{ complex conjugates} \end{aligned}$$

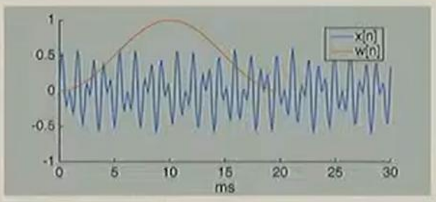
$$X(\omega) = \frac{A_1}{2} e^{j\phi_1}\delta(\omega - \omega_1) + \frac{A_2}{2} e^{j\phi_2}\delta(\omega - \omega_2) + \text{negative frequency part}$$

Question:


$\omega_1 = 2\pi f_1 T = 2\pi \frac{f_1}{f_s}$
$\omega_2 = 2\pi f_2 / f_s$
x[n]
$f_1 = \frac{1}{T_1}$
$X(e^{j\omega})$
$-\omega_1$
ω_1
ω


Spectral effects of windowing the pure tones
$x[n] = \frac{A_1}{2} e^{j(2\pi f_1 nT + \phi_1)} + \frac{A_2}{2} e^{j(2\pi f_2 nT + \phi_2)} + \text{ c.c.}$
$\mathcal{F}\{x[n]w[n]\} = \frac{A_1}{2} e^{j\phi_1} W(\omega - \omega_1) + \frac{A_2}{2} e^{j\phi_2} W(\omega - \omega_2) +$
negative freq parts
To compare:
$X(\omega) = \frac{A_1}{2} e^{j\phi_1} \delta(\omega - \omega_1) + \frac{A_2}{2} e^{j\phi_2} \delta(\omega - \omega_2) +$
negative frequency part


$\omega_1 = 2\pi f_1 T = 2\pi \frac{f_1}{f_s}$
$\omega_2 = 2\pi f_2 / f_s$
x[n]
$f_1 = \frac{1}{T_1}$
$X(e^{j\omega})$
$x[n] = A_1 e^{j\omega_1 n}$
w[n]
x[n]w[n]
Magnitude
$-\omega_1$
ω_1



$$x[n] = \frac{A_1}{2} e^{j(2\pi f_1 nT + \phi_1)} + \frac{A_2}{2} e^{j(2\pi f_2 nT + \phi_2)} + \text{ c.c.}$$

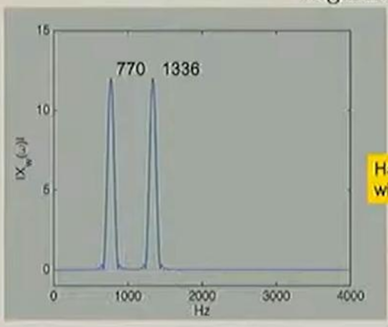
$$\mathcal{F}\{x[n]w[n]\} = \frac{A_1}{2} e^{j\phi_1} W(\omega - \omega_1) + \frac{A_2}{2} e^{j\phi_2} W(\omega - \omega_2) +$$

negative freq parts

$$x[n] = \frac{A_1}{2} e^{j(2\pi f_1 nT + \phi_1)} + \frac{A_2}{2} e^{j(2\pi f_2 nT + \phi_2)} + \text{ c.c.}$$

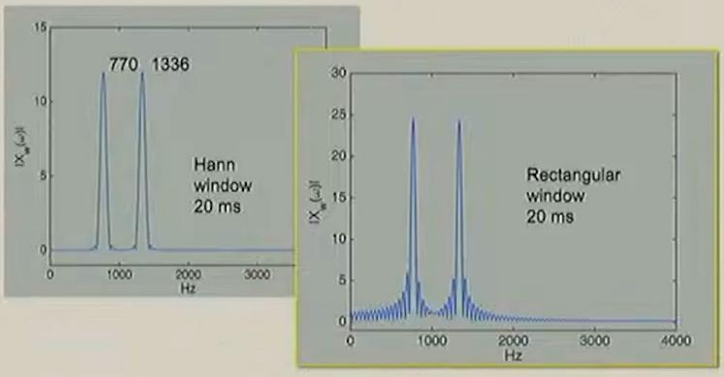
$$\mathcal{F}\{x[n]w[n]\} = \frac{A_1}{2} e^{j\phi_1} W(\omega - \omega_1) + \frac{A_2}{2} e^{j\phi_2} W(\omega - \omega_2) +$$

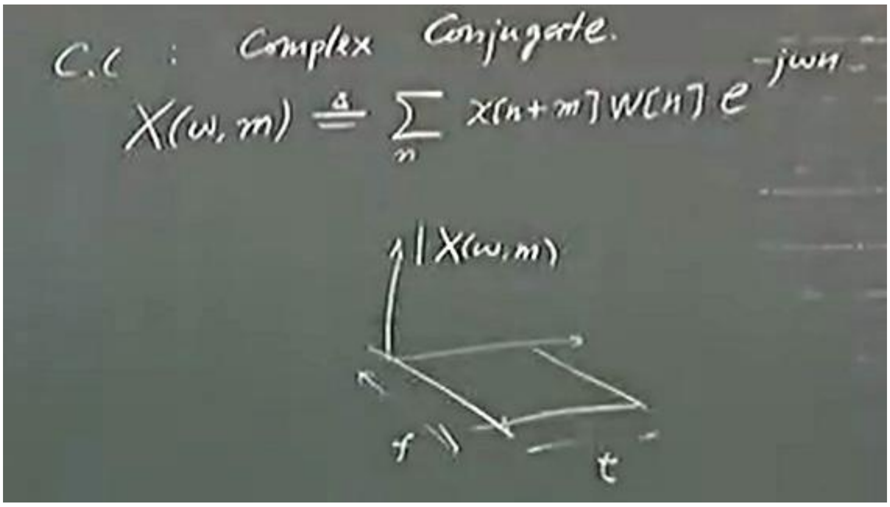
negative freq parts

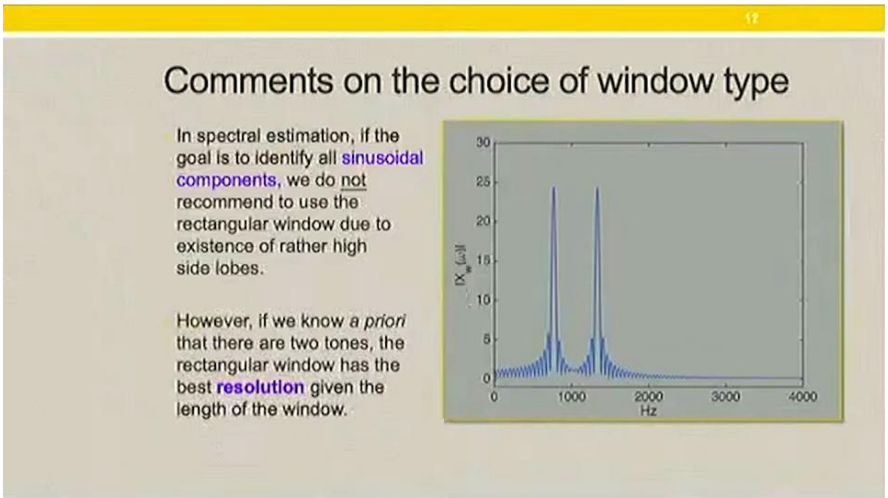


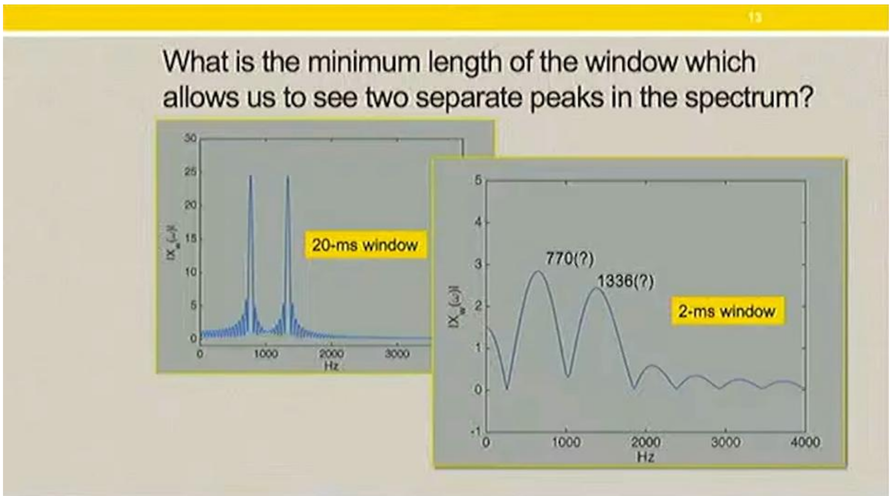
Hann window
window length = 20 ms

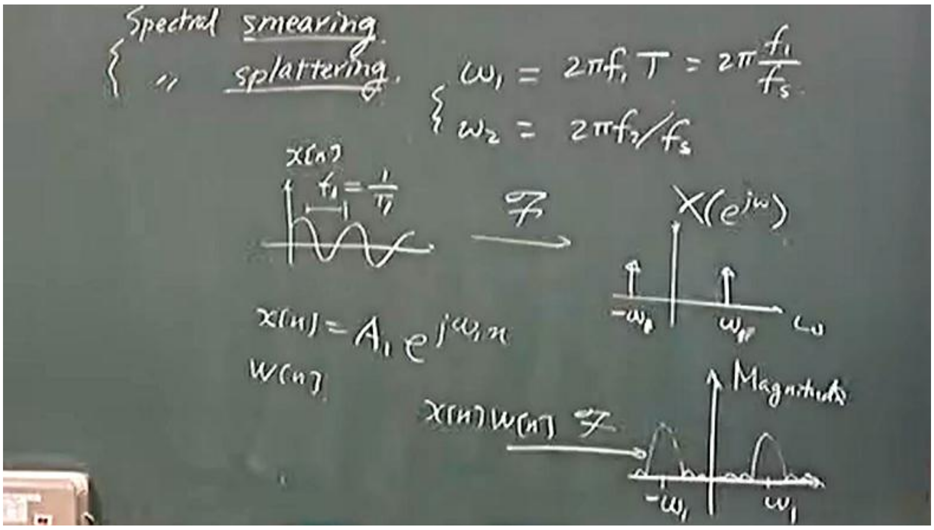
Choice of the window




C.C : Complex Conjugate.
$X(\omega, m) \triangleq \sum_{n} x[n+m]\,w[n]\,e^{-j\omega n}$
$|X(\omega, m)|$
f
t

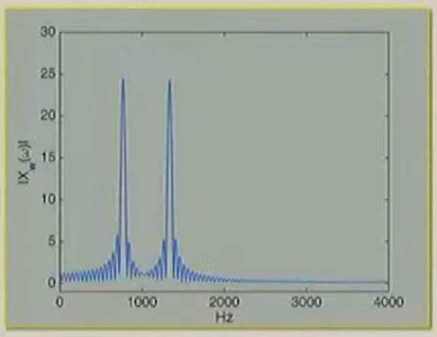

12
Comments on the choice of window type
In spectral estimation, if the goal is to identify all sinusoidal components, we do not recommend to use the rectangular window due to existence of rather high side lobes.
However, if we know a priori that there are two tones, the rectangular window has the best resolution given the length of the window.
Hz


13
What is the minimum length of the window which allows us to see two separate peaks in the spectrum?
20-ms window
770(?)
1336(?)
2-ms window
Hz



Comments on the choice of window type

- In spectral estimation, if the goal is to identify all sinusoidal components, we do not recommend to use the rectangular window due to existence of rather high side lobes.
- However, if we know *a priori* that there are two tones, the rectangular window has the best **resolution** given the length of the window.

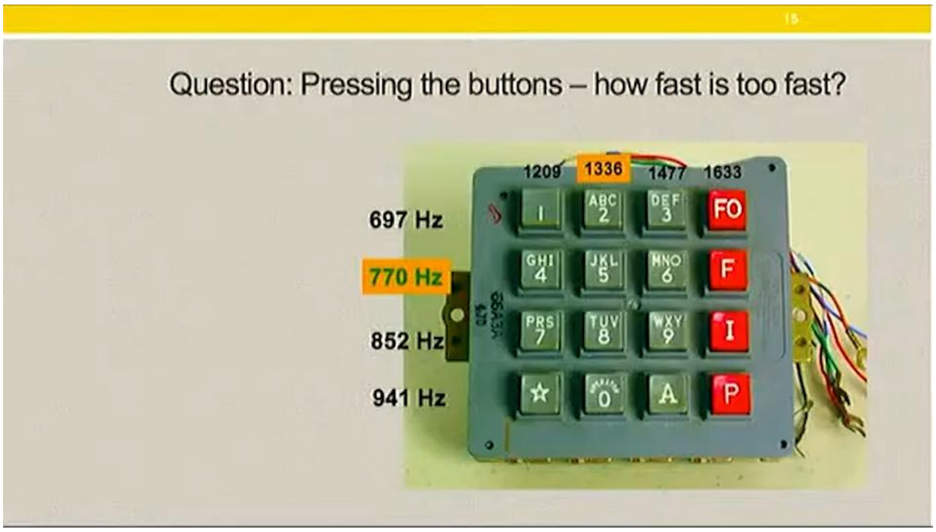


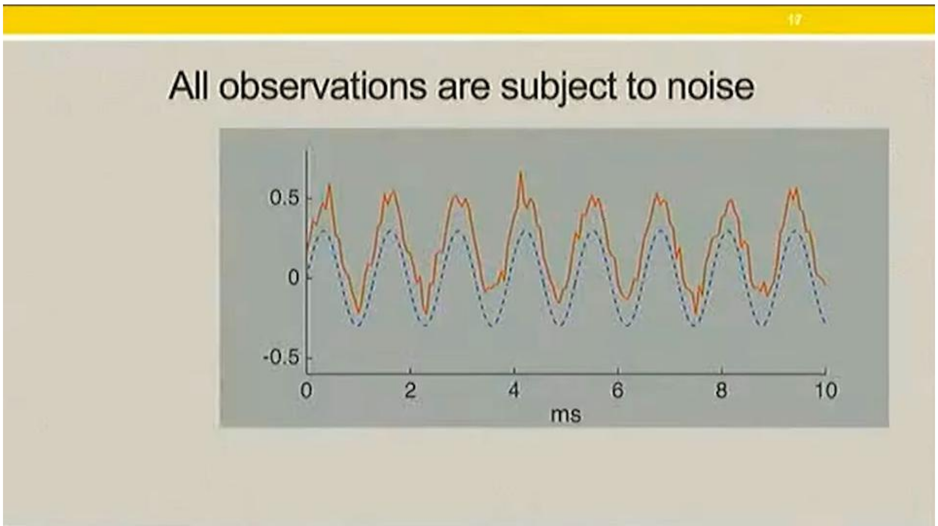
Remarks on the resolution in frequency

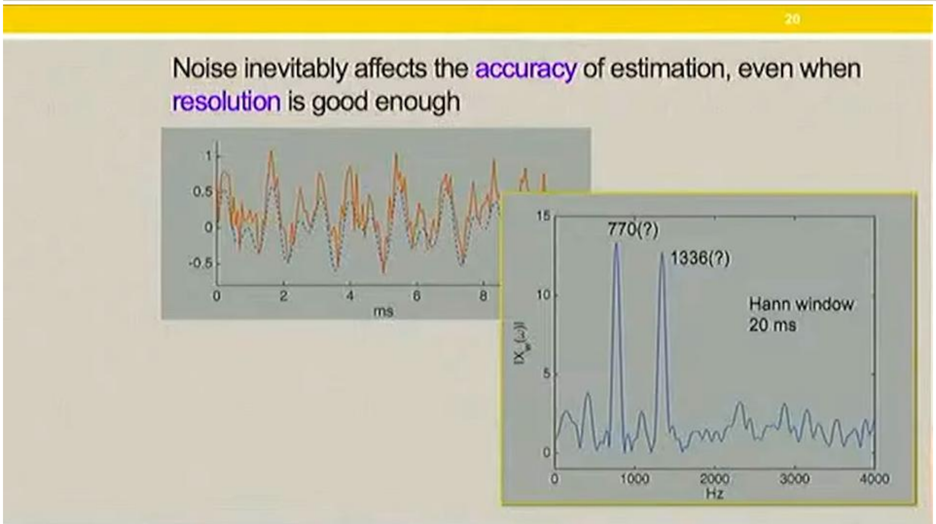
- Since the mainlobe width in the spectrum of a window is inversely proportional to the duration of the window in time, we need to use a sufficiently long window so as to resolve two spectral components.

$$(\Delta f) \propto (\Delta t)^{-1}$$

E.g., f_1 = 770 Hz, f_2 = 1336 Hz
=> Δf = 566 Hz, $1/\Delta f$ = 1.8 ms.


Question: Pressing the buttons – how fast is too fast?
1209
1336
1477
1633
697 Hz
770 Hz
852 Hz
941 Hz


All observations are subject to noise
0.5
0
-0.5
0
2
4
6
8
10
ms


Noise inevitably affects the accuracy of estimation, even when resolution is good enough
770(?)
1336(?)
Hann window
20 ms
Hz

21

Summary:
Resolution and accuracy are two different issues

- Resolution is a concern against interference from concurrent sinusoidal components.
- Accuracy is a concern against noise.
- Resolution: $(\Delta f) \propto (\Delta t)^{-1}$
- Accuracy: $(\Delta f)^2 \propto (\Delta t)^{-3} \cdot \frac{\sigma^2}{A^2}$

Remarks:
A^2/σ^2 is the **signal to noise ratio** (SNR)

22

Question:
how do we approach the limit of (frequency) estimation?

$$(\Delta f)^2 \propto (\Delta t)^{-3} \cdot \frac{\sigma^2}{A^2}$$

- Suggestion: use quadratic interpolation in the *log spectrum*
- Fisher information and Cramer-Rao bound (CRB)

23

4D: The Fisher information and the Cramer-Rao bound

Problem formulation:

Parameter estimation from noisy observation

$$\mathbf{x} = \mathbf{s} + \mathbf{u} \in \mathbb{R}^N$$

$$\mathbf{s} = \mathbf{s}(\theta) \qquad \mathbf{u} \sim f_{\mathbf{u}}(.)$$

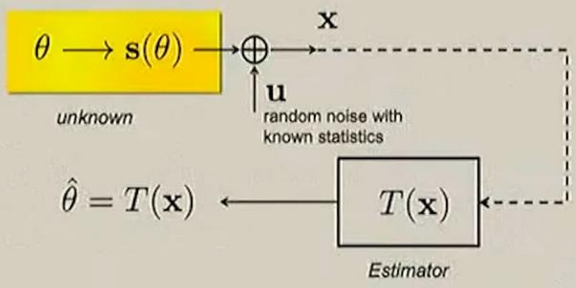
θ: parameter.
s: parameterized signal.

random vector, noise

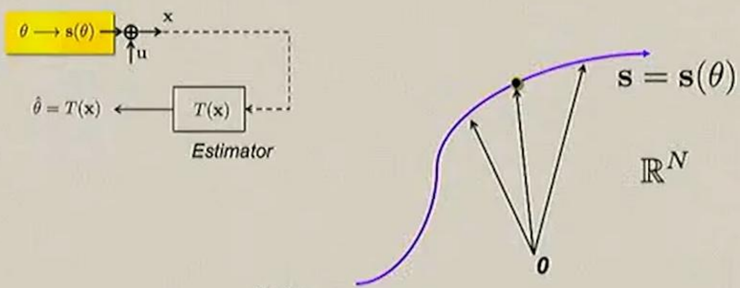
$$\mathbf{x} \sim f(\mathbf{x}; \theta)$$

Question:
How do we best estimate θ by looking at **x**?

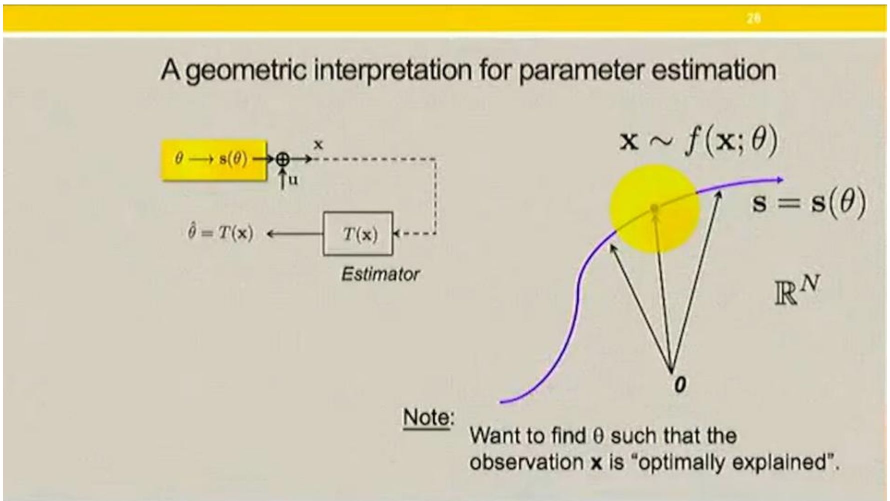
Parameter estimation block diagram

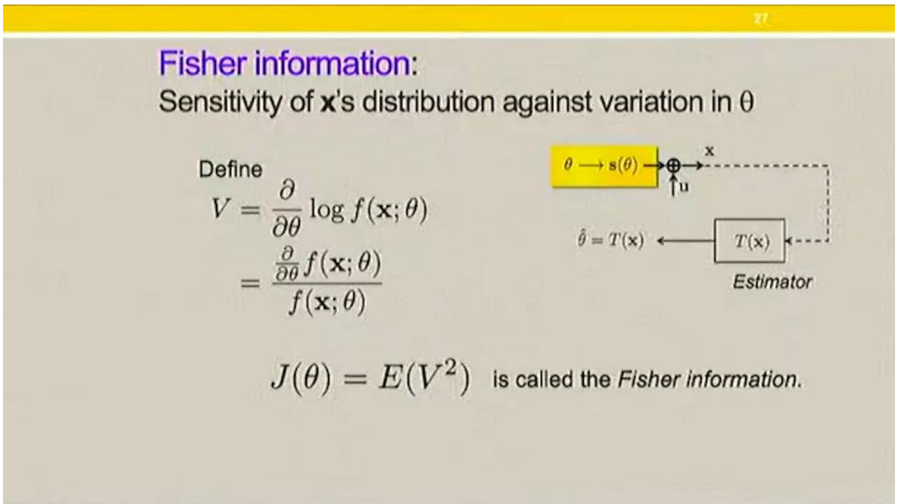


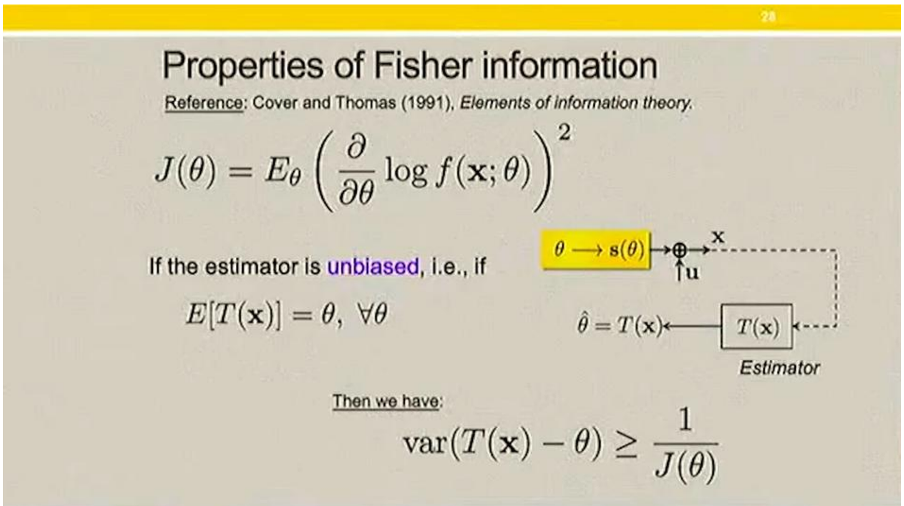
A geometric interpretation for parameter estimation



Note:


26
A geometric interpretation for parameter estimation
$\theta \longrightarrow \mathbf{s}(\theta)$
x
u
$\hat{\theta} = T(\mathbf{x})$
$T(\mathbf{x})$
Estimator
$\mathbf{x} \sim f(\mathbf{x};\theta)$
$\mathbf{s} = \mathbf{s}(\theta)$
$\mathbb{R}^N$
0
Note: Want to find θ such that the observation **x** is "optimally explained".


27
Fisher information:
Sensitivity of **x**'s distribution against variation in θ
Define
$V = \frac{\partial}{\partial\theta} \log f(\mathbf{x};\theta) = \frac{\frac{\partial}{\partial\theta} f(\mathbf{x};\theta)}{f(\mathbf{x};\theta)}$
$J(\theta) = E(V^2)$ is called the *Fisher information*.
$\theta \longrightarrow \mathbf{s}(\theta)$
x
u
$\hat{\theta} = T(\mathbf{x})$
$T(\mathbf{x})$
Estimator


28
Properties of Fisher information
Reference: Cover and Thomas (1991), *Elements of information theory.*
$J(\theta) = E_\theta \left(\frac{\partial}{\partial\theta} \log f(\mathbf{x};\theta) \right)^2$
If the estimator is unbiased, i.e., if
$E[T(\mathbf{x})] = \theta, \ \forall\theta$
$\theta \longrightarrow \mathbf{s}(\theta)$
x
u
$\hat{\theta} = T(\mathbf{x})$
$T(\mathbf{x})$
Estimator
Then we have:
$\mathrm{var}(T(\mathbf{x}) - \theta) \geq \frac{1}{J(\theta)}$

29

Example: single-frequency estimation

Signal model:

$$s[n] = A\cos(\omega n + \phi), n = 0, 1, 2, ..., N-1$$

Noise model: Additive white Gaussian noise (AWGN)

$$u[n] \sim \mathcal{N}(0, \sigma^2), \text{i.i.d.}$$

Mission:

$$\text{from } \mathbf{x} = (x[0], x[1], ..., x[N-1])^T, \text{ find } \omega.$$

30

Single-frequency estimation under AWGN

Reference: Rife & Boorstyn (1974). IEEE Trans. Info. Theory, 20(5), 591-598.

$$\log f(\mathbf{x};\omega) = C + \frac{-1}{2\sigma^2}\sum_{n=0}^{N-1}\left(x[n] - A\cos(\omega n + \phi)\right)^2$$

$$J(\omega) = E\Big(\partial_\omega \log f(\mathbf{x};\omega)\Big)^2$$

$$= \frac{A^2}{\sigma^2}Q_3(N)$$

← $Q_3(N)$ is a third-order polynomial.

Therefore,

$$E(\hat{\theta} - \theta)^2 \geq \text{CRB} = \frac{1}{J(\omega)} \propto \frac{\sigma^2}{A^2N^3}$$

when N is sufficiently large.