University of Padua
Degree Course in
ICT for Internet and Multimedia
Academic Year 2017/2018

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

1^{st} Homework

SANTI GIOVANNI - 1177739

VANIN EDOARDO - 1179018

PROBLEM 1

Given a single realization of 800 i.i.d. samples of the r.p $\{x(k)\}$ defined by:

$$x(k) = e^{j(2\pi f_1 k + \varphi_1)} + 0.8 e^{j(2\pi f_2 k + \varphi_2)} + w(k)$$

where $w(k) \sim \mathcal{CN}(0, \sigma_w^2)$, $f_1 = 0.17$, $f_2 = 0.78$, $\sigma_w^2 = 2.0$ $\varphi_{1,2} \sim \mathcal{U}(0, 2\pi)$ are statistically independent, we want to estimate the $\mathcal{PSD}$ of $\{x(k)\}$ using the following models.
The signal was generated using the *in-phase* and *in-quadrature components* notation:

$$x(k) = x_I(k) + j x_Q(k)$$

where respectively

$$x_I(k) = \cos(2\pi f_1 k \varphi_1) + 0.8 \cos(2\pi f_2 k \varphi_2) + w(k)$$
$$x_Q(k) = \sin(2\pi f_1 k \varphi_1) + 0.8 \sin(2\pi f_2 k \varphi_2) + w(k)$$

BLACKMAN AND TUKEY CORRELOGRAM

$$\mathcal{P}(F) = T_c \sum_{n=-L}^{L} w(n) \hat{r}_x(n) e^{-j2\pi f n T_c} \tag{1}$$

where $\hat{r}_x(n)$ is the autocorrelation of x taken from -L to L and $w(n)$ is a Hamming window of length 2L+1. The autocorrelation function is evaluated using the *unbiased estimate* give by:

$$\hat{r}_x(n) = \frac{1}{K-n} \sum_{k=n}^{K-1} x(k) x^*(k-n) \qquad n = 0, 1, ..., K-1$$

where K is the number of samples. In order to reduce the variance of the estimate, we chose $L = \lfloor \frac{K}{5} \rfloor$. The hamming window is defined by:

$$w(k) = \begin{cases} 0.54 + 0.46 \cos\left(2\pi \frac{k - \frac{L-1}{2}}{L-1}\right) & k = 0, 1, ..., L-1 \\ 0 & elsewhere \end{cases} \tag{2}$$

PERIODOGRAM

$$\mathcal{P}_{PER}(f) = \frac{1}{KT_c}|\tilde{\mathcal{X}}(f)| \tag{3}$$

where $\tilde{\mathcal{X}}(f) = T_c\mathcal{X}(f)$, $\mathcal{X}(f)$ is the Fourier Transform of $\{x(k)\}$, $k = 0, 1, ..., K-1$.

WELCH PERIODOGRAM

$$\mathcal{P}_{WE}(F) = \frac{1}{N_s}\sum_{x=0}^{N_s-1}\mathcal{P}^{(s)}_{PER}(f) \tag{4}$$

where $\mathcal{P}^{(s)}_{PER}(f)$ is the periodogram estimate given by Eq.3. The model extracts different subsequences of consecutive D samples which eventually overlap. If $x^{(s)}$ is the s-th subsequence, characterized by S samples in common with the previous subsequence $x^{(s-1)}$ and with the following $x^{(s+1)}$, the total number of subsequences is $N_s = \lfloor \frac{K-D}{D-S} - 1 \rfloor$=160 since in our analysis, we chose S=100 and D=200.
More in details, we computed:

$$\begin{aligned}
x^{(s)}(k) &= w(k)x(k + s(D_S)) \qquad k = 0, 1, ..., D-1 \quad s = 0, 1, ..., N_s - 1 \\
\tilde{\mathcal{X}}^{(s)}(f) &= \mathcal{F}[x^{(s)}(k)] = T_c\sum_{k=0}^{D-1} x^{(s)}(k)e^{-j2\pi fkT_c} \\
\mathcal{P}^{(s)}_{PER}(f) &= \frac{1}{DT_cM_w}\left|\tilde{\mathcal{X}}^{(s)}(f)\right|^2
\end{aligned}$$

where $w(k)$ is the hamming window difined by Eq. 2 and $M_w = \frac{1}{D}\sum_{k=0}^{D-1} w^2(k)$ is the normalized energy of the window.

AR(N) MODEL

$$\mathcal{P}_{AR}(f) = \frac{T_c\sigma_w^2}{|\mathcal{A}(f)|^2} \tag{5}$$

The output process is described by the following recursive e equation

$$x(k) = -\sum_{n=1}^{N} a_n x(k-n) + w(k) \tag{6}$$

where w is white noise with variance σ_w^2. The transfer function of this system is given by $H_{AR} = \frac{1}{A(z)}$ where $A(z) = 1 + \sum_{n=1}^{N} a_n z^{-n}$. The ACS of x in z-domain is easily evaluated using

$$P_x(z) = \frac{\sigma_w^2}{A(z)A^*(\frac{1}{z^*})}$$

from which Eq. 5 is computed letting $\mathcal{A}(f) = A(e^{j2\pi fT_c})$.
Coefficients $a_1, a_2, ..., a_N$ are computed exploiting the *Yule-Walker equation* $\mathbf{a} = -\mathbf{R}^{-1}\mathbf{r}$, where $r = [r(1), r(2), \ldots, r(N)]$ and the *autocorrelation matrix* is defined as

$$\mathbf{R} = \begin{pmatrix} r(0) & r(-1) & \ldots & r(_N{+}1) \\ r(1) & r(0) & \ldots & r(-N+2) \\ \vdots & \vdots & \ddots & \vdots \\ r(N-1) & r(N-2) & \ldots & r(0) \end{pmatrix}$$

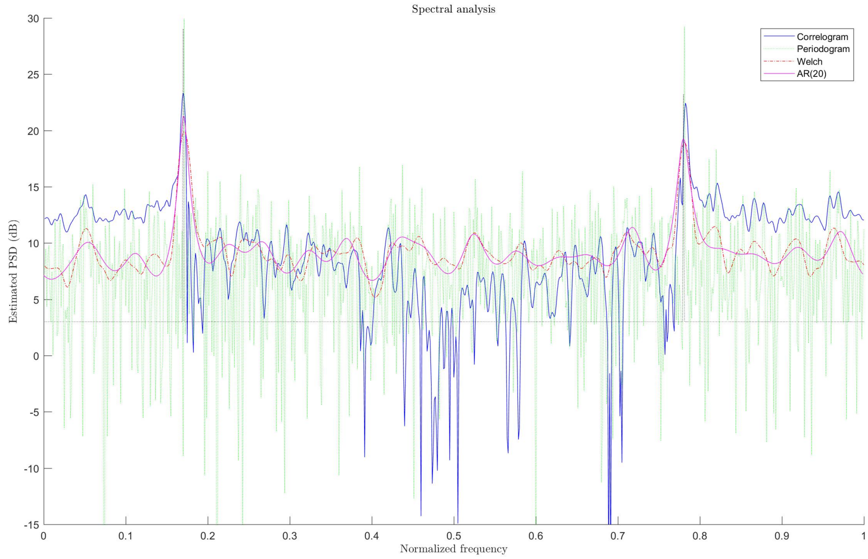

Figure 1: PSD estimates.

The noise variance deriving from the model is

$$\sigma_w^2 = r_x(0) + \mathbf{r}^H \mathbf{a} \tag{7}$$

CONCLUSIONS

A comparison between different PSD estimates of $\{x(k)\}$ is now discussed. The difficulties in this problem were related to the fact that the noise variance is double with respect to the maximum amplitude of the useful signal. For this reason, a good analysis can be achieved only by considering a generated sequence $\{x(k)\}$ which is not too corrupted at the carrier frequencies f_1 and f_2. The four different PSD estimates are shown in Figure [1].

The parameters of the different models are as follow. The order of the autocorrelation estimator for the Correlogram is $L = \lfloor K/5 \rfloor = 160$, and the window used is a Hamming window definde in Eq. 2. For the Periodogram the window used is the rectangular one, while for the Welch Periodogram we set $D = 200$ and $S = 100$ still using an Hamming window in order to emphasize the discrete components. This choise was made by observing different PSD estimates with varying values of S and D, as shown in Figure [2]. The order evaluation for the AR model was carried out by analysing the behaviour of the noise variance σ_w^2 with varying N defined in eq. 7 and graphically shown in Fig. [3].
It is easy to see that no significant keen occurs neither for high value of N, thus for this problem there is not a clean value that should be chosen. On the other hand, a large N may result in an ill-conditioned autocorrelation matrix $\mathbf{R}$, this because of the presence of spectral lines in the original process. For all these reasons, we decided that N=25 was a good trade-off.

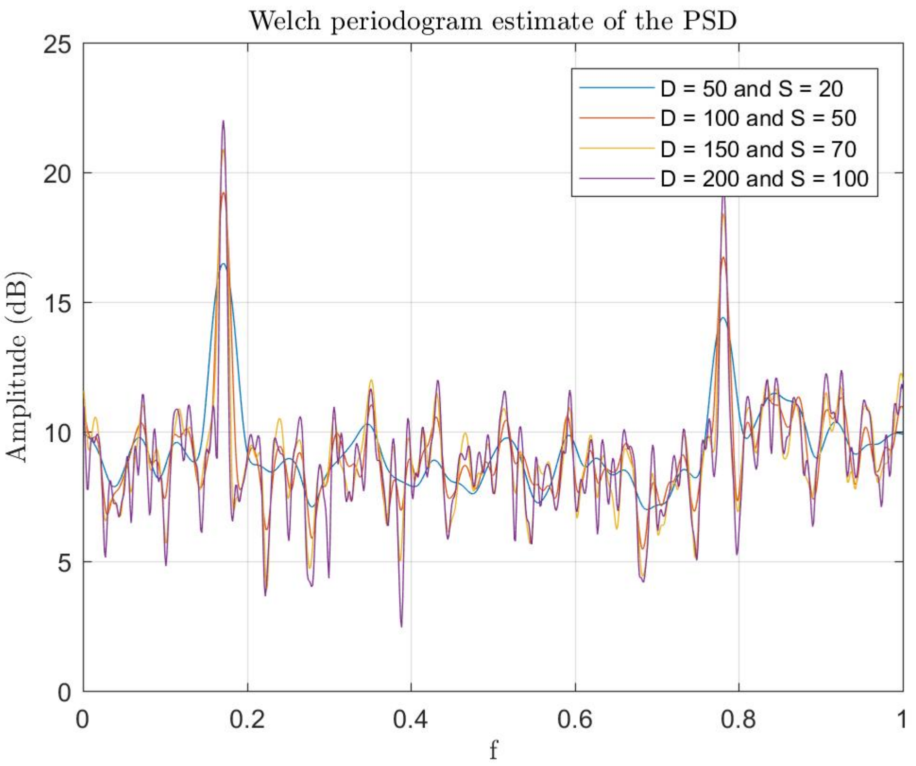

Figure 2: Welch periodogram estimates as a function of different values of D and S.

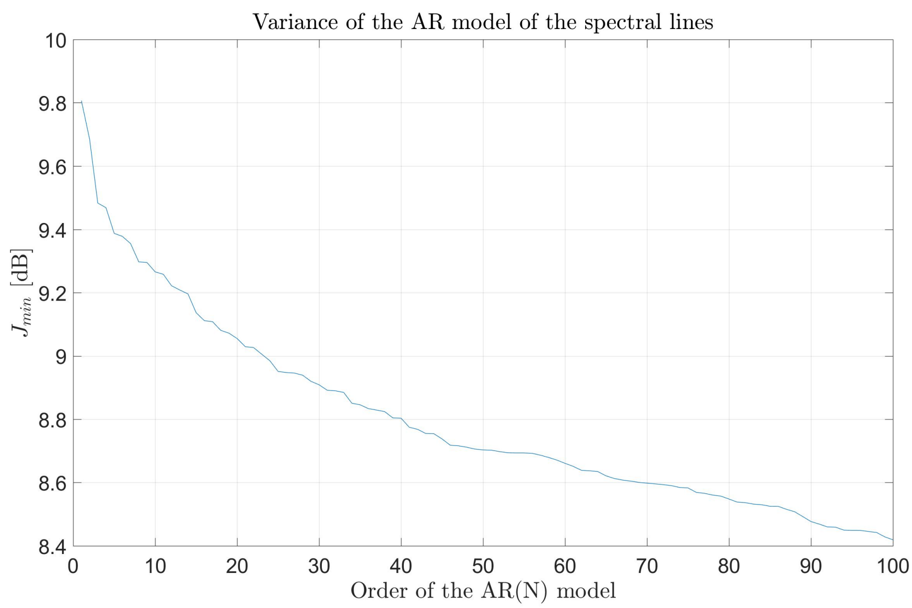

Figure 3: Variance of the AR model white noise as a function of the order N.

PROBLEM 2

The only different with Problem 1 is the noise variance, which is $\sigma_w = 0.1$. Now the noise component is almost negligible with respect the usful signal, and then we were able to provide a much more accurate analysis. The different spectral estimates are given in Figure [4].

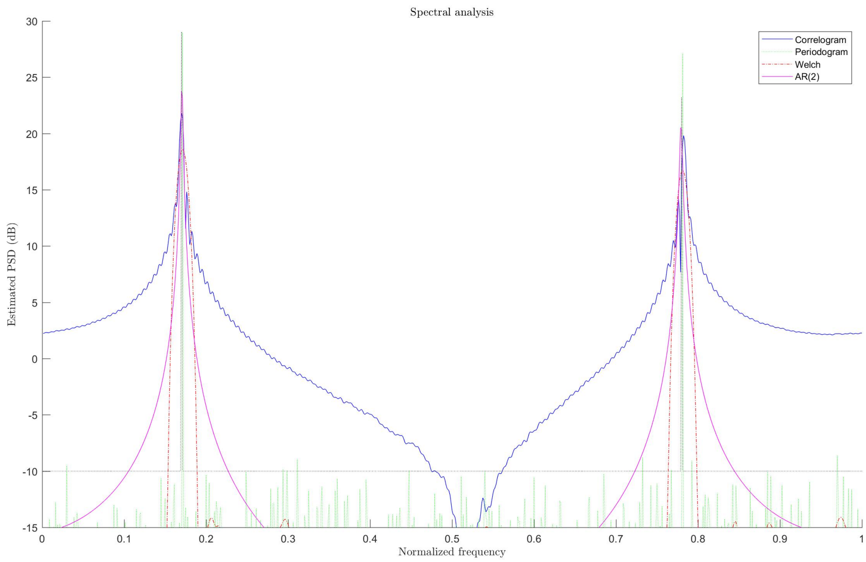

Figure 4: PSD estimates.

The parameters of the model are as follow. The order of the autocorrelation estimator for the Correlogram is $L = \lfloor K/5 \rfloor = 160$, and the window used is a Bartlet window definde as

$$w(k) = \begin{cases} 1 - 2\left|\frac{k - \frac{L-1}{2}}{L-1}\right| & k = 0, 1, ..., L-1 \\ 0 & elsewhere \end{cases} \tag{8}$$

This choise was taken by comparing the PSD estimates obtained with different windows, respectively: *Hamming*, *Rectangular* and *Bartlett*. As shown in Figure [5], the last one has a much more smooth behaviour.

For the Periodogram the window used is the rectangular one, while for the Welch Periodogram we set $D = 400$ and $S = 200$ using an Bartlett window. Again, this choise was taken by comparing the PSD estimates with varying values of D and S, given in Figure [6]. Note that for higher values of S and D the Welch estimate tend to approximate with high confidence the analytical PSD, and at the same time it tends to follow the Periodogram estimate. This is a consequence of Welch model, since by considering longer subsequences the total number of this decrease and then the average is computed taking less Periodograms. Finally, the order evaluation for the AR model was carried out by analysing the behaviour of the noise variance σ_w^2 with varying N defined in eq. 7 and graphically shown in Fig. [7]. Here it can be observed a knee in the function, which corresponds to a value of N=2. Note that also N=5 can be a good choise, this because it yelds to a lower noise variance and at the same time the AR model approximates in a better way the analytical PSD, having more

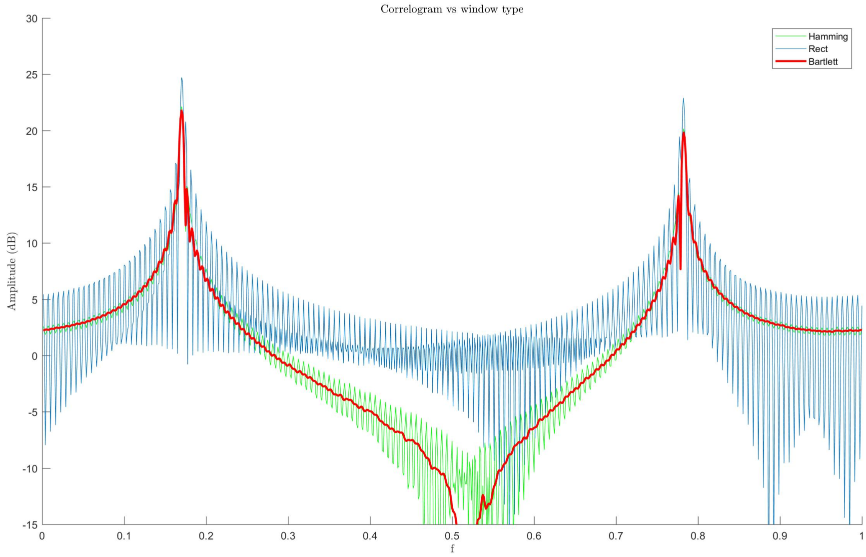

Figure 5: PSD estimates.

degrees of freedom. However N=2 is sufficent to describe a model with two discrete components and also it is more feasible in the remaining problems.

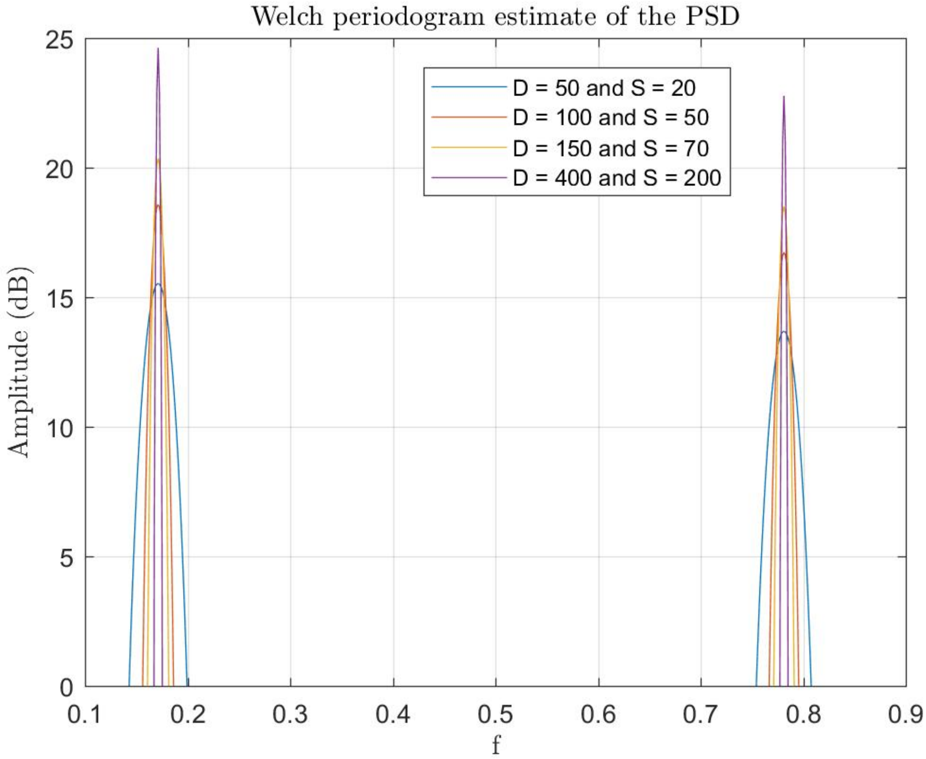

Figure 6: Welch periodogram estimates as a function of different values of D and S.

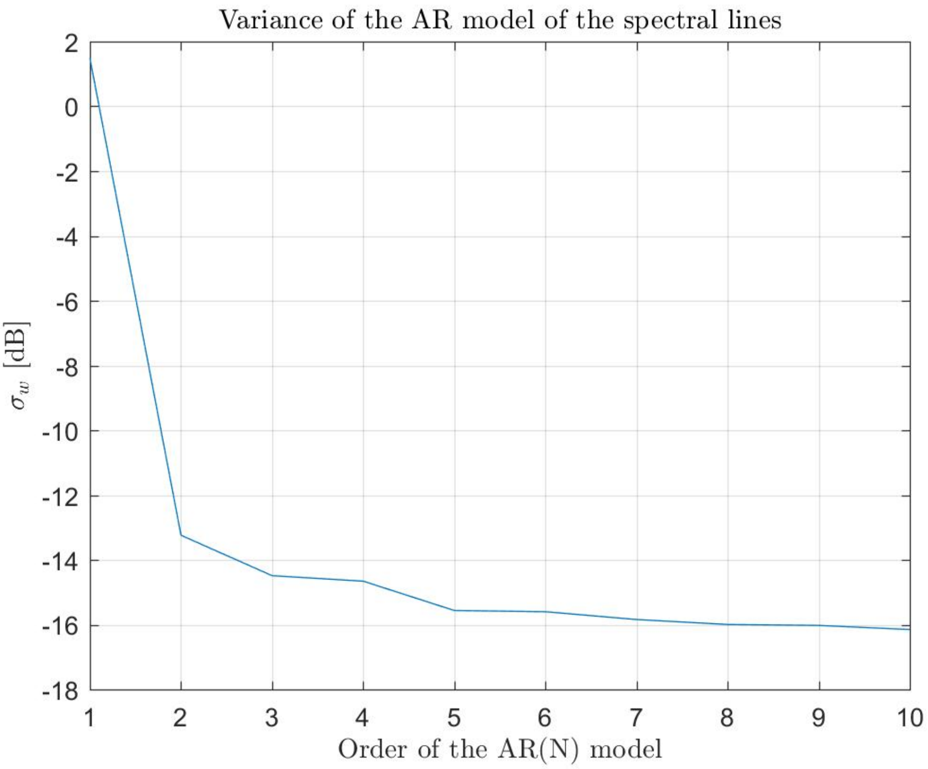

Figure 7: Variance of the AR model white noise as a function of the order N.

PROBLEM 3

Given a discrete-time WSS random process *x(k)* with zero mean at time *k-1*, i.e. $\mathbf{x}^T(k-1) = [x(k-1), x(k-2), \ldots, x(k-N)]$, the *one-step forward predictor of order* N attempts to estimate the value of *x(k)* using $\mathbf{x}^T(k-1)$. This problem can be solved by considering the signal $\mathbf{x}^T(k-1)$ as the input of a Wiener-Hopf filter of order N and *x(k)* as the reference signal. The solution is computed by the Wiener-Hopf equation

$$\mathbf{R}_N \mathbf{c}_{opt} = \mathbf{r}_N$$

It is important to underline that this solution holds if and only if the autocorrelation matrix $\mathbf{R}$ is invertible, i.e det$\mathbf{R} \neq 0$. In this case the coefficent vector is optimum and it can be computed by

$$\mathbf{c}_{opt} = \mathbf{R}_N^{-1} \mathbf{r}_N$$

Moreover, the minimum value of the cost function J is

$$J_{min} = r_x(0) - \mathbf{r}_N^H \mathbf{c}_{opt}$$

Knowing that, given an AR process x of order N, the optimum prediction coefficients $\mathbf{c}_{opt}$ coincide with the parameter $-\mathbf{a}$ of the process and, moreover, that $J_{min} = \sigma_w^2$, we can use the results of Problem 2 and get respectively:

- N=2;
- $\mathbf{c}_{opt} = -\mathbf{a} = [+0,657 - 0,101i, -0,935 + 0,305i]$;
- $J_{min} = \sigma_w^2 = r_x(0) + \mathbf{r}^H \mathbf{a}$.

The transfer function of the model is $H_{AR} = \frac{1}{A(z)}$ where $A(z) = 1 + \sum_{n=1}^{2} a_n z^{-n}$ and the noise variance is $\sigma_w^2 = -13.2397dB$. Poles and zeros of *A(z)* are shown in Figure [8].

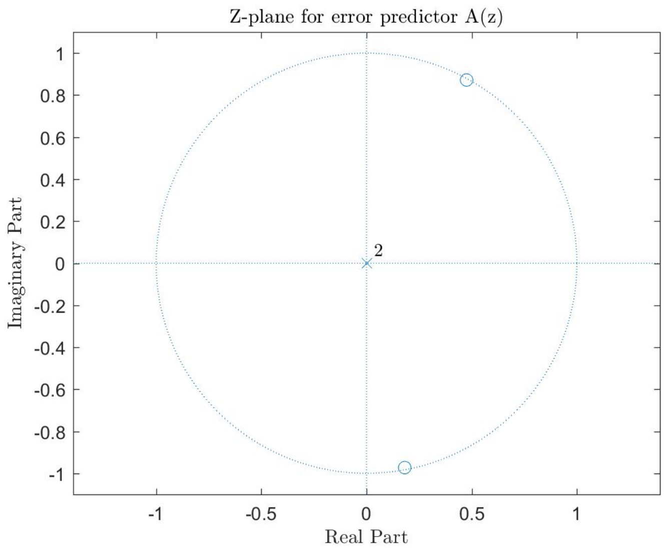

Figure 8: Zeros of A(z) in the complex plane.

Notice that zeros of *A(z)* are poles of H_{AR}. For this reason, a zero near the unit circle of *A(z)* corresponds to a pole of H_{AR} with magnitude almost equal to 1, which brings $|H_{AR}| \to \infty$. This can be easly observed in the frequensy domain, in which a the condition of a pole with magnitude almost equal to 1 coincides with a peak centered at frequency $\frac{\angle(\Theta)}{2\pi}$, where the pole is $p = \rho e^{j\Theta}$, $\rho \approx 1$. The same reasoning is applied to the PSD estimate given by the AR model of equation 5, therefore knowing the coefficient vector **a** we can compute the frequensies of the 'discrete components'. The values we obtained are the following:

$$z_1 = \rho_1 e^{j\Theta_1} = 0.9932 e^{+1,0714j} \qquad f_1 = \frac{\Theta_1}{2\pi} = 0.1705$$

$$z_2 = \rho_2 e^{j\Theta_2} = 0.9897 e^{-1,3864j} \qquad f_2 = \frac{\Theta_2}{2\pi} = 0.7793$$

It is easy to see that the carrier frequencies evaluated using the AR model looks very similar to the original ones used to create the $\{x\,(k)\}$ process

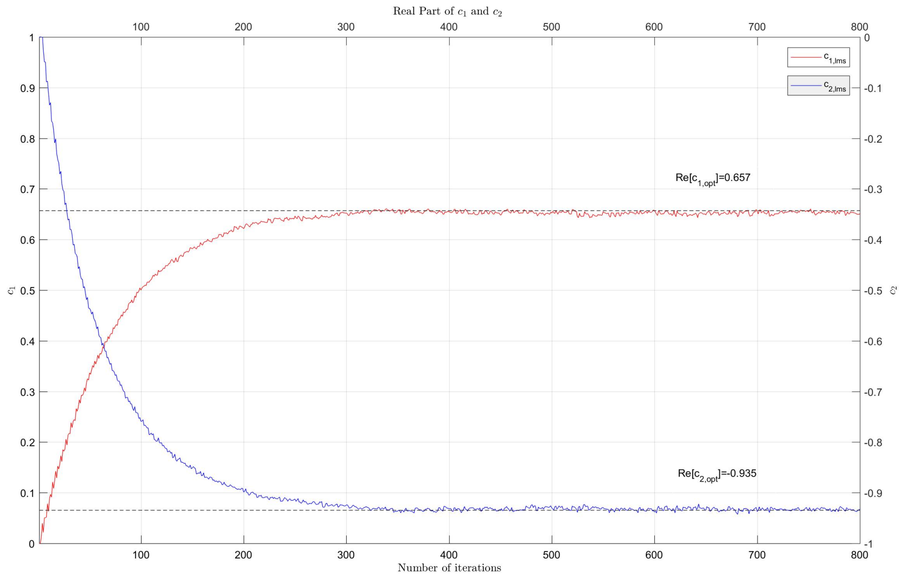

Figure 9

PROBLEM 4

The setup for the *least mean-square* (LMS) is the one described in Problem 3. It provides a recursive method to compute an approximation of the optimal Wiener-Hopf solution for the predictor problem without using the autocorrelation matrix **R** and the vector **r**, which are instead necessary in order to compute the *optima* coefficients.
The equation for the adaptation of the filter coefficients is

$$\mathbf{c}(k+1) = \mathbf{c}(k) + \mu e(k)\mathbf{x}^*(k) \tag{9}$$

where k is the considered instant. The two parameters that must be determined are the filter order N and the update coefficient $\mu = \frac{\tilde{\mu}}{N r_x(0)}$, for which $\tilde{\mu}$ should be in the range $(0, 2)$ to garantee convergency.
Since the sequence $\{x(k)\}$ is the same as Problem 3, we chose N=2 and starting from the all-zero initial condition and $\tilde{\mu} = 0.06$ we run the algorithm:

1. given $[x(k), x(k-1), x(k-2), \ldots, x(k-N+1)]^T$, $x(k) = 0 \quad \forall k < 0$, we computed the predictor value
$$y(k) = \sum_{n=1}^{N} c_n x(k-n)$$
2. we evaluated the error $f_N(k) = d(k) - y(k) = x(k+1) - y(k)$;
3. we updated the coefficients using Equation 9.

The convergence curves for $\Re[c_i]$ and $\Im[c_i]$, $i = 1, 2$, are given in Figures 9 and 10 both showing also the theoretical value of convergence. In Figure 11 is also given the cost function of the Wiener filter approximated by $|e(k)|^2$ in the LMS algorithm.
As expected, both the coefficients vector **c** and the squared-error $|e(k)|^2$ converge to $-\mathbf{a}$ and σ_w^2 respectively. This is true not only for a single realization of $\{x(k)\}$

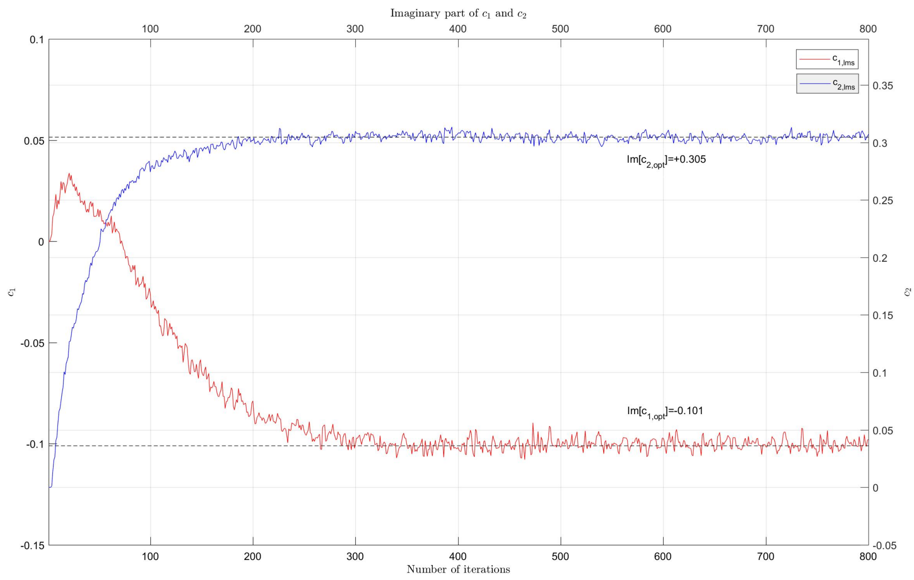

Figure 10

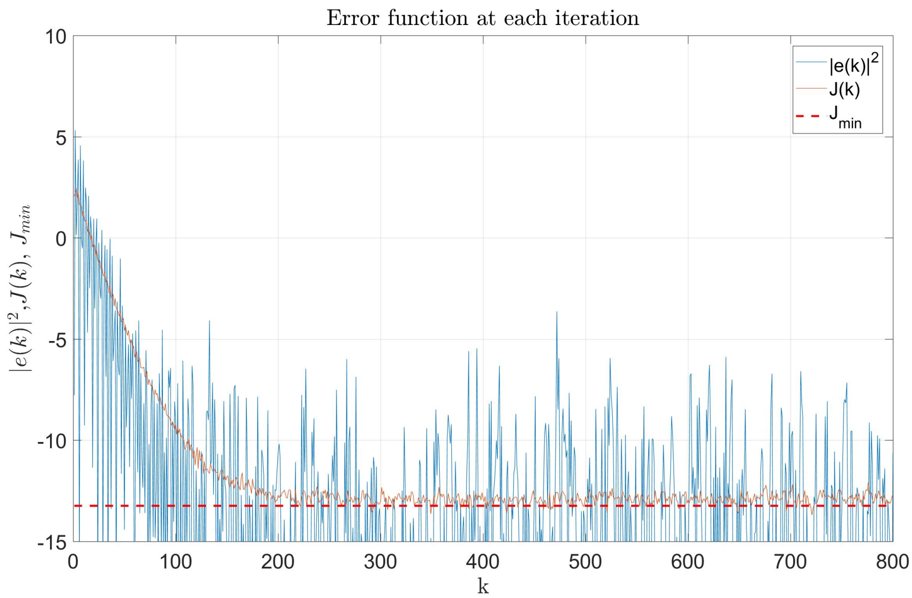

Figure 11

but also in mean, assuming the process to be ergotic. This can be easily seen in Figure 11, where the orange curve represents the averege error evaluated over 300 realizations. Note that here the convergence is achieved after almost 300 iterations both for the real and imaginary part of c_1 and c_2, but this is related to our choise of $\tilde{\mu} = 0.06$. We observed that this was a good trade-off between fast-convergence and grassy behaviour of the coefficients around the optimum values, since higher values ensured faster convergency but also higher fluctuations.
The values we obtained are the following:

$\mathbf{c_1(600)}$	$\mathbf{c_2(600)}$]	**mean over last 100 values**	
0.6579 - i0.0968	-0.9393 + i0.3075	$\bar{c_1}$=0.6535 - i0.1004	$\bar{c_2}$=-0.9330 + i0.3051