Advanced signal processing techniques
Second assignment

Theodoros Katzalis
AEM: 9282
katzalis@auth.gr

Aristotle University of Thessaloniki

April 22, 2021

1 Description

Consider the real discrete process given by

$$X[k] = \sum_{i=1}^{6} \cos(\omega_i k + \varphi_i)\,, k = 0,1,\dots,N-1,$$

where $\omega_i = 2\pi\lambda_i$, $\lambda_3 = \lambda_1 + \lambda_2$ and $\lambda_6 = \lambda_4 + \lambda_5$, $\varphi_3 = \varphi_1 + \varphi_2$, $\varphi_6 = \varphi_4 + \varphi_5$ and $\varphi_1, \varphi_2, \varphi_4, \varphi_5$ are independent and uniformly distributed random variables on $[0,2\pi]$. Consider that $\lambda_1 = 0.12Hz$, $\lambda_2 = 0.30Hz$, $\lambda_4 = 0.19Hz$ and $\lambda_5 = 0.17Hz$ (hence, $\lambda_3 = 0.42Hz$ and $\lambda_6 = 0.36Hz$). Moreover, let $N = 8192$ as the data length.

1. Construct the $X[k]$.
2. Estimate the power spectrum $C_2^x(f)$. Use $L_2 = 128$ max shiftings for autocorrelation.
3. Estimate the bispectrum (only in the primary area) $C_3^x(f_1, f_2)$ using
   a) the indirect method with $K = 32$ and $M = 256$. Use $L_3 = 64$ max shiftings for the third-order cumulants. Use: a_1) rectangular window and a_2) Parzen window.
   b) the direct method with $K = 32$ and $M = 256$. Use $J = 0$.
4. Plot $X[k]$, $C_2^x(f)$, $C_3^x(f_1, f_2)$ (all estimations).

5. Compare the estimations of $C_3^x(f_1, f_2)$ amongst {a_1, a_2, b} settings. Comment on the comparisons.
6. What can you deduce regarding the frequency content from the comparison of $C_2^x(f)$ and $C_3^x(f_1, f_2)$ (all estimations)?
7. How the results will change if you repeat the process from 1 to 5 taking into account:
   a) different segment length: i) $K = 16$ and $M = 512$ ii) $K = 64$ and $M = 128$?
   b) 50 realizations of the $X[k]$ and comparing the mean values of the estimated $C_2^x(f)$, $C_3^x(f_1, f_2)$?

2 Solutions

The Matlab files for the following can be found here.

2.1 Time series

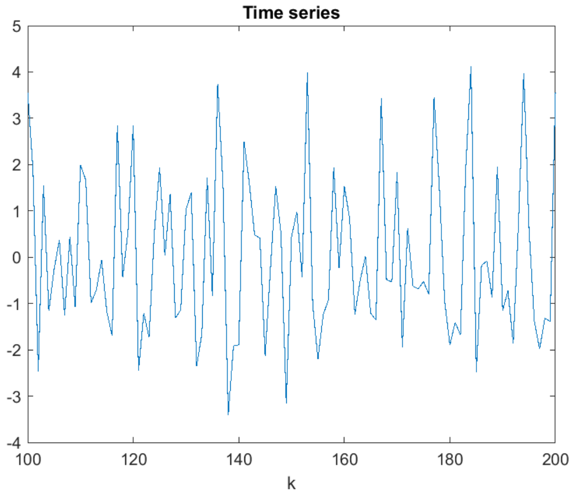


Figure 1: A portion of the time series (100:1:200 of total 0:1:8191)

2.2 Power Spectrum Density

About **Power Spectrum Density** (PSD) estimation, we first compute the autocorrelation function with 128 max shifts (thus range -128:128) and then its Fourier Transform with $2 * 128 + 1 = 257$ samples. As we can observe, the PSD can unfold the frequencies of the signal.

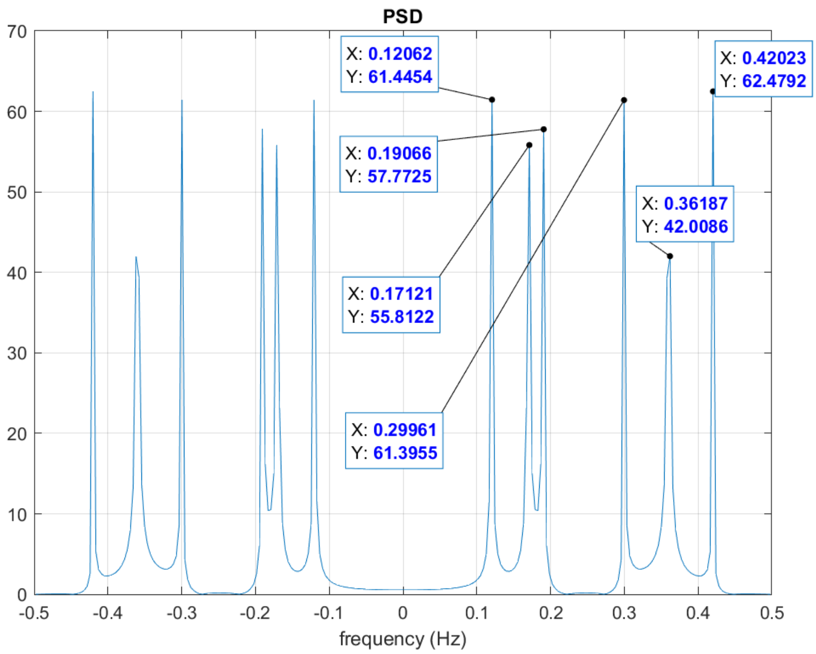


Figure 2: Power Spectrum Density

2.3 Bispectrum

About the following bispectrum estimations, it should be noted that a) for the Welch's method we used zero overlap between sample segments, b) regarding the plots, taking into account symmetries, we focused only in one quarter of the bispectrum domain including a red line for the $y = x$ symmetry and c) all the estimations are based on the HOSA toolbox.

2.3.1 Indirect L = 64, M = 256

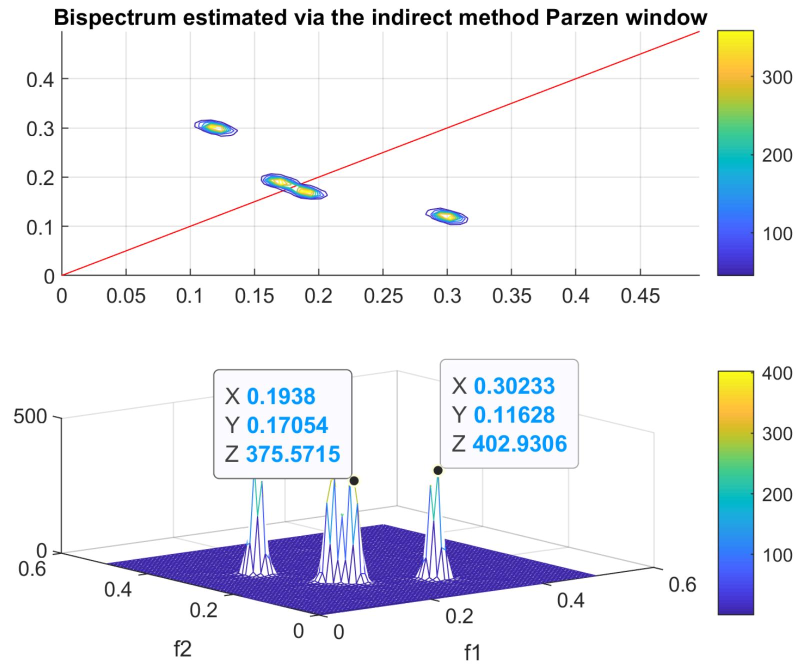


Figure 3: Bispectrum Indirect with Parzen window

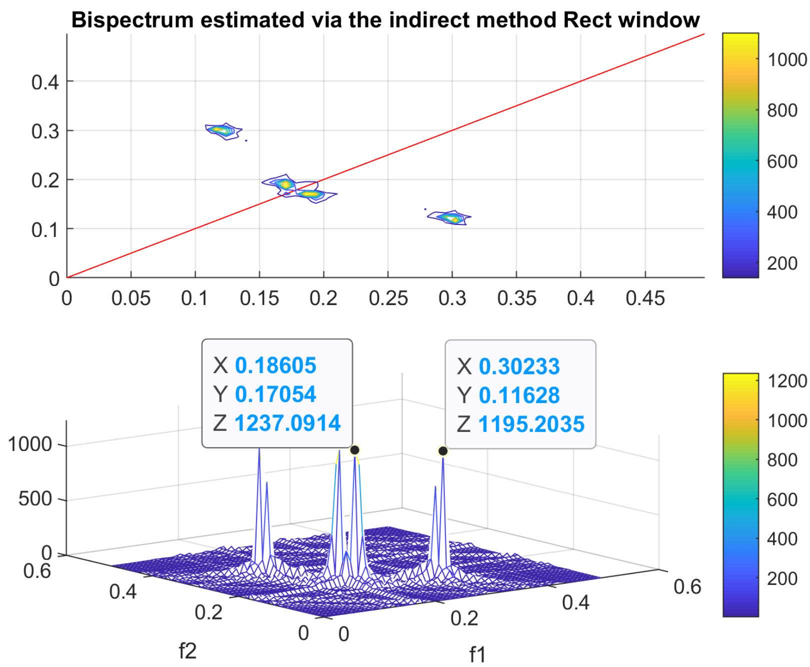


Figure 4: Bispectrum Indirect with Rectangular window

2.3.2 Direct M = 256

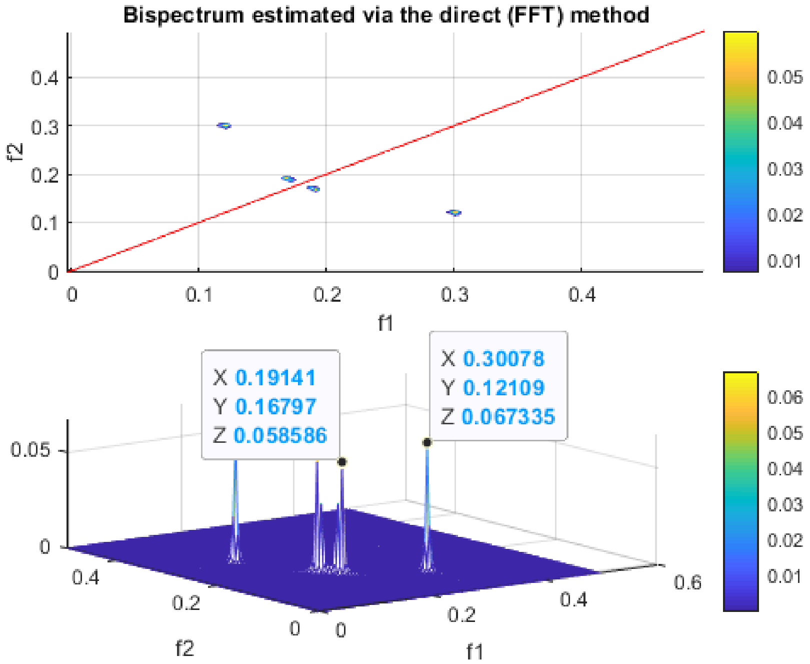


Figure 5: Bispectrum Direct without tapering

2.4 Conclusions

It should be noted that since we are using bispectrum estimations and not bicoherence, we are not interested particularly in the quantification of the coupling, so we will not make conclusions about the amplitude of the bispectrum rather than just the peaks observation.

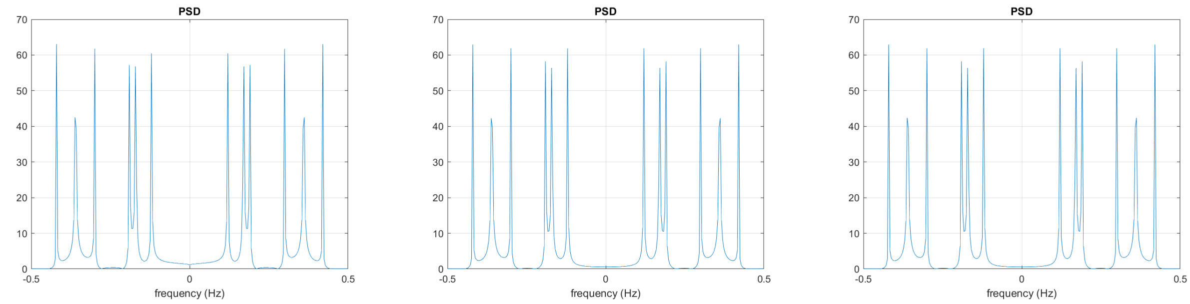
- Among indirect bispectrum estimation techniques, Parzen as a window function handles better **spectral leakage** rather than the Rectangular (no window) option and spectrum is smoother as it is expected. On the other hand since we don't use overlap between the segments, the tapering hides the edge information of the signal thus the estimation can't pinpoint the exact frequencies, creating a wider base of the peak.

- Direct estimation has better frequency resolution. This is reasonable because using 64 max lag in the indirect method leads to $2 * (64) + 1 = 127$ data points for the fast fourier transform since the transform is applied to the third order cumulant and not directly to the 256 samples as the Direct method. A fair comparison should probably be based on an equal number of samples. Testing locally the aforementioned comparison, increasing the max lag to 128 for the indirect, spectral leakage had been reduced but still the Direct method had sharper peaks.

- The (f1,f2) coordinates of the peaks from the bispectrum plots indicate the frequency components that are coupled to another component. First, inspecting the power spectrum, we can discover the frequency content of our signal. Then we can spot if some of them can be expressed as the sum of two other frequencies. Indeed $0.19 + 0.17 = 0.36 = f_3$ and $0.3 + 0.12 = 0.42 = f_6$. So for the f_3 and f_6 components to be quadratic coupled, meaning that $phase_3 = phase_2 + phase_1$ and $phase_6 = phase_5 + phase_4$, we should observe a sharp peak in the amplitude of the bispectrum. Indeed we can validate this based on the original signal and the plots. Thus we can conclude that bispectrum estimation can unfold phase coupling that is a phase information that power spectrum hides.

2.5 Repeating the process

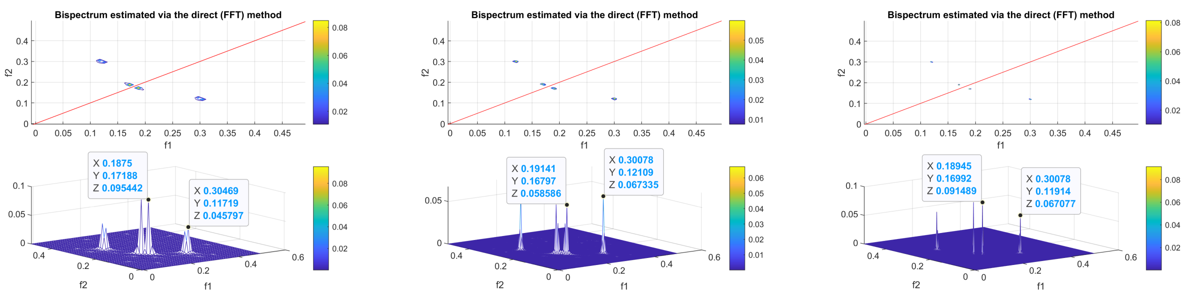
2.5.1 Power Spectrum Density varying M and L = 128



(a) M = 128 (b) M = 256 (c) M = 512

Figure 6: Power Spectrum Density estimations for different length (M) of the sample segments

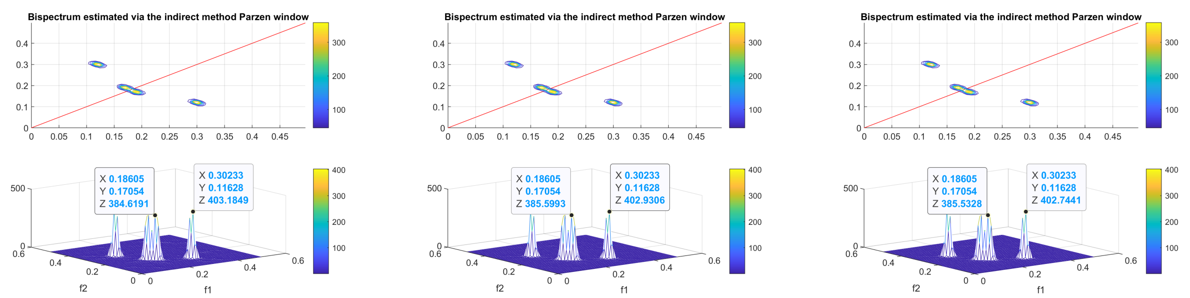
2.5.2 Bispectrum Direct varying M



(a) M = 128 (b) M = 256 (c) M = 512

Figure 7: Bispectrum estimations using Direct method for different length (M) of the sample segments

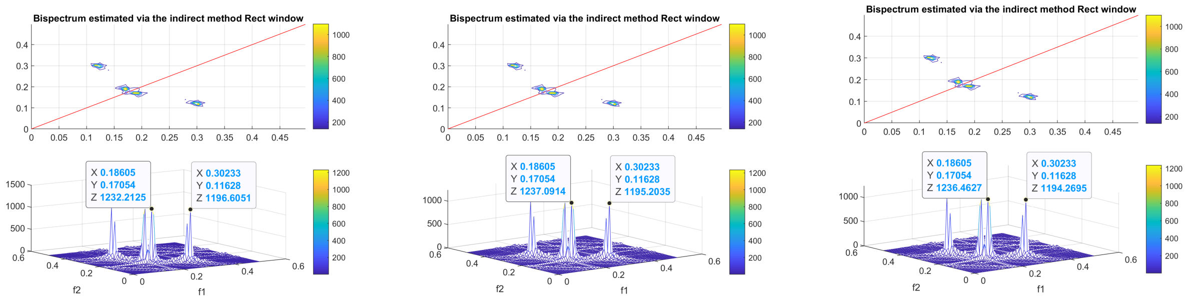
2.5.3 Bispectrum Indirect Parzen varying M and L = 64



(a) M = 128 (b) M = 256 (c) M = 512

Figure 8: Bispectrum estimations using Indirect-Parzen for different length (M) of the sample segments

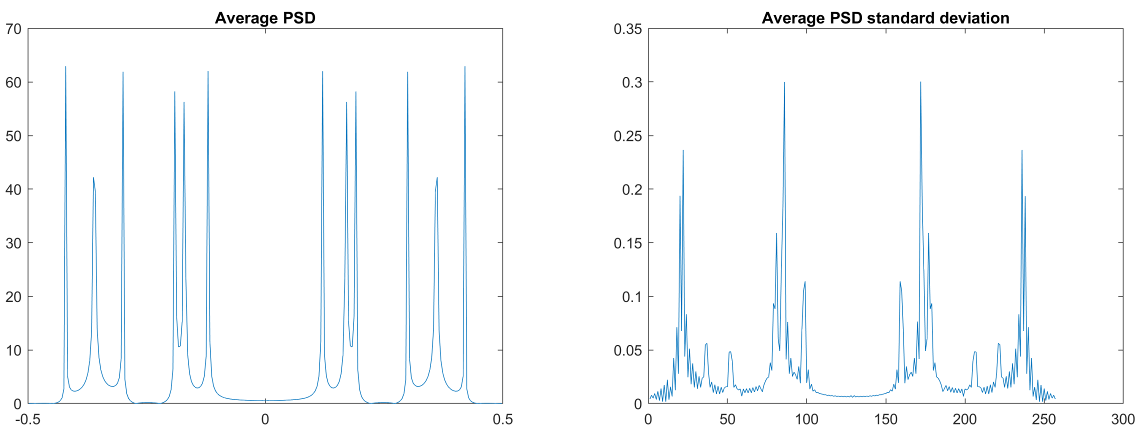
2.5.4 Bispectrum Indirect Rectangular varying M and L = 64



(a) M = 128 (b) M = 256 (c) M = 512

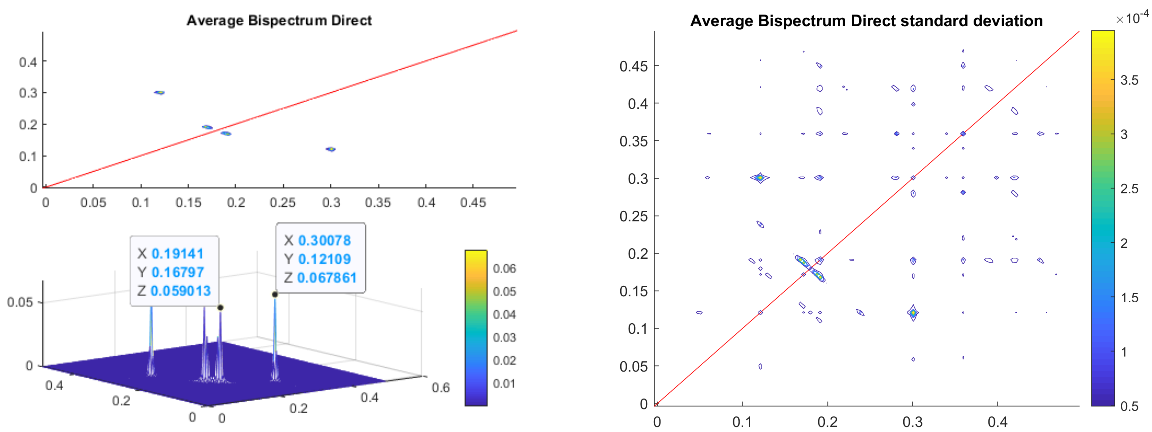
Figure 9: Bispectrum estimations using Indirect-Rectangular for different length (M) of the sample segments

2.5.5 Averaging estimations for 50 realizations



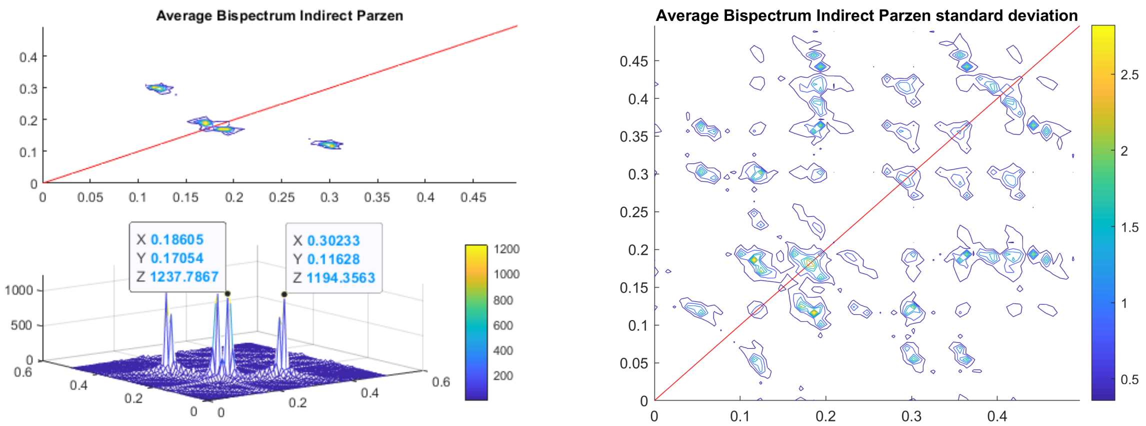
(a) M = 0, L = 128 (b) Standard Deviation

Figure 10: Mean of Power Spectrum Density for 50 realizations



(a) M = 256 (b) Standard Deviation

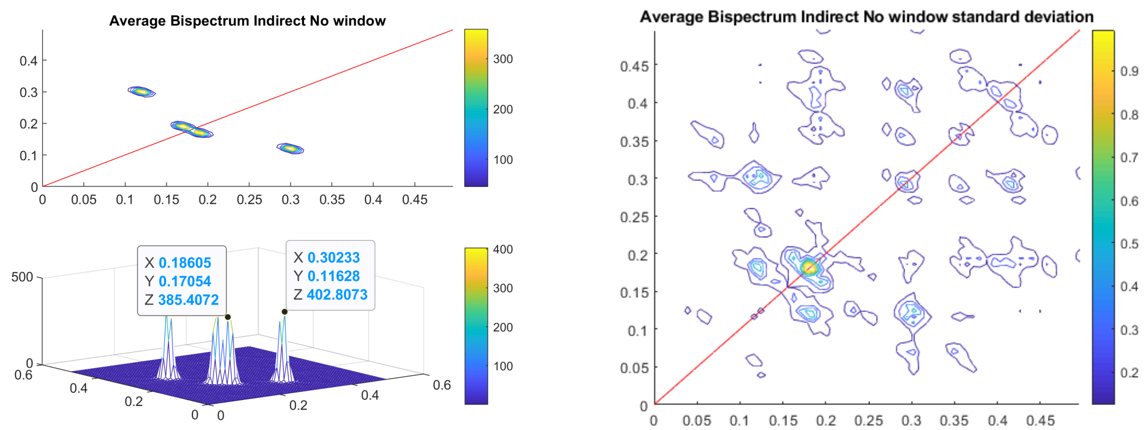
Figure 11: Mean of Bispectrum using Direct method for 50 realizations



(a) M = 256, L = 64

(b) Standard Deviation

Figure 12: Mean of Bispectrum using Indirect-Parzen window for 50 realizations



(a) M = 256, L = 64

(b) Standard Deviation

Figure 13: Mean of Bispectrum using Indirect-Rectangular for 50 realizations

2.5.6 Conclusions

- Varying the length (M) of the sample segments, we couldn't spot any significant difference for the estimations except the Bispectrum using the Direct method. The number of fft samples for this one is based on the length of the segments, so the frequency resolution increased, thus the peaks become sharper. On the other hand, since we kept the max lag constant, indirect estimations had the same resolution and changing the size of the slices didn't change the outcome.

- Averaging the estimations, we can conclude that they have small variance.