

# Lab Topic 4

# SNR of LR test

- ▶ Joint pdf of noise:

$$p_{\bar{E}}(\bar{y}) = \frac{1}{(2\pi)^{N/2} |\mathbf{C}|^{1/2}} \exp\left(-\frac{1}{2} \|\bar{y}\|^2\right)$$
$$\|\bar{y}\|^2 = \bar{y} \mathbf{C}^{-1} \bar{y}^T$$

- ▶ Let  $\langle \bar{z}, \bar{y} \rangle = \bar{z} \mathbf{C}^{-1} \bar{y}^T$
- ▶ LR statistic :  $\Lambda = \frac{1}{2} \|\bar{y}\|^2 - \frac{1}{2} \|\bar{y} - \bar{s}\|^2 = \langle \bar{y}, \bar{s} \rangle - \frac{1}{2} \langle \bar{s}, \bar{s} \rangle \rightarrow \Lambda = \langle \bar{y}, \bar{s} \rangle$
- ▶ SNR:  $\|\bar{s}\|$

# Inner product for stationary noise

- White Gaussian Noise

$$\langle \bar{x}, \bar{y} \rangle \rightarrow \sum_{k=0}^{N-1} x_k y_k$$

- Stationary Gaussian noise with Power Spectral Density (PSD)  $S_n(f)$

$$\langle \bar{x}, \bar{y} \rangle \rightarrow \frac{\Delta}{N} \tilde{x} (\tilde{y}^\dagger ./ \bar{S}_n^T) = \frac{1}{T} \tilde{x} (\tilde{y}^\dagger ./ \bar{S}_n^T)$$

Where  $\tilde{x} = F\bar{x}$  is the DFT

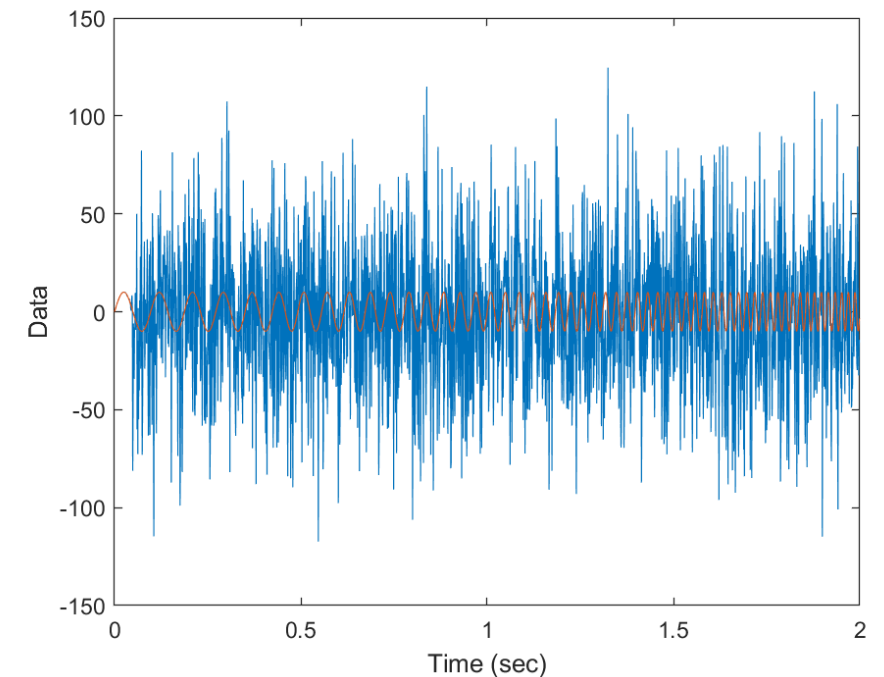
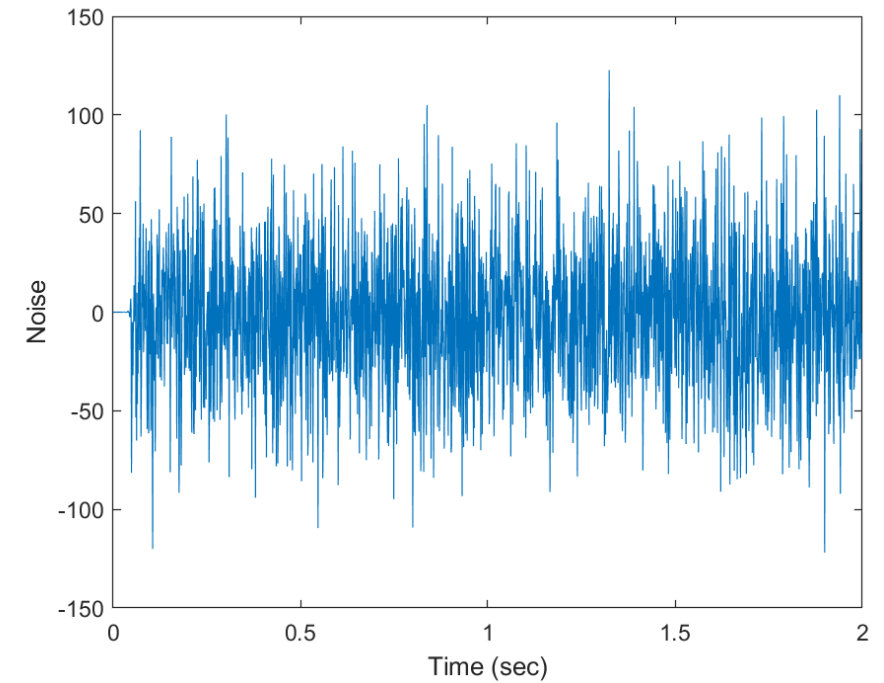
$T$  is the length of the data ( $\frac{\Delta}{N} = \frac{1}{N f_s} = \frac{1}{T}$ )

$\bar{S}_n$  is  $S_n(f)$  evaluated at the DFT frequencies

# Normalization of signal amplitude: SNR

## ► Objective:

1. Generate an  $N$  sample realization of colored gaussian noise with some specified Power Spectral Density (PSD)
2. Normalize a given signal vector of  $N$  samples such that the Likelihood Ratio test for noise with the above PSD has a specified Signal to Noise Ratio (SNR)



# Codes

- ▶ Generating stationary Gaussian noise: **NOISE/statgaussnoisegen.m**
  - ▶ Script **NOISE/colGaussNoiseDemo.m** has been modified to use this function
  - ▶ The function supplies the right scale factor for the generated noise (earlier we only wanted to get the shape of the PSD right, not its overall scale)
- ▶ Taking an inner product: **DETEST/innerprodpsd.m**
- ▶ Demo for normalization of signal: **DETEST/SNRcalc.m**

# Exercise #1

- ▶ Follow the script **DETEST/SNRcalc.m** and write a function that calculates the normalization factor for a given signal and noise PSD
  - ▶ Inputs:
    - ▶ Signal vector
    - ▶ Sampling frequency
    - ▶ PSD at positive DFT frequencies
    - ▶ SNR
  - ▶ Outputs:
    - ▶ Normalized signal vector
    - ▶ Normalization factor
- ▶ **Code should do a sanity check:** Number of PSD values should match the number of positive DFT frequencies based on the length of the signal vector

## Exercise #2

- ▶ Generate a realization of initial LIGO noise using **NOISE/statgaussnoisegen.m**
- ▶ Add the signal your team was assigned with an SNR=10 to the noise realization
- ▶ Time domain: Plot the data (signal+noise) realization and the signal
- ▶ Frequency domain: Plot the periodogram of the noise and data together
- ▶ Time-frequency domain: Plot the spectrogram of the data
- ▶ Can you see the signal in any of the three domains? Namely, the time, or the Fourier, or the time-frequency domains

# Exercise #3

- ▶ Follow **DETEST/SNRcalc.m** to write a function to calculate the Likelihood Ratio for a given data vector
  - ▶ What are the inputs needed?
- ▶ Run this function over  $M$  realizations of data under  $H_0$  and  $M$  realizations of data under  $H_1 \rightarrow$  Estimate the SNR of the LR test as shown in SNRcalc.m
- ▶ Does the estimated SNR come close to the one you had normalized the signal with? (It should!)



# Test significance

- ▶ For any detection statistic  $\Gamma(\bar{y})$ , a function of the data  $\bar{y}$ , the **significance** of detection made with the statistic is defined as :
  - ▶ “The probability of observing a value of  $\Gamma \geq$  its observed value for given data if the null hypothesis is true”

$$\Pr\{\Gamma \geq \Gamma^{(observed)} | H_0\} = \int_{\Gamma^{(observed)}}^{\infty} \underbrace{p_{\Gamma}(x|H_0)}_{Estimated} dx$$

$\Gamma^{(observed)} = \Gamma(\bar{y})$ : The value of  $\Gamma$  obtained for the given data  $\bar{y}$

# Estimating probability

$$\Pr\{\Gamma \geq \Gamma^{(observed)} | H_0\} = \int_{\Gamma^{(observed)}}^{\infty} \underbrace{p_{\Gamma}(x|H_0)}_{\text{Estimated}} dx$$

- ▶ Frequentist probability: Obtain a large number of trial values of  $\Gamma$ 
  - ▶ Trial value  $\Rightarrow$  generate a data realization and compute the  $\Gamma$  value for it
  - ▶ Then,

$$\Pr\{\Gamma \geq x\} \approx \frac{\text{Number of times } \Gamma \geq x}{\text{Total number of data realizations}}$$

- ▶ The approximation above becomes better as the total number of trials increases

# Exercise #4: Part 1

- ▶ Write a function to calculate the GLRT for the unknown amplitude case

$$L_G = \langle \bar{y}, \bar{q} \rangle^2 \text{ where } \bar{q} = \frac{\bar{s}}{\|\bar{s}\|}$$

- ▶ Inputs:
  - ▶ Data vector  $\bar{y}$
  - ▶ Noise PSD at positive DFT frequencies
- ▶ Output:
  - ▶ GLRT value
- ▶ This function is similar to the one you wrote earlier in Exercise #3 except that the data inner product is computed with  $\bar{q}$ , which is the signal vector normalized to have SNR=1
- ▶ **Note:** A new function **DETEST/normsig4psd.m** is available for normalizing a signal (it is the [solution](#) to Lab Topic 4→Exercise #1)

## Exercise #4: Part 2

- ▶ You have been provided 3 data realizations named **DETEST/data<n>.mat** where  $n = 0, 1, 3$
- ▶ You are told that the signal in each data realization is a quadratic chirp but the amplitude is unknown (maybe zero also!)
- ▶ Signal parameters:  $a_1=10$ ;  $a_2=3$ ;  $a_3=3$ ;
- ▶ You will calculate the GLRT for this signal, using the function coded in **Exercise #4: Part1**, for each of the 3 data realizations
  - ▶ The quadratic chirp signal is the one coded in **DSP/crcbgenqcsig.m**
  - ▶ and the noise PSD used is the one in **DETEST/SNRCalc.m**  
`noisePSD = @(f) (f>=100 & f<=300).*(f-100).*(300-f)/10000 + 1;`
- ▶ Use simulations to estimate the significance of the GLRT values for the 3 data realizations
- ▶ Finally, **git push** a text files into **DETEST** called **Team<n>\_testSig.txt** containing the significances you found for the 3 data realizations (each value on a new line starting from  $n = 1$ ). In the last line of the file, write the number of data realizations your simulated to estimate the significance

# Exercise #5

- ▶ Extend the function in Exercise #4 Part 1 to **create a new function**:
  - ▶ Inputs:
    - ▶ Data vector
    - ▶ **Parameter values**  $a_1, a_2, a_3$  for a quadratic chirp
    - ▶ Noise PSD vector at positive DFT frequencies
  - ▶ Output:
    - ▶ value of Likelihood ratio maximized over amplitude:  $\langle \bar{y}, \bar{q}(\Theta) \rangle^2$   
where  $\Theta$  is the set of input parameters  $a_1, a_2, a_3$
- ▶ See the script **DETEST/LRcalc.m** for an example of the steps involved for a specific set of values for  $\Theta$
- ▶ **Test**: Repeat **Exercise #4 Part 2** using this function: You should get identical values for the significance

# Conclusion: Lab Topic 4

- ▶ The function created in Exercise #5 will serve as the **fitness function** that will be optimized by PSO
- ▶ This optimization problem and the results associated with it are fully discussed in the textbook
- ▶ The main difference is that we have generalized the textbook problem to the case of **colored Gaussian noise**