

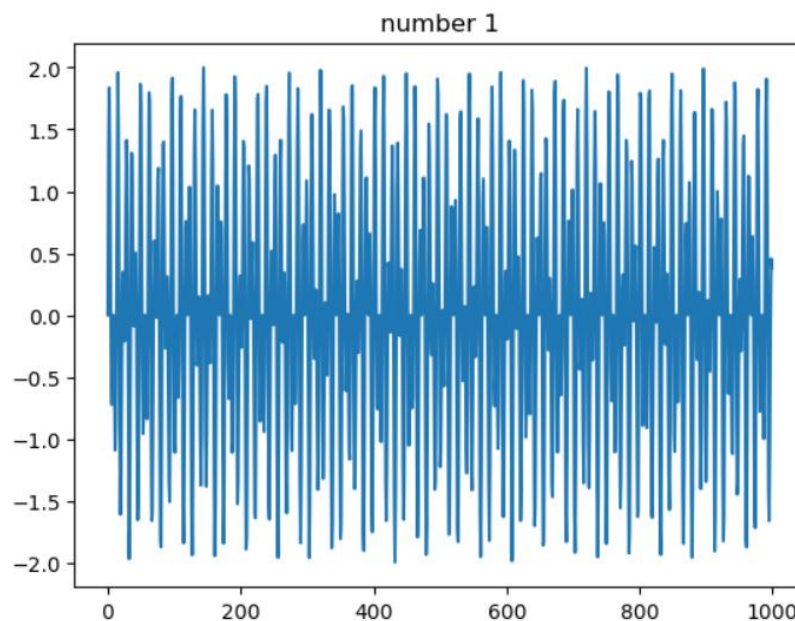
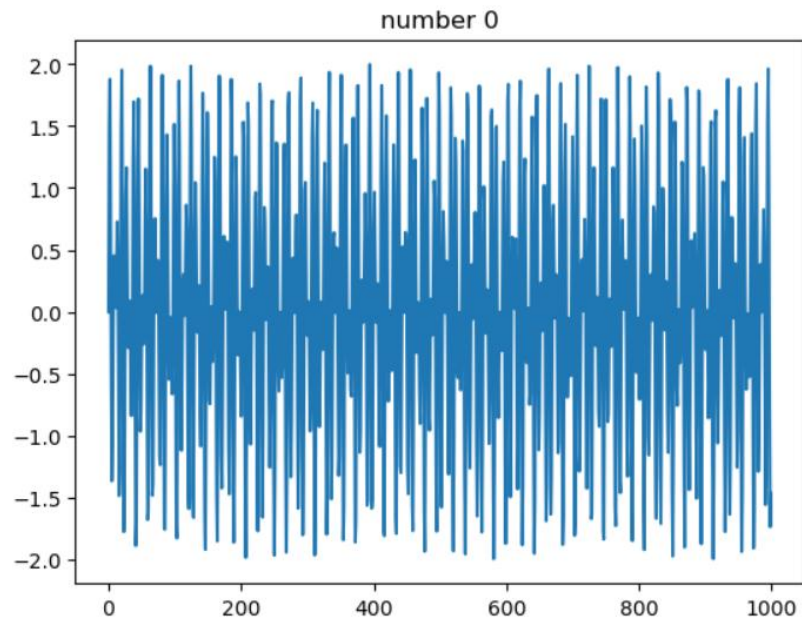
Ψηφιακή Εργασία Σήματος

1^η Εργαστηριακή Άσκηση

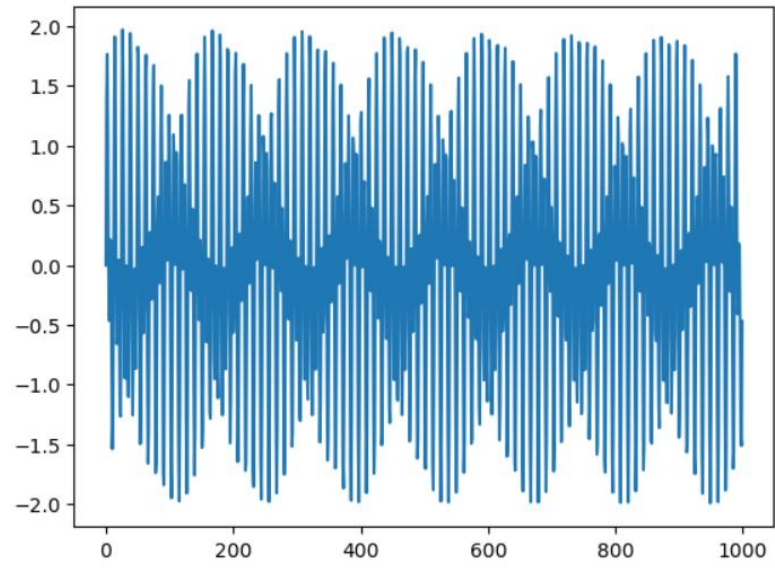
Άννα Κουτσώνη-03120019

Μέρος 1^ο – Σύστημα Εντοπισμού Τηλεφωνικών Τόνων (Telephone Touch Tones)

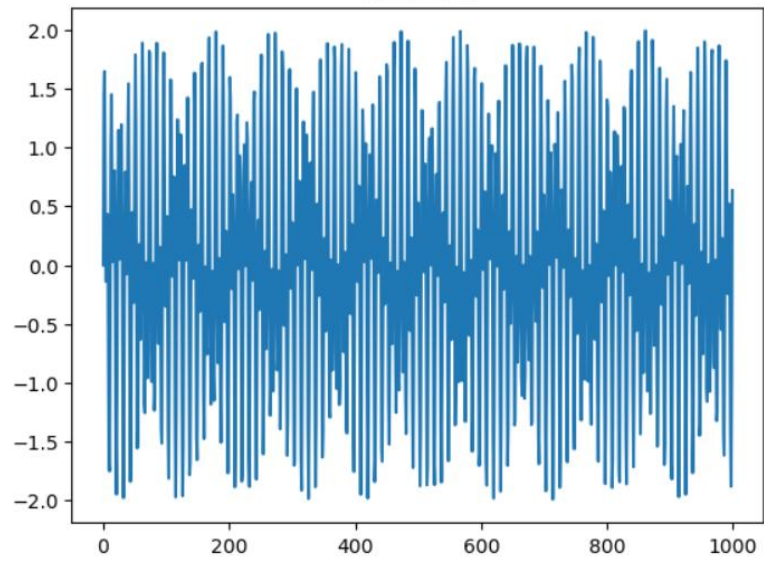
- 1.1 Δημιουργούμε 10 διαφορετικούς τόνους σύμφωνα με τον πίνακα συχνοτήτων που δίνεται. Κάθε τόνος έχει μήκος $L=1000$ και είναι άθροισμα 2 ημιτόνων, το υψίσυχο ημίτονο (στήλη πίνακα) και το χαμηλόσυχο ημίτονο (γραμμή πίνακα). Με την εντολή `IPython.display.Audio()` και συχνότητα δειγματοληψίας 8.192 kHz ακούμε τον τόνο d1 και διαπιστώνουμε την ορθή λειτουργία του. Οι γραφικές παραστάσεις των 10 τόνων:



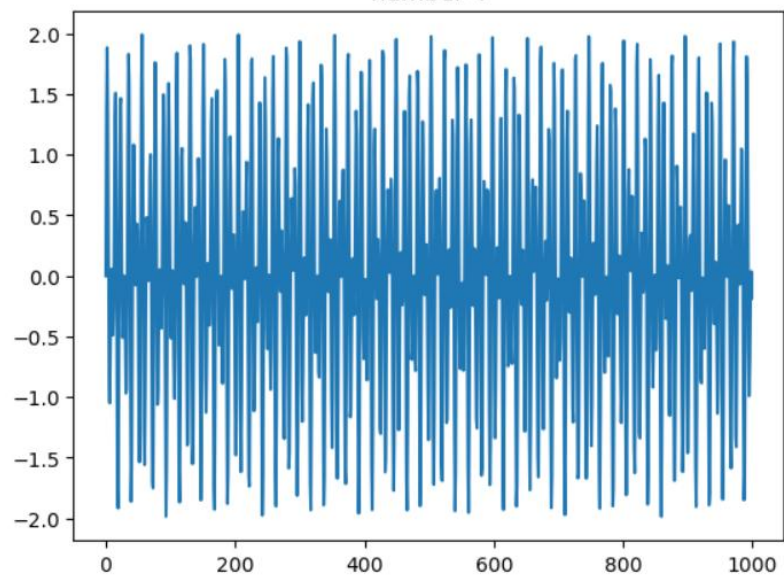
number 2



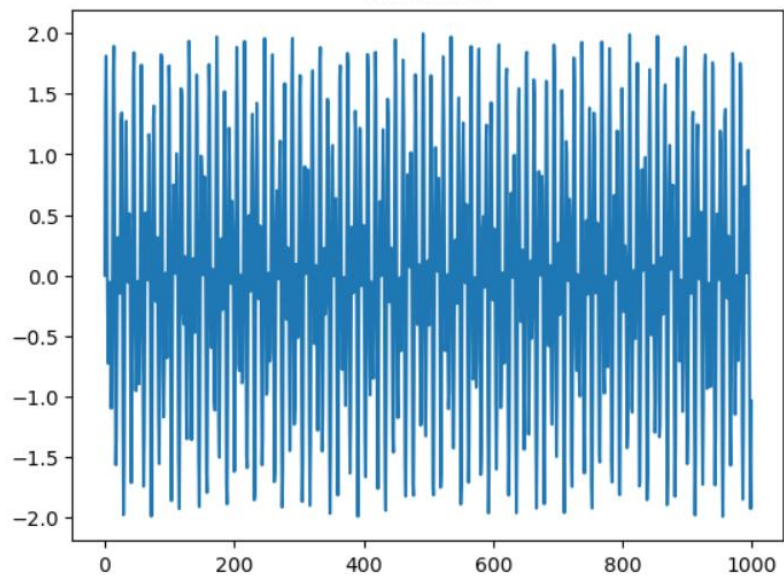
number 3



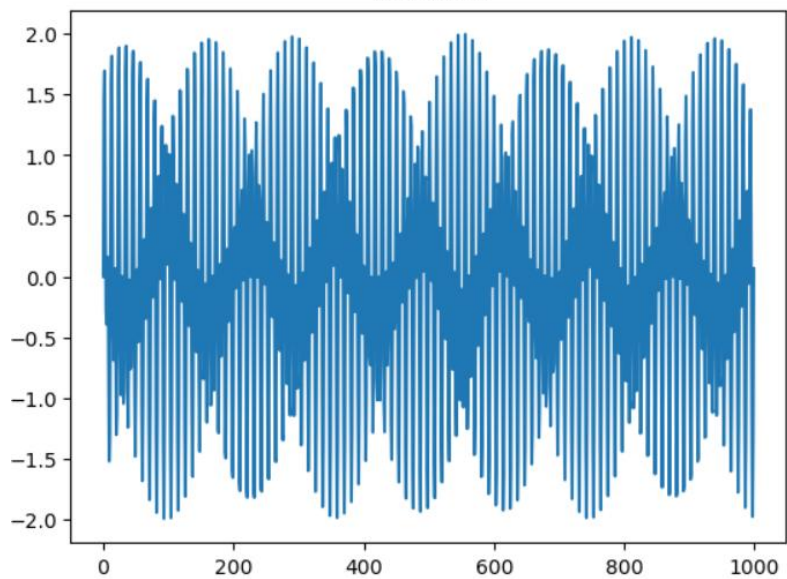
number 4



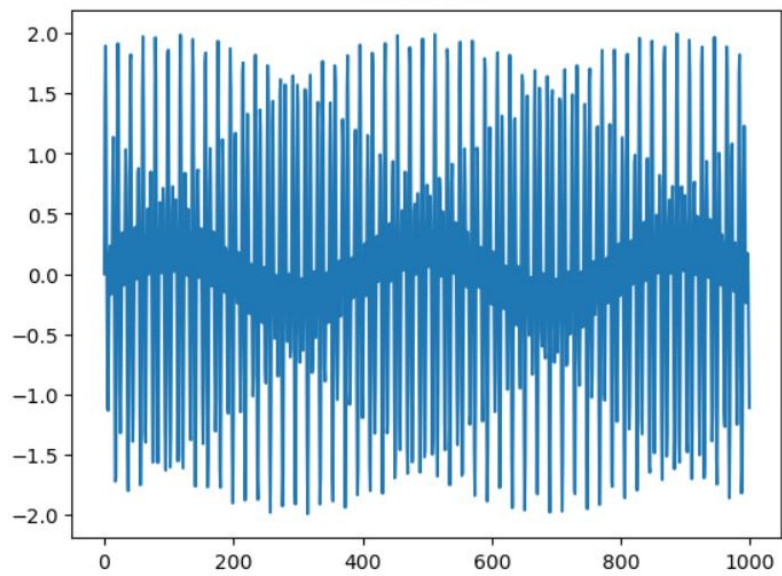
number 5



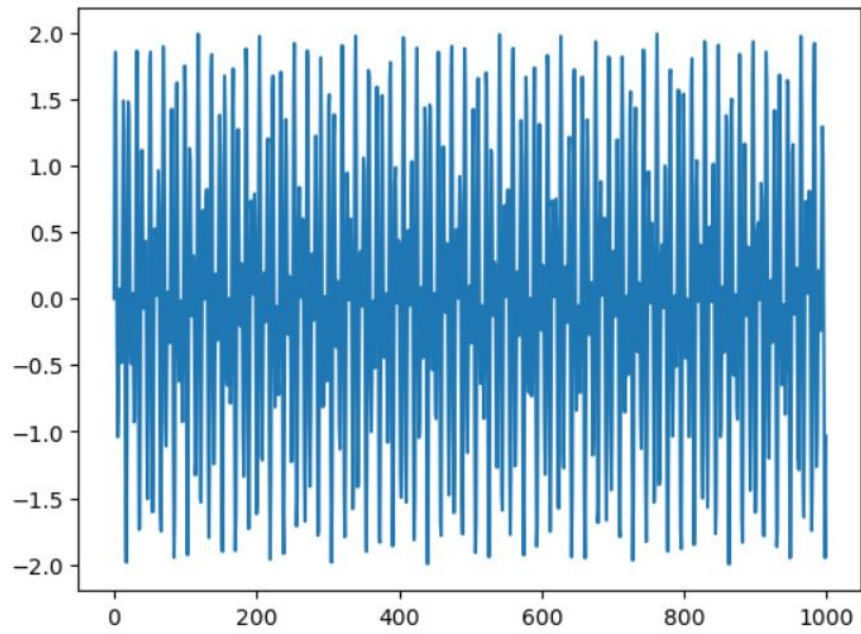
number 6



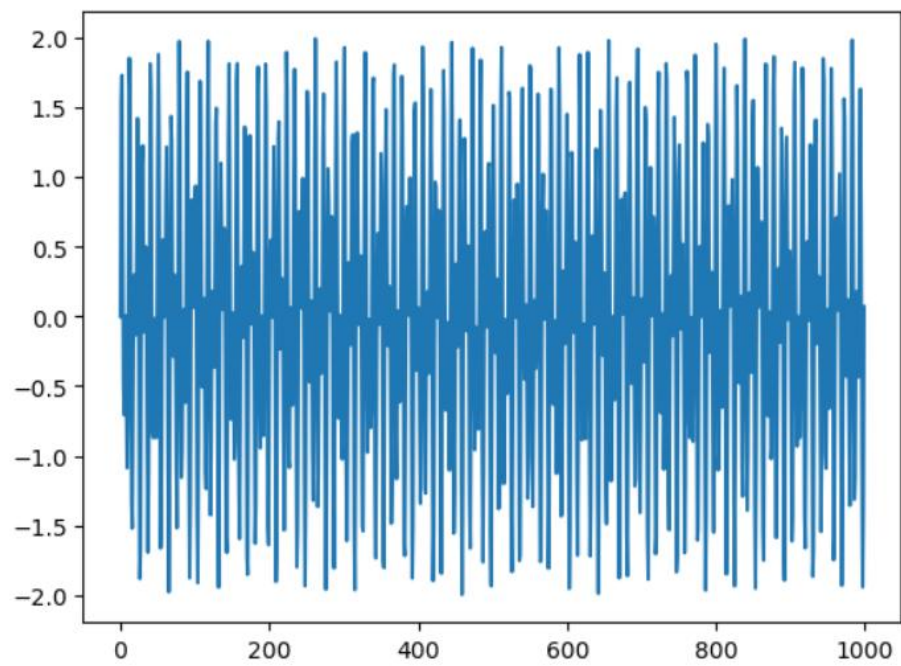
number 7



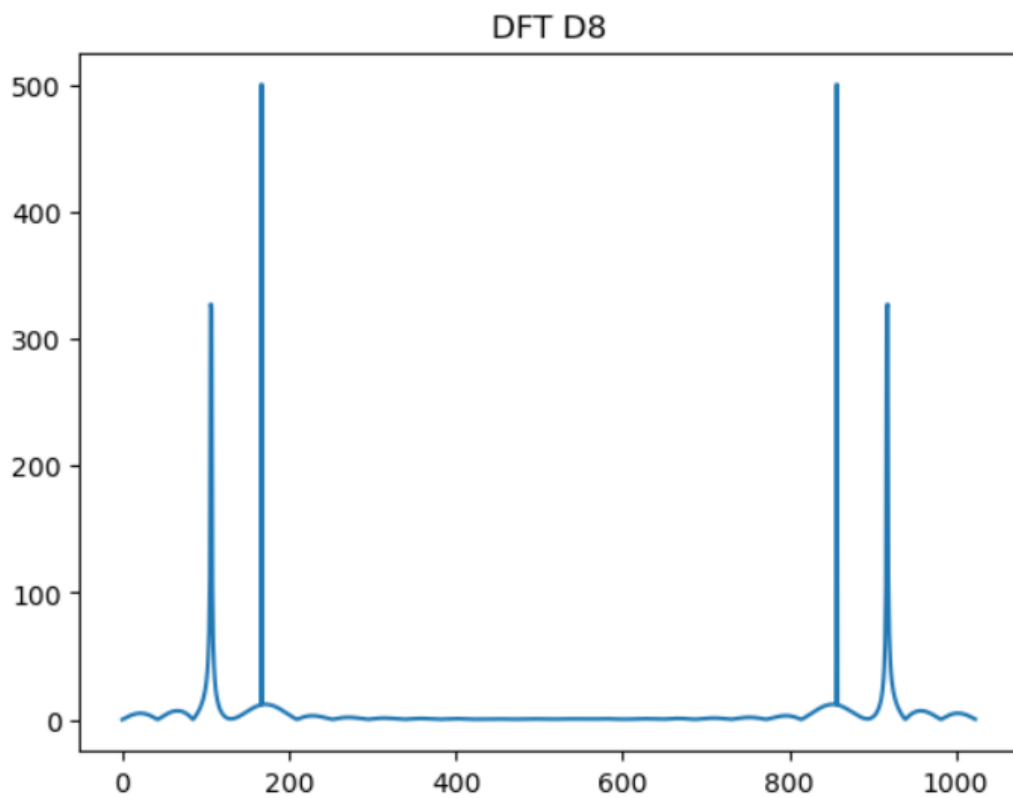
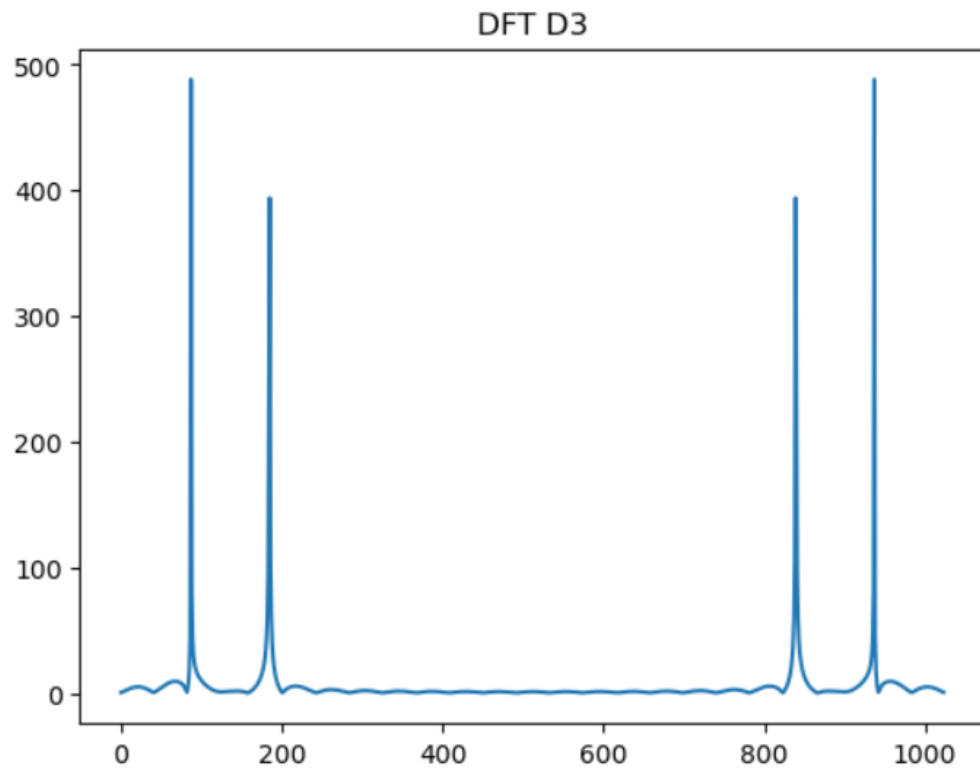
number 8

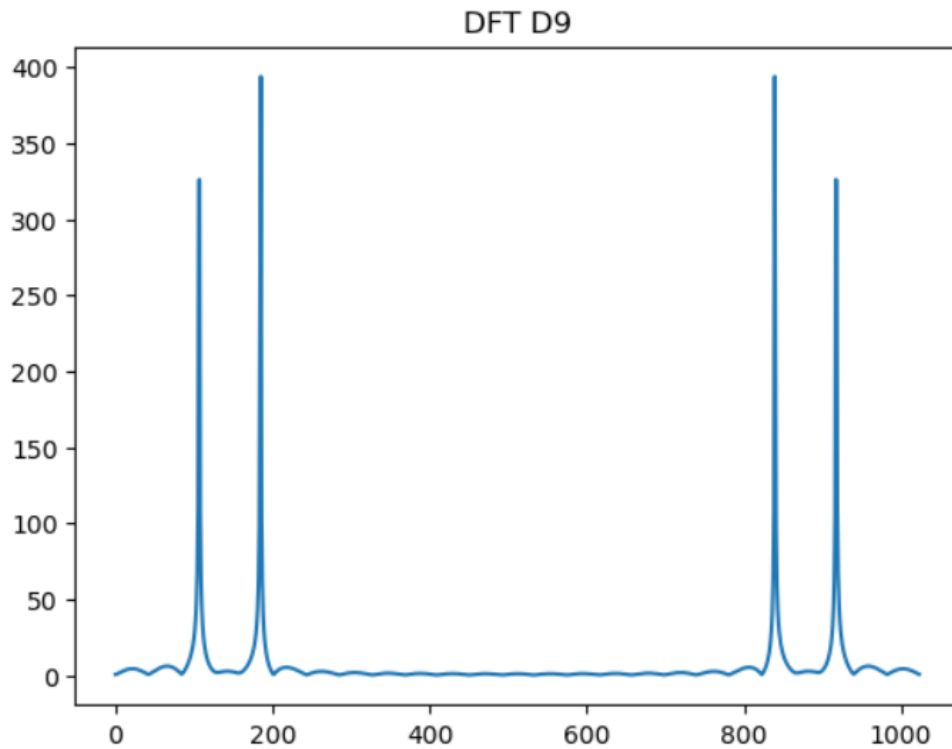


number 9

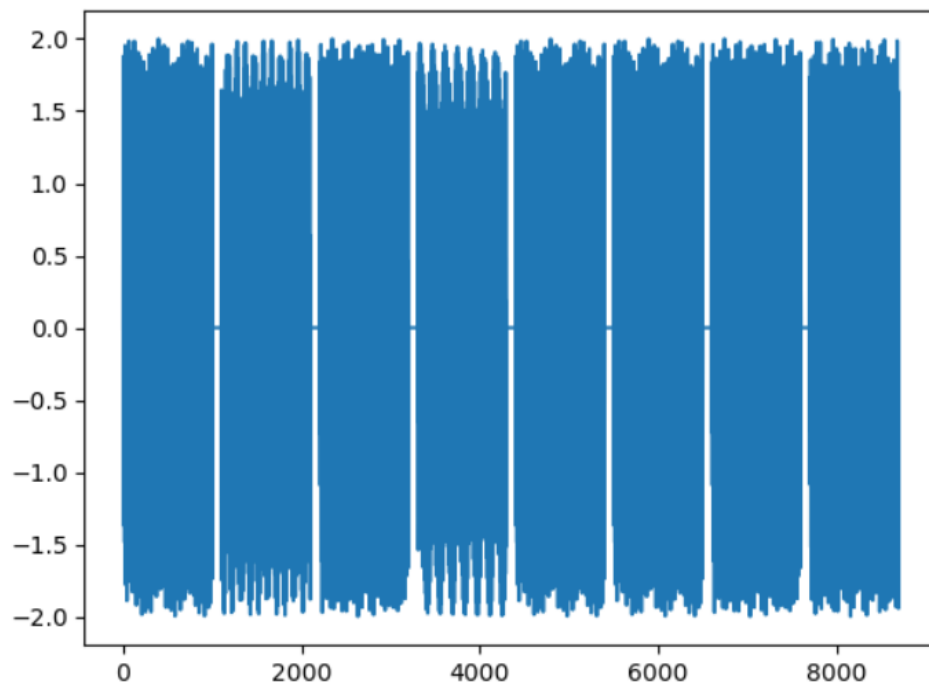


1.2 Με την εντολή `np.fft.fft()` της `numpy` σχεδιάζουμε τον DFT 1024 δειγμάτων των τόνων `d3[n]`, `d8[n]` και `d9[n]` και προκύπτουν οι εξής γραφικές παραστάσεις:

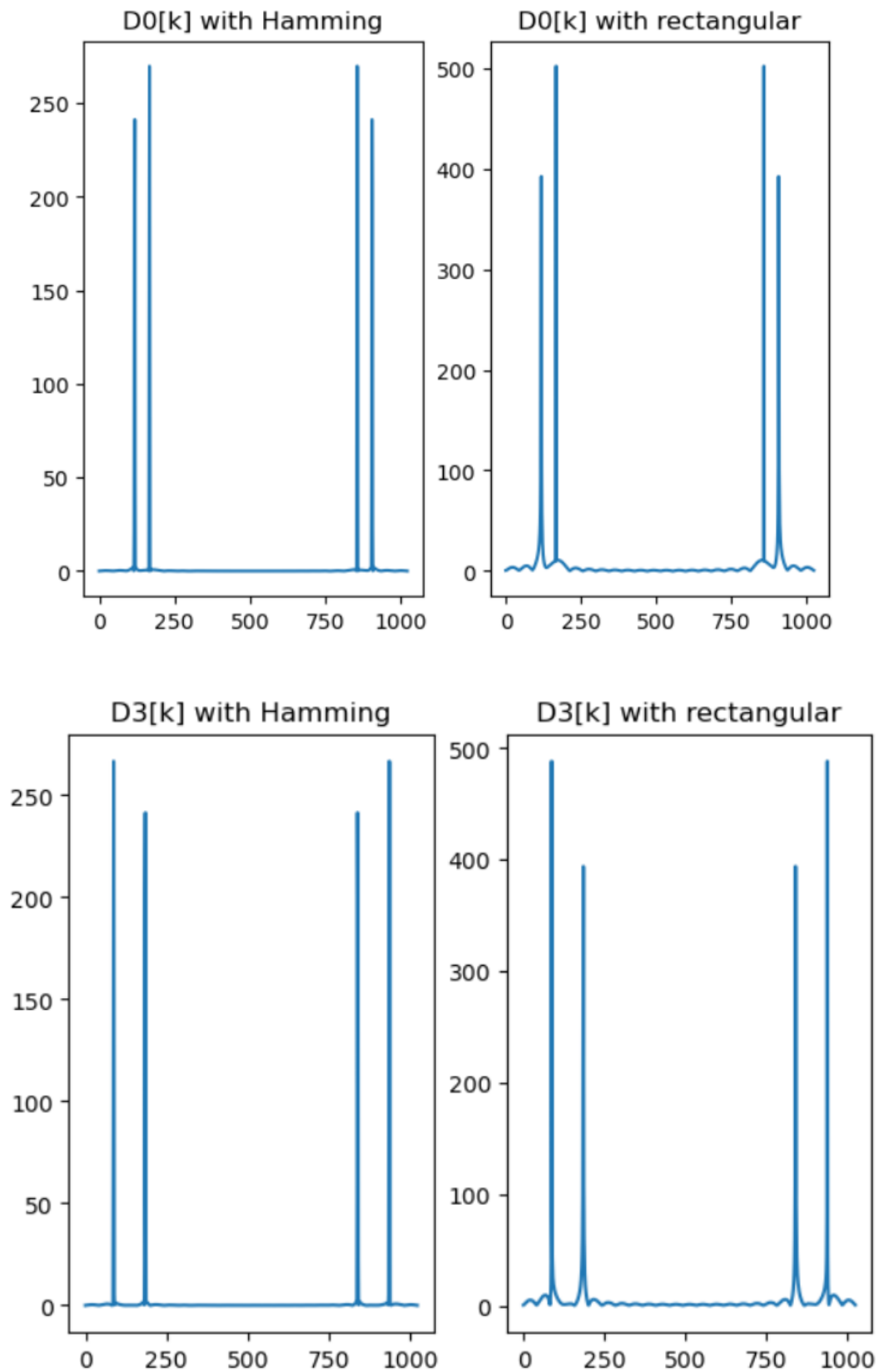


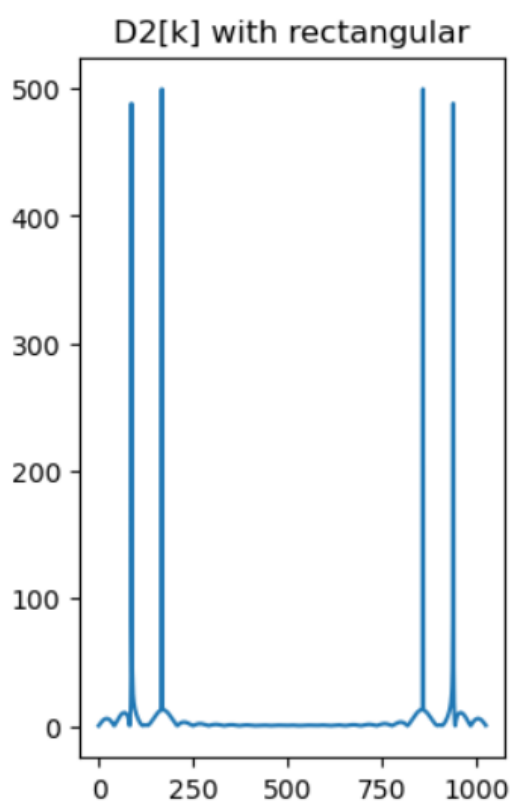
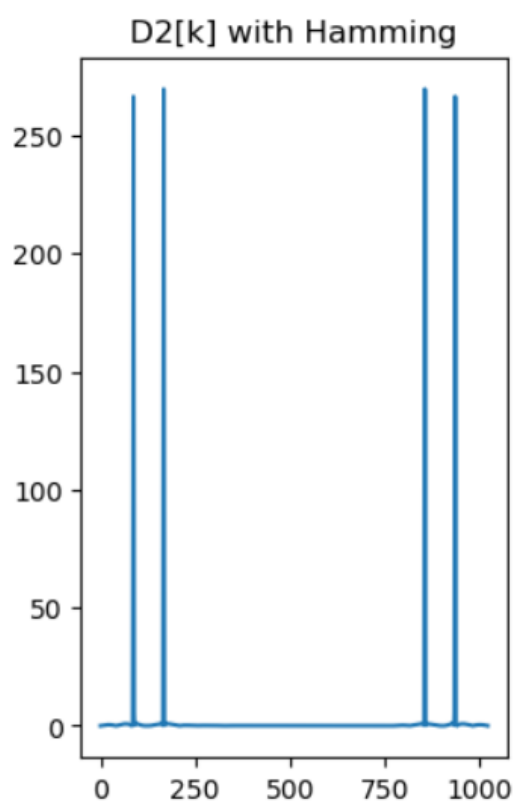
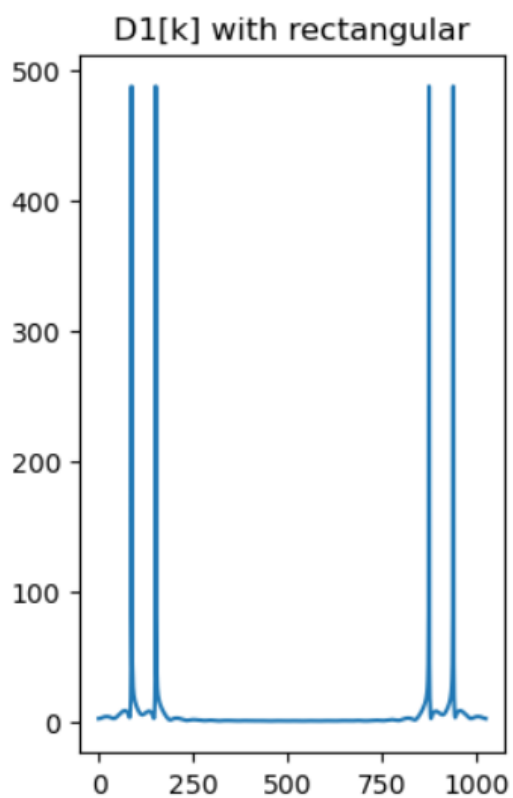
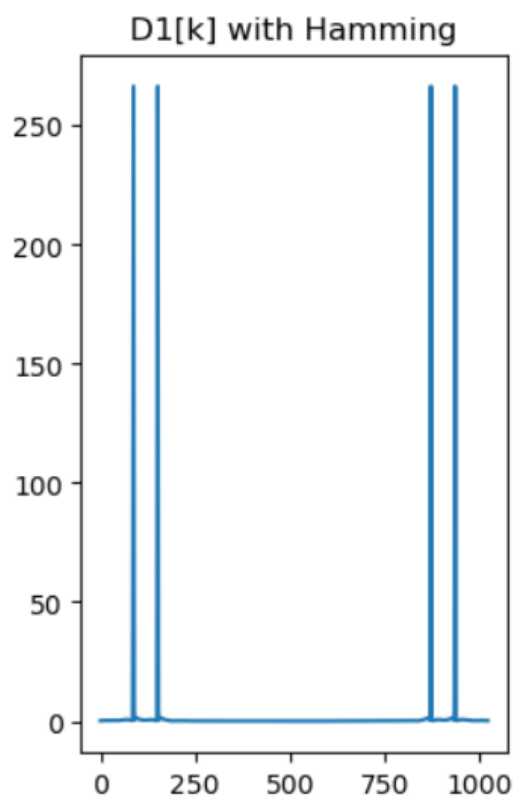


1.3 Ο αριθμός μητρώου είναι 03120019. Δημιουργήσαμε το σήμα και το αποθηκεύσαμε σε ένα αρχείο "tone_sequence.wav". Χρησιμοποιούμε την εντολή play της sounddevice για να ακούσουμε το σήμα μας. Παρακάτω φαίνεται η γραφική παράσταση των τονικών σημάτων που αντιστοιχούν σε κάθε ψηφίο του αριθμού μητρώου. Η προσθήκη μηδενικών μετά από κάθε τόνο τους κάνει πιο ευδιάκριτους στο ηχητικό σήμα.

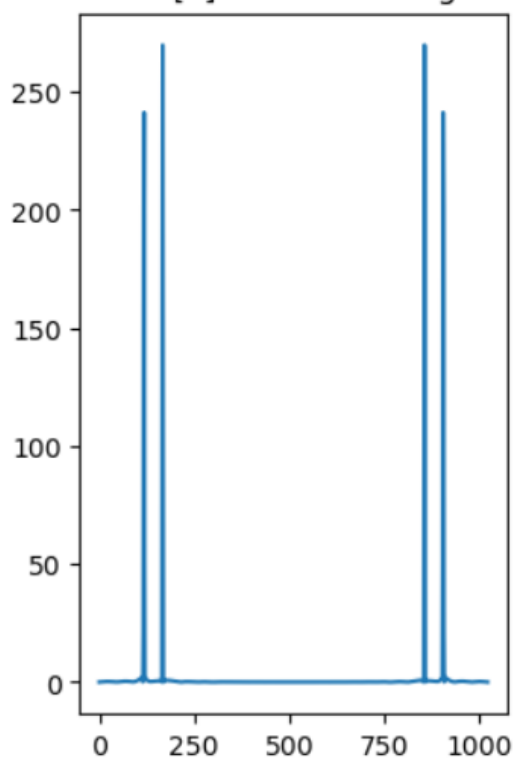


1.4 Με την εντολή `fft()` της `numpy` υπολογίζουμε τον DFT 1024 δειγμάτων των παραθυροποιημένων σημάτων (τετραγωνικά παράθυρα και παράθυρα Hamming) και προκύπτουν οι εξής γραφικές παραστάσεις:

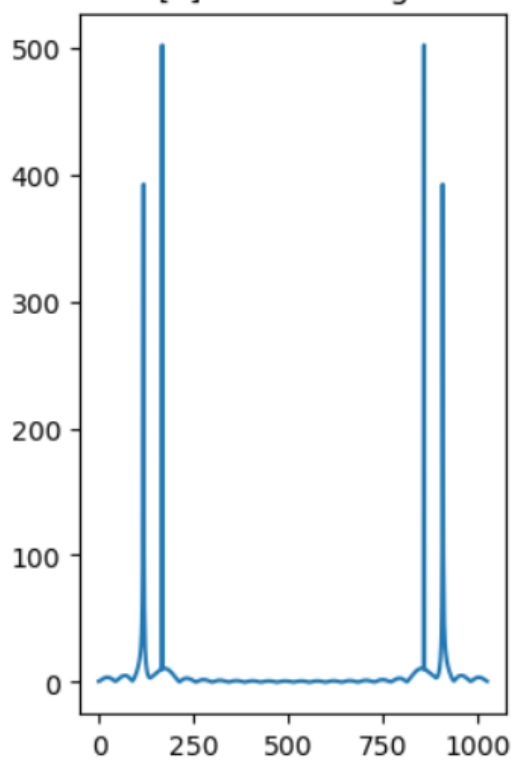




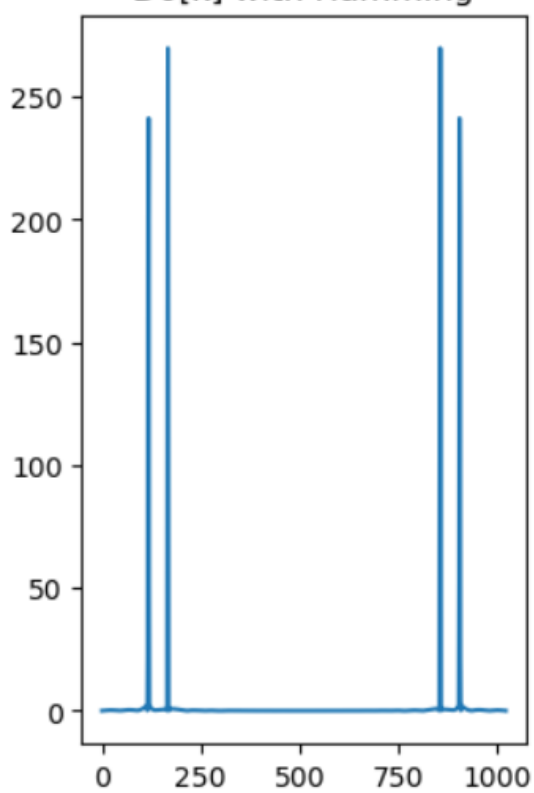
D0[k] with Hamming



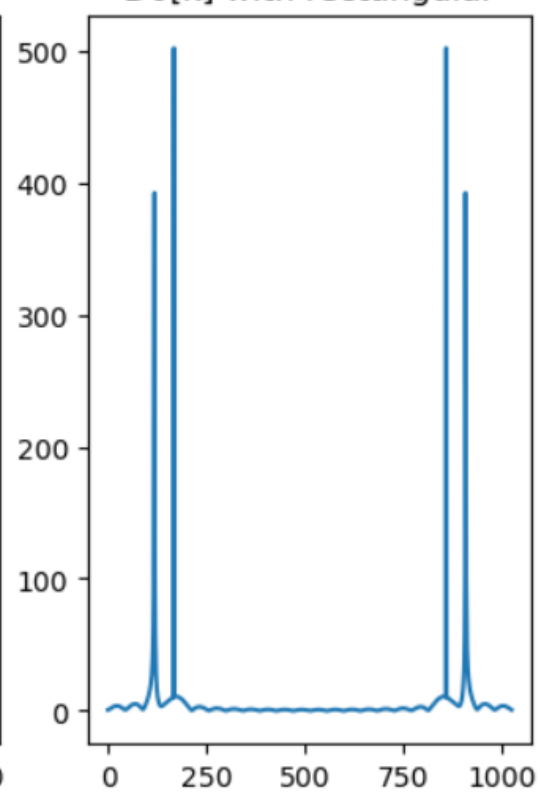
D0[k] with rectangular

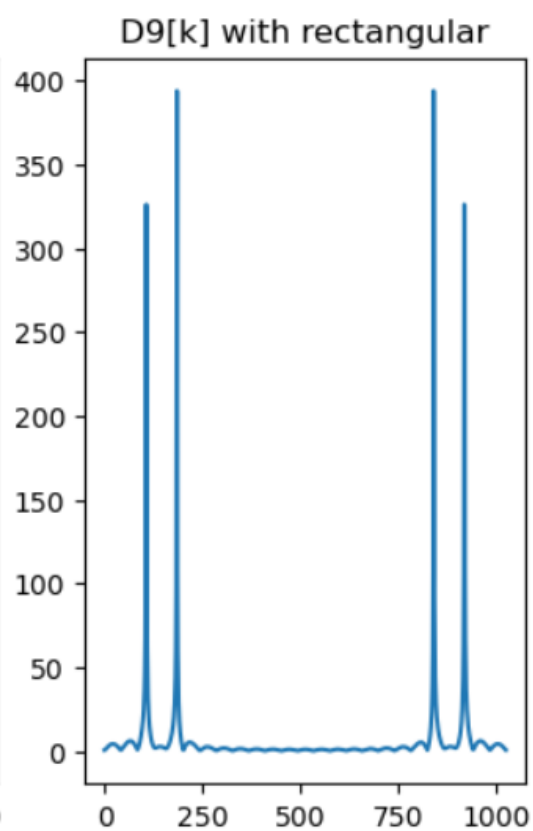
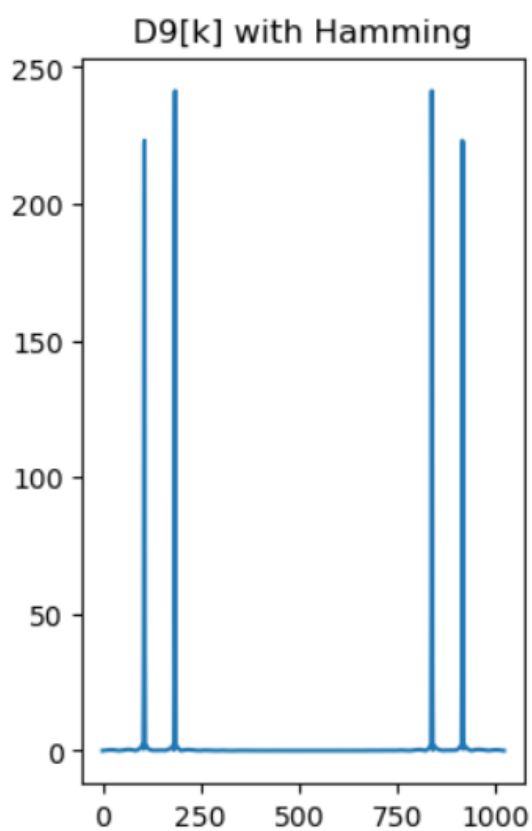
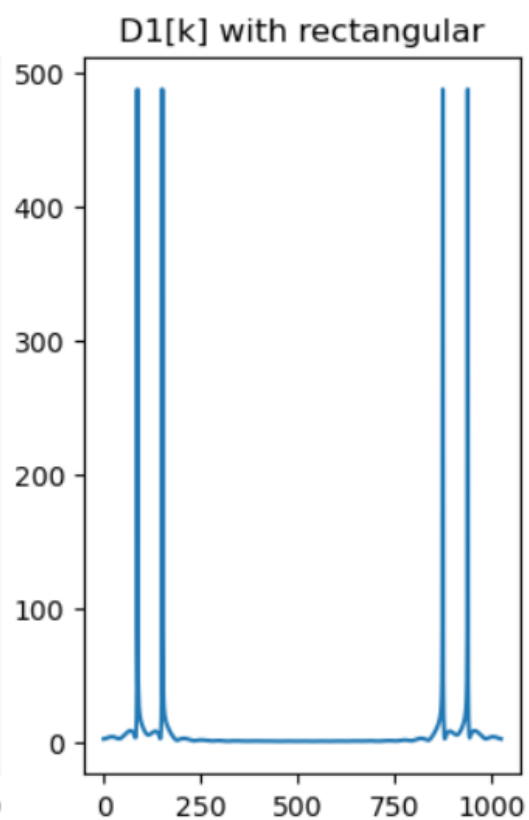
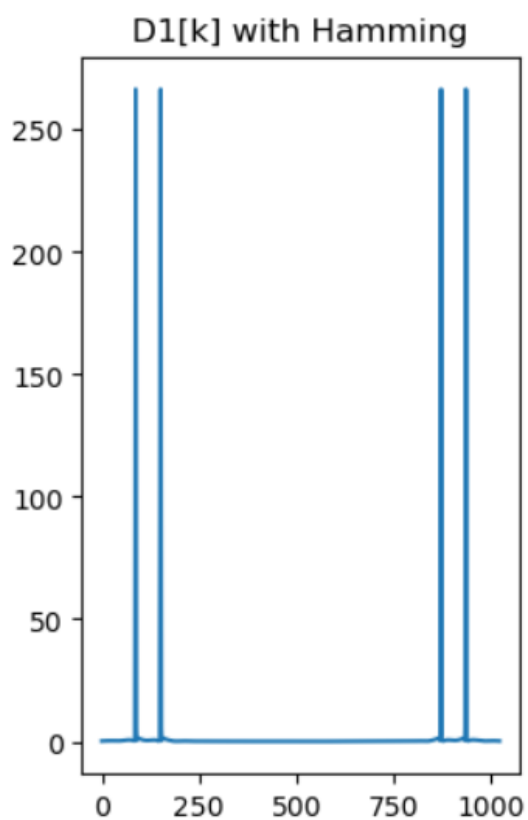


D0[k] with Hamming

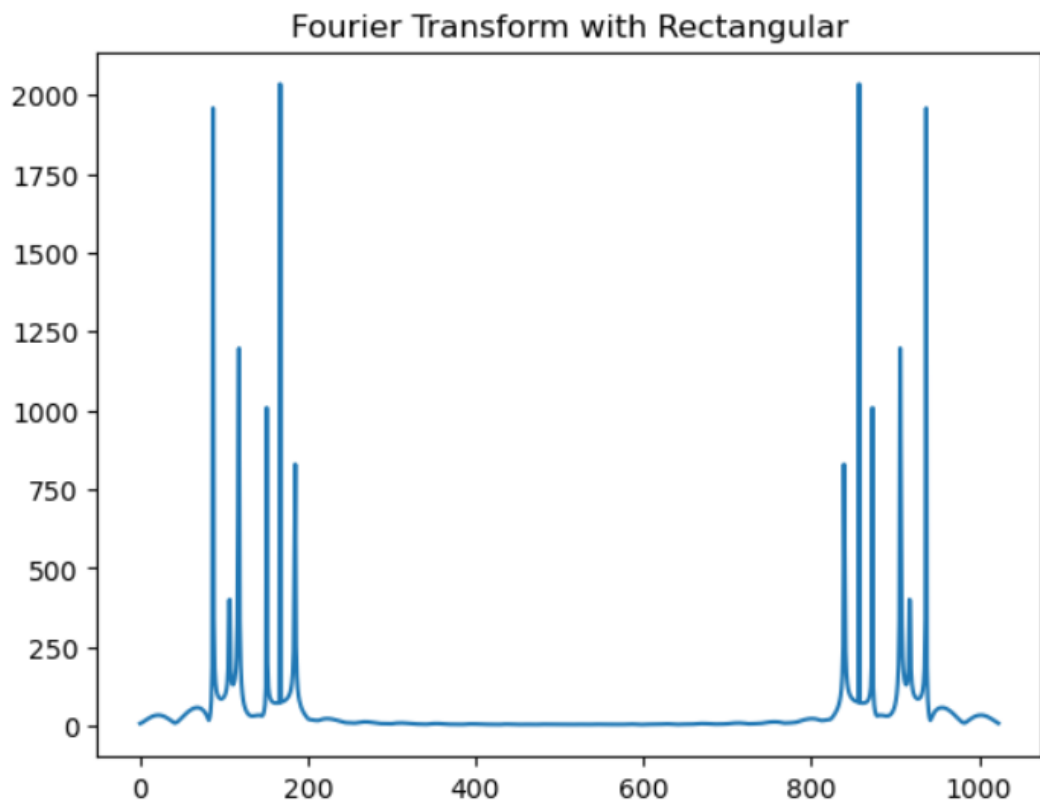
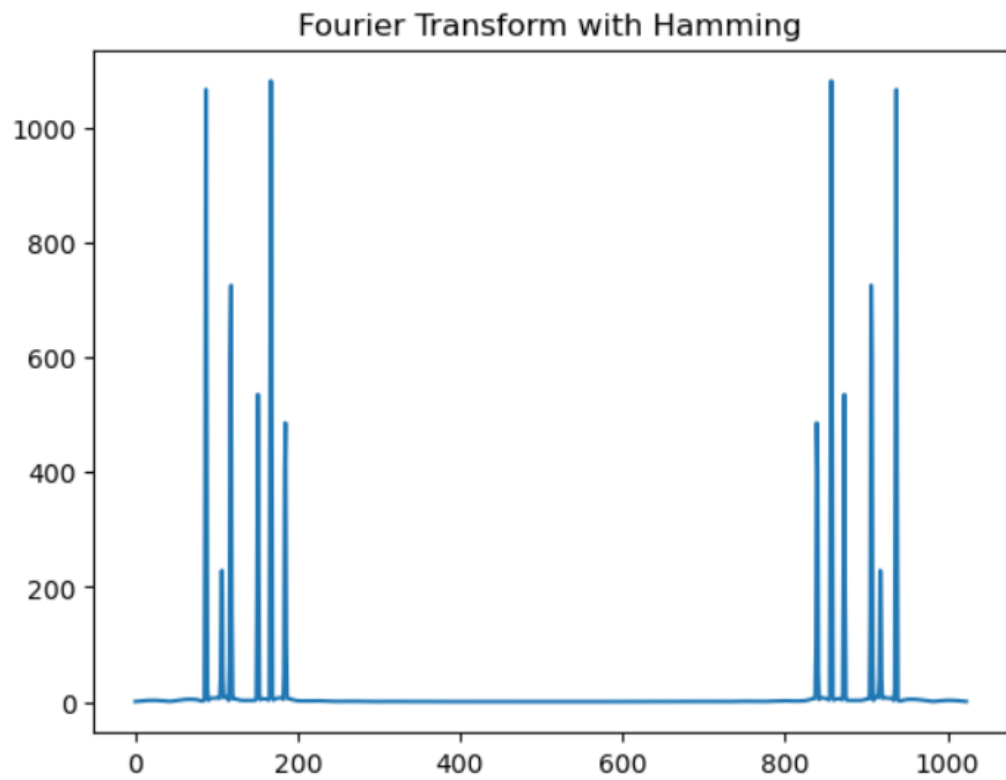


D0[k] with rectangular





Συνολική αναπαράσταση όλων των ψηφίων:



- 1.5 Δημιουργούμε μια συνάρτηση `getIndexes()` με την οποία εντοπίζουμε ποιες συχνότητες από ένα σύνολο δεδομένων FFT αντιστοιχούν στις προκαθορισμένες συχνότητες που χρησιμοποιούνται στα τηλεφωνικά πλήκτρα. Για να το κάνουμε αυτό μετατρέπουμε τους δείκτες συχνοτήτων του FFT σε κανονικοποιημένες τιμές και υπολογίζουμε ενδιάμεσα όρια για την ταξινόμησή τους. Στη συνέχεια, συγκρίνουμε κάθε συχνότητα με τα όρια αυτών των ζωνών και αντιστοιχίζουμε την πιο κοντινή τιμή προκαθορισμένης συχνότητας, αποθηκεύοντας μόνο μοναδικές συχνότητες. Για επαλήθευση δοκιμάζουμε τη λειτουργία του με ένα σετ δοκιμαστικών τιμών παραπλήσιες στις touch tone συχνότητες.

```
#testing
u = 1/(np.pi * 2)
ks = [0.53, 0.58, 0.65, 0.70, 0.92, 1, 1.1]
ks = [(x * u * N) for x in ks]

print(getIndexes(ks,N))

{0.5346, 0.5906, 0.7217, 0.6535, 0.9273, 1.0247, 1.1328}
```

- 1.6 Κατασκευάζουμε τη συνάρτηση `getTonesEnergy()` η οποία χωρίζει το σήμα σε τμήματα μήκους 1000 δειγμάτων, εφαρμόζει ένα Hamming παράθυρο για μείωση των παραμορφώσεων και υπολογίζει τον FFT κάθε τμήματος για να βρει την ενεργειακή κατανομή των συχνοτήτων. Στη συνέχεια κατασκευάζουμε την `ttdecode()` η οποία εντοπίζει τις δύο κυρίαρχες συχνότητες κάθε τμήματος, τις κανονικοποιεί και τις αντιστοιχίζει στις προδιαγεγραμμένες συχνότητες μέσω της συνάρτησης `getIndexes()`. Τέλος, αυτές οι συχνότητες αντιστοιχίζονται στα αριθμητικά πλήκτρα του τηλεφώνου και επιστρέφεται η ακολουθία των ψηφίων που περιείχε το σήμα. Επιβεβαιώνουμε την ορθή λειτουργία της συνάρτησης θέτοντας σαν είσοδο το σήμα του ερωτήματος 1.3.

```
print('Digits of AM:', ttdecode(signall))

Digits of AM: [0, 3, 1, 2, 0, 0, 1, 9]
```

1.7 Με χρήση της εντολής `load()` της `numpy` φορτώνουμε τα αρχεία `easy_sig.npy`, `medium_sig.npy`, `hard_sig.npy` και τα αντιστοιχίσαμε σε τρία διαφορετικά σήματα, `easySig`, `mediumSig`, `hardSig`. Προσδιορίζουμε τα ψηφία στα οποία αντιστοιχούν οι τόνοι των 3 σημάτων με την χρήση της `ttdecode()` που δημιουργήσαμε πριν.

```
easysig = np.load('easy_sig.npy')
print ('easysig digits:', ttdecode(easysig)) # εφαρμόζουμε την παραπάνω συνάρτηση για easys

easysig digits: [2, 0, 1, 3, 6, 5, 1, 5]

mediumsig = np.load('medium_sig.npy')
print ('mediumsig digits:', ttdecode(mediumsig)) # εφαρμόζουμε την παραπάνω συνάρτηση για m

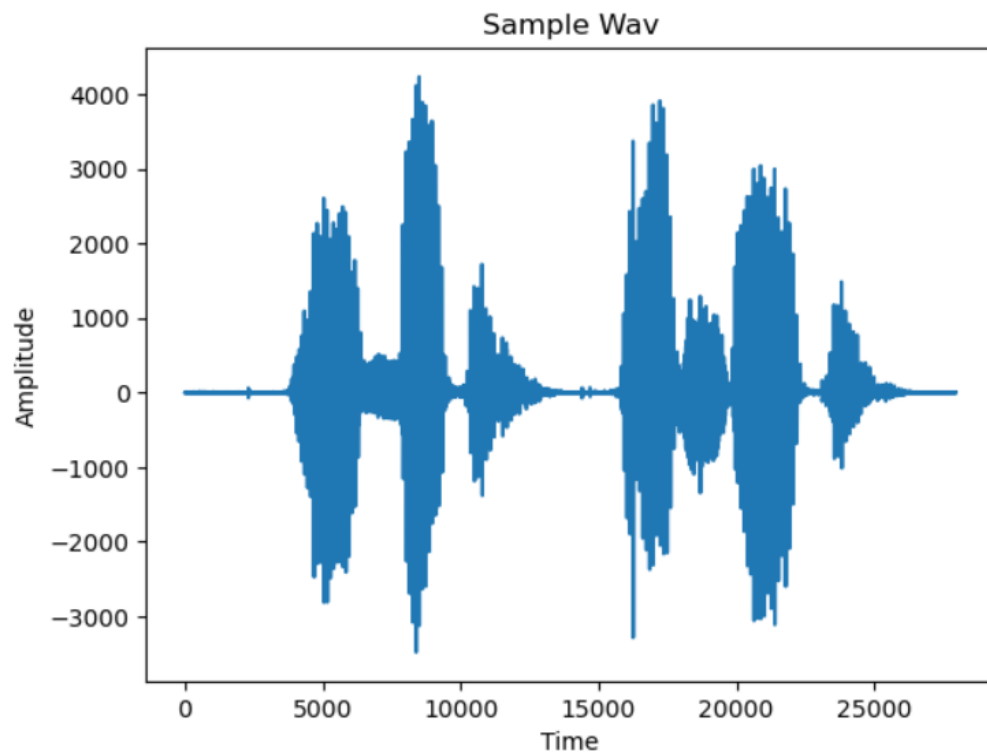
mediumsig digits: [4, 7, 4, 8, 0, 1, 1, 8, 2, 0]

hardsig = np.load('hard_sig.npy')
print ('hardsig digits:', ttdecode(hardsig)) # εφαρμόζουμε την παραπάνω συνάρτηση για hards

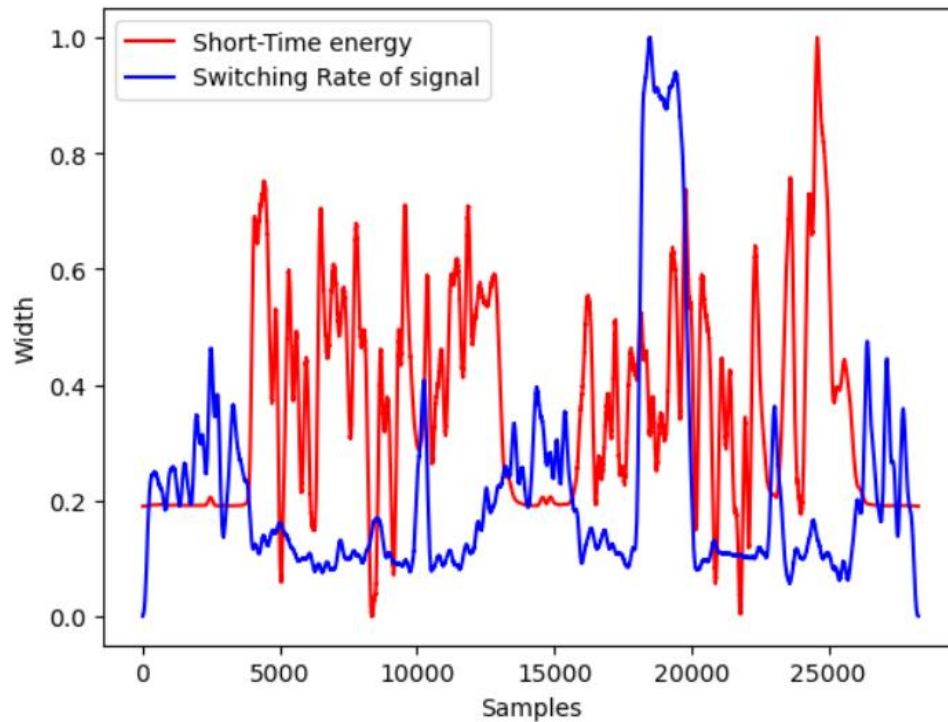
hardsig digits: [0, 3, 6, 0, 9, 3, 7, 3, 0, 4]
```

Μέρος 2° – Χαρακτηριστικά Βραχέος Χρόνου Σημάτων Φωνής και Μουσικής

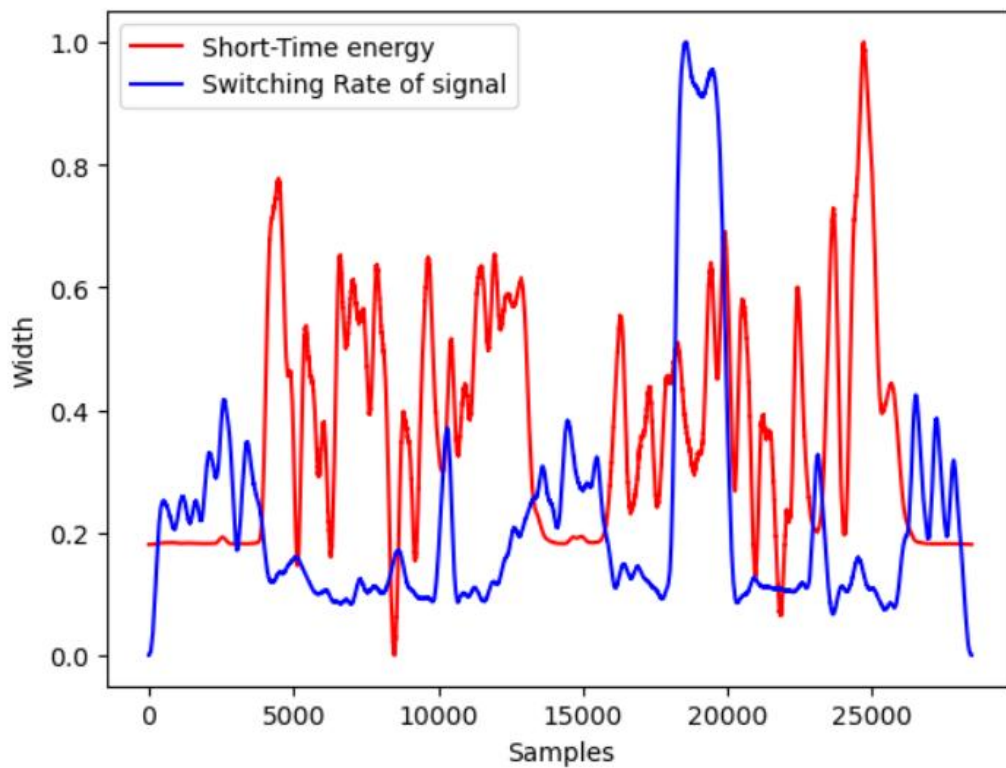
2.1 Ανοίγουμε το “speech_utterance.wav” από το υλικό που υπάρχει στο “dsp25_lab1_Data.zip” και παρακάτω απεικονίζεται το σήμα στο πεδίο του χρόνου:



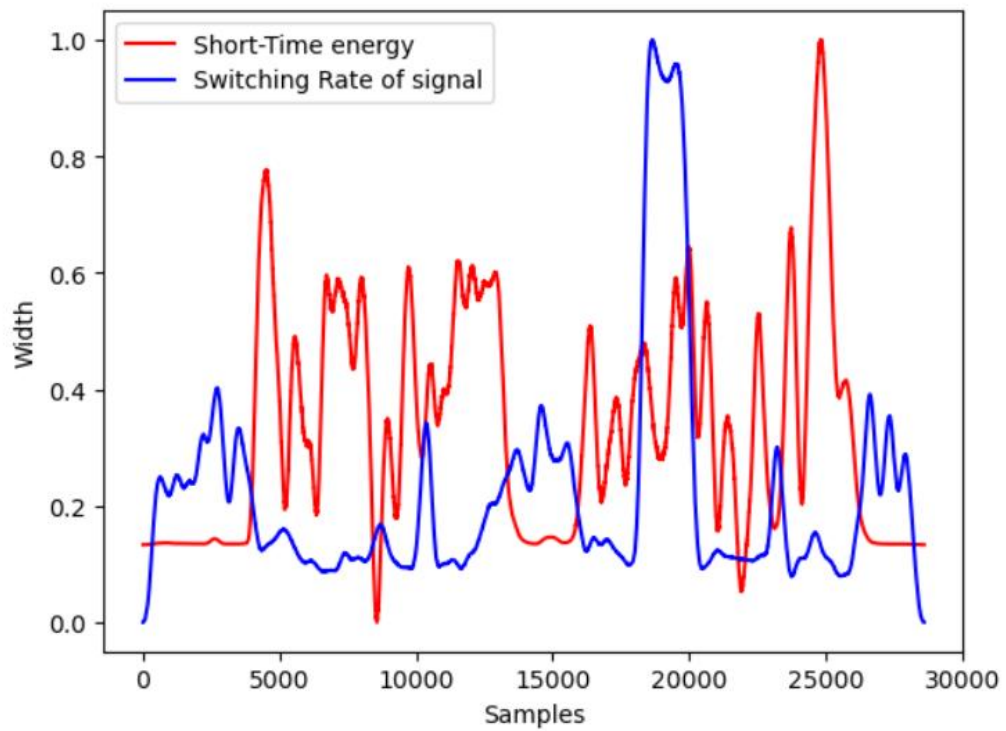
2.2 Για συχνότητα δειγματοληψίας 16kHz, και παράθυρο μήκους έστω 20ms, θα λάβουμε παράθυρο hamming μήκους 320 δειγμάτων.



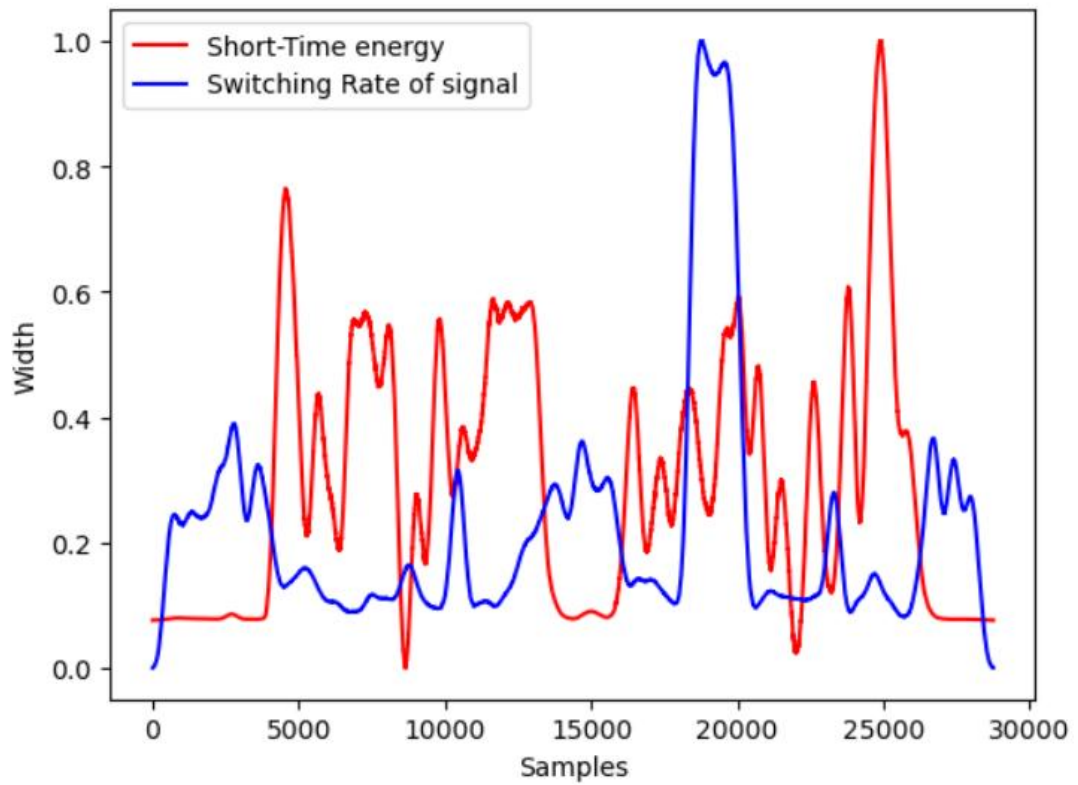
Για συχνότητα δειγματοληψίας 16kHz, και παράθυρο μήκους έστω 30ms, θα λάβουμε παράθυρο hamming μήκους 480 δειγμάτων.



Για συχνότητα δειγματοληψίας 16kHz, και παράθυρο μήκους έστω 40ms, θα λάβουμε παράθυρο hamming μήκους 640 δειγμάτων

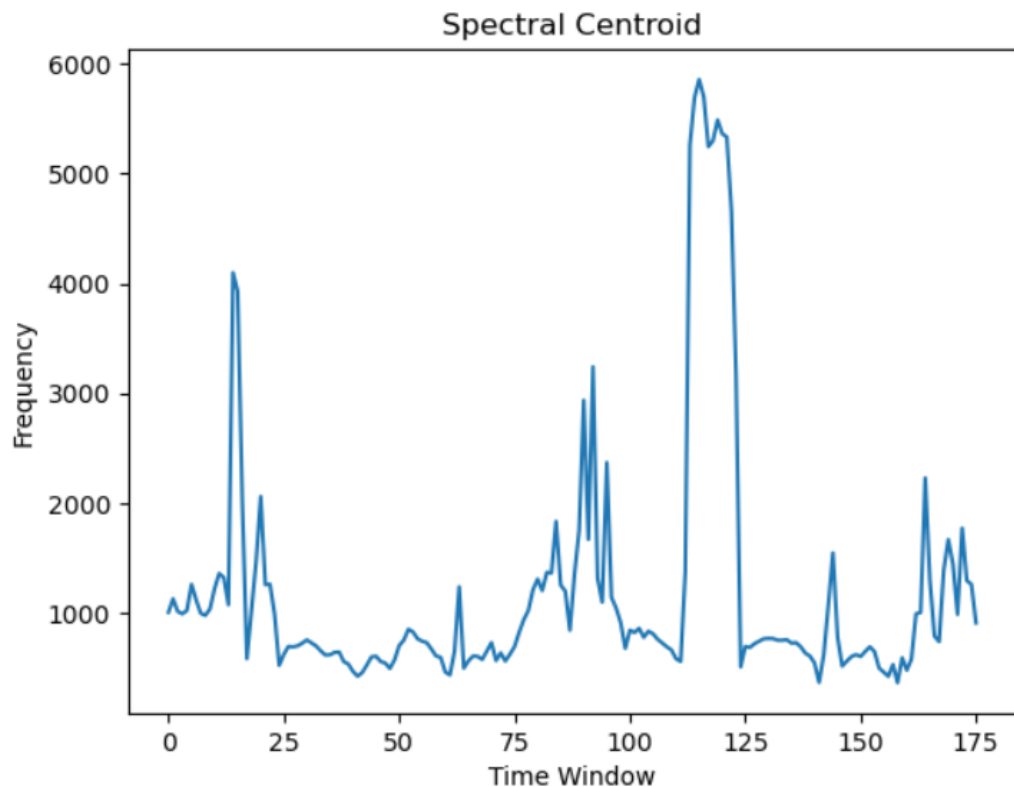


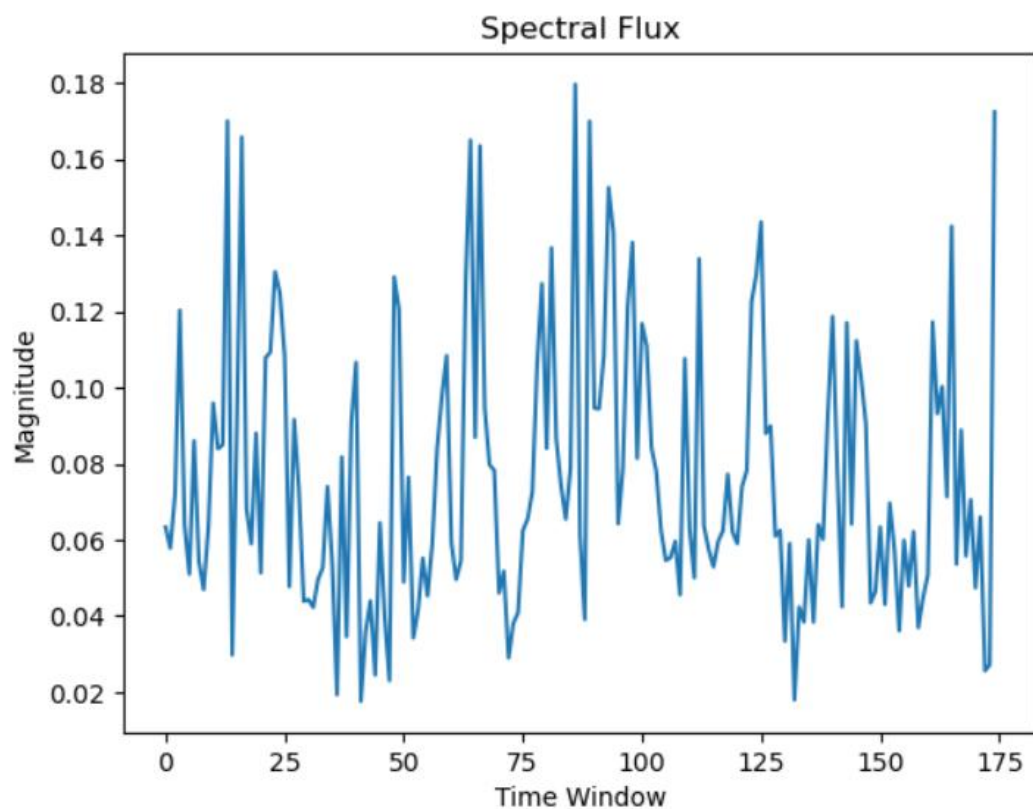
Για συχνότητα δειγματοληψίας 16kHz, και παράθυρο μήκους έστω 50ms, θα λάβουμε παράθυρο hamming μήκους 800 δειγμάτων



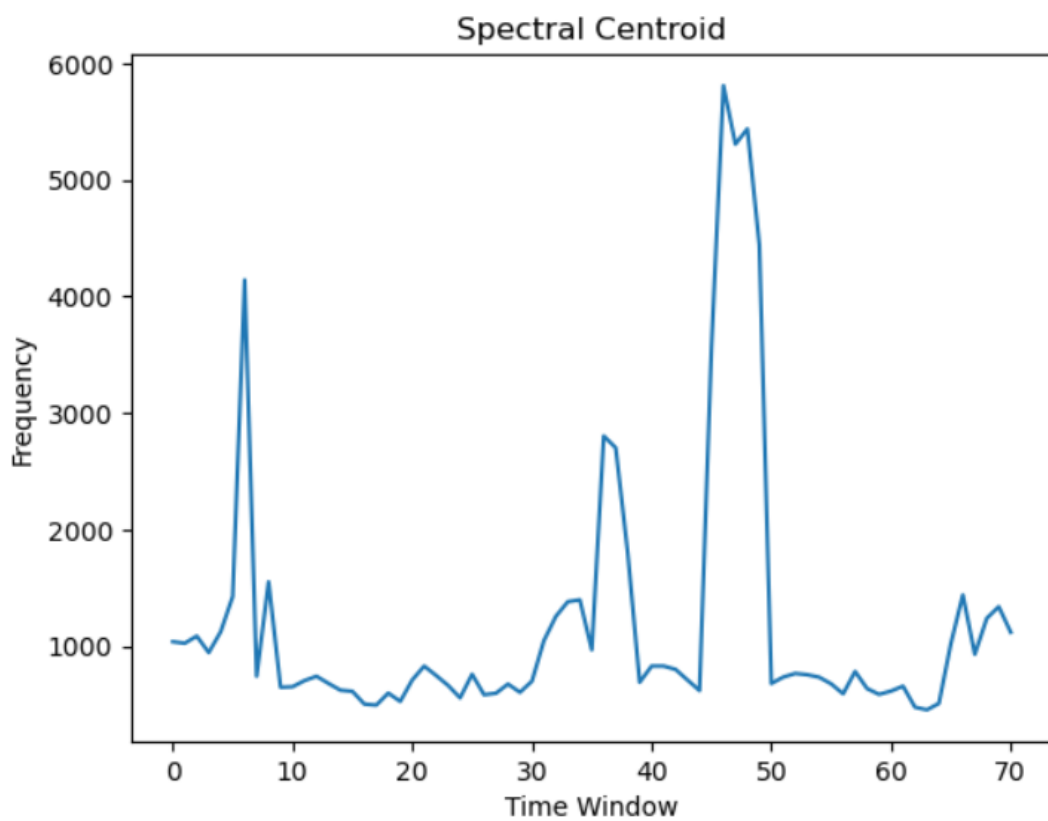
Παρατηρούμε πως όσο αυξάνεται το μήκος του παραθύρου τα διαγράμματα παρουσιάζουν σημαντική εξομάλυνση, όπως αναμέναμε. Παρά την εξομάλυνση ωστόσο, δεν υπάρχει σημαντική απώλεια πληροφορίας όσον αφορά την παρουσία έμφωνων ή άφωνων ήχων. Ακόμα σε σημεία όπου η ενέργεια βραχέος χρόνου είναι μη μηδενική, αντιστοιχίζουμε φωνήματα. Είναι λογικό ακόμα να υποθέσουμε πως σε σημεία που ο ρυθμός αλλαγής προσήμου κορυφώνεται (δηλαδή σε σημεία όπου το σήμα δείχνει να έχει υψηλή συχνότητα), αντιστοιχίζονται άφωνοι ήχοι (αφού αυτοί είναι πιο υψίσυχνοι), ενώ στις περιοχές με την υψηλότερη ενέργεια βραχέος χρόνου να αντιστοιχίζονται έμφωνοι ήχοι (αφού αυτοί είναι πιο δυνατοί). Ακόμα, σε σημεία με μηδενική σχεδόν ενέργεια αλλά μεγάλο zero-crossing rate έχουμε περιόδους σιωπής, στις οποίες λόγω εξωτερικών θορύβων και μικροφωνισμού παρατηρούμε μεγάλο zerocrossing rate.

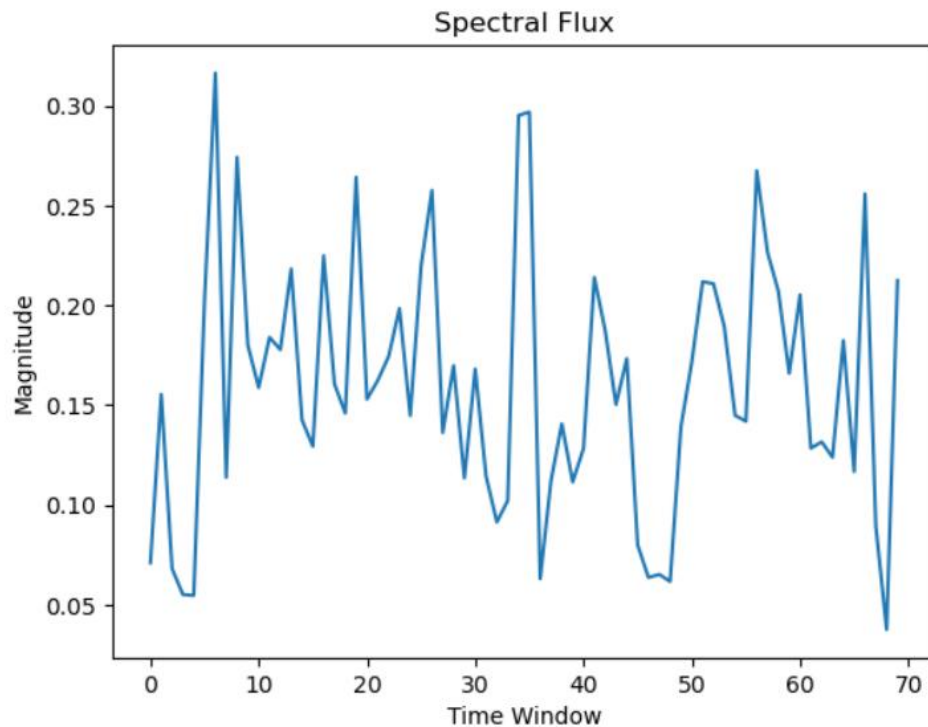
2.3 Για παράθυρο μήκους 20ms





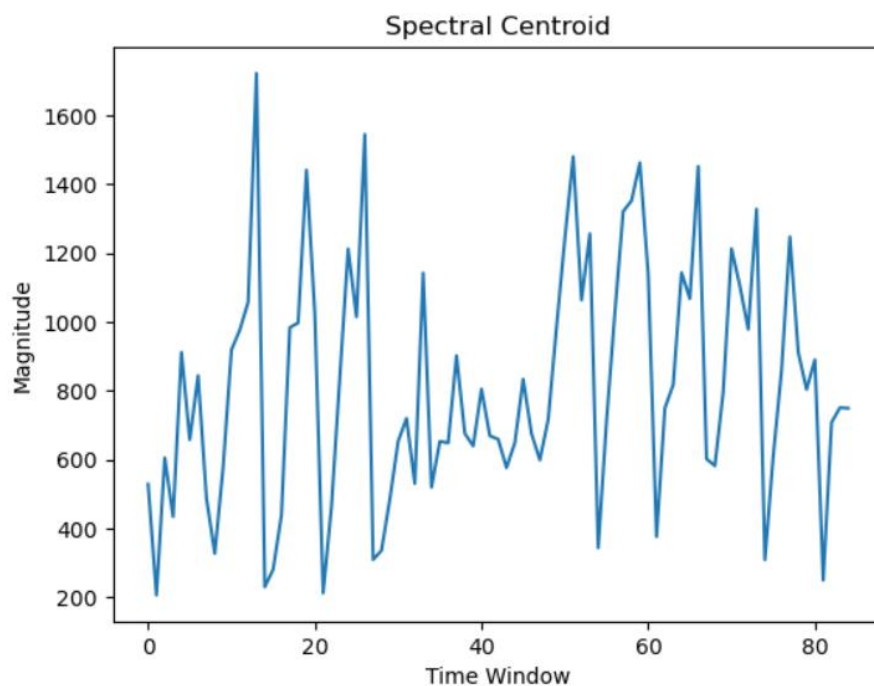
Για παράθυρο μήκους 50ms

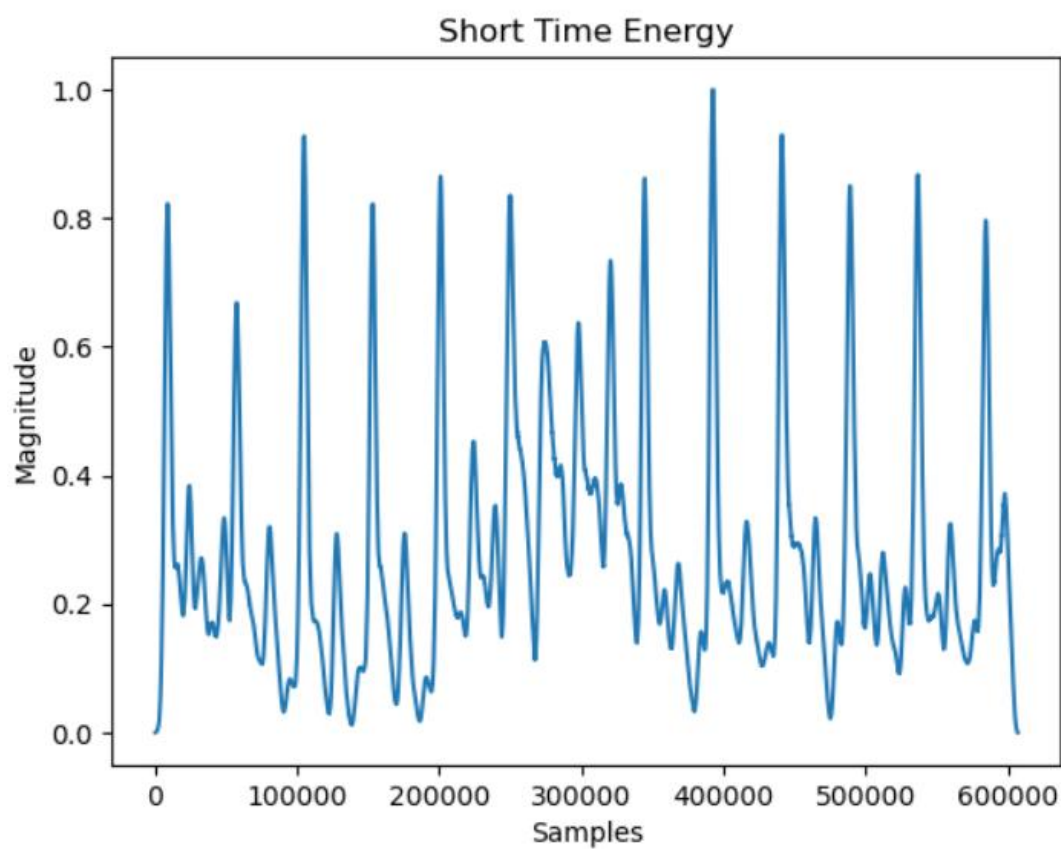
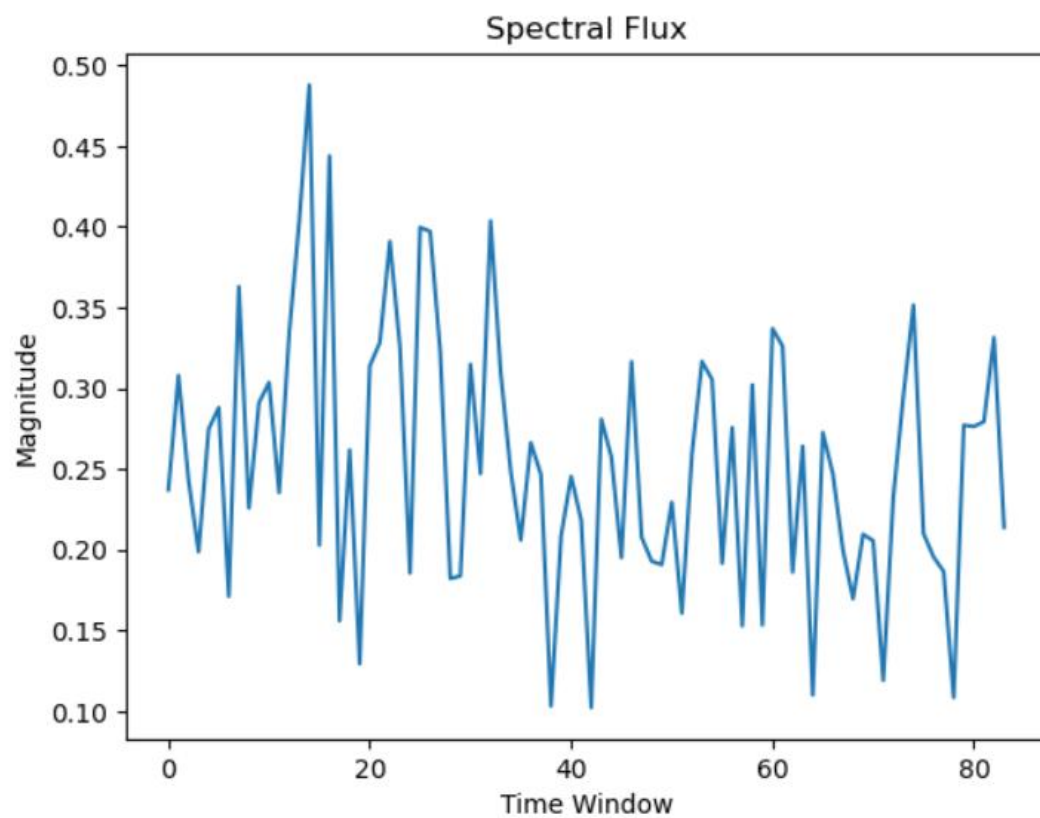


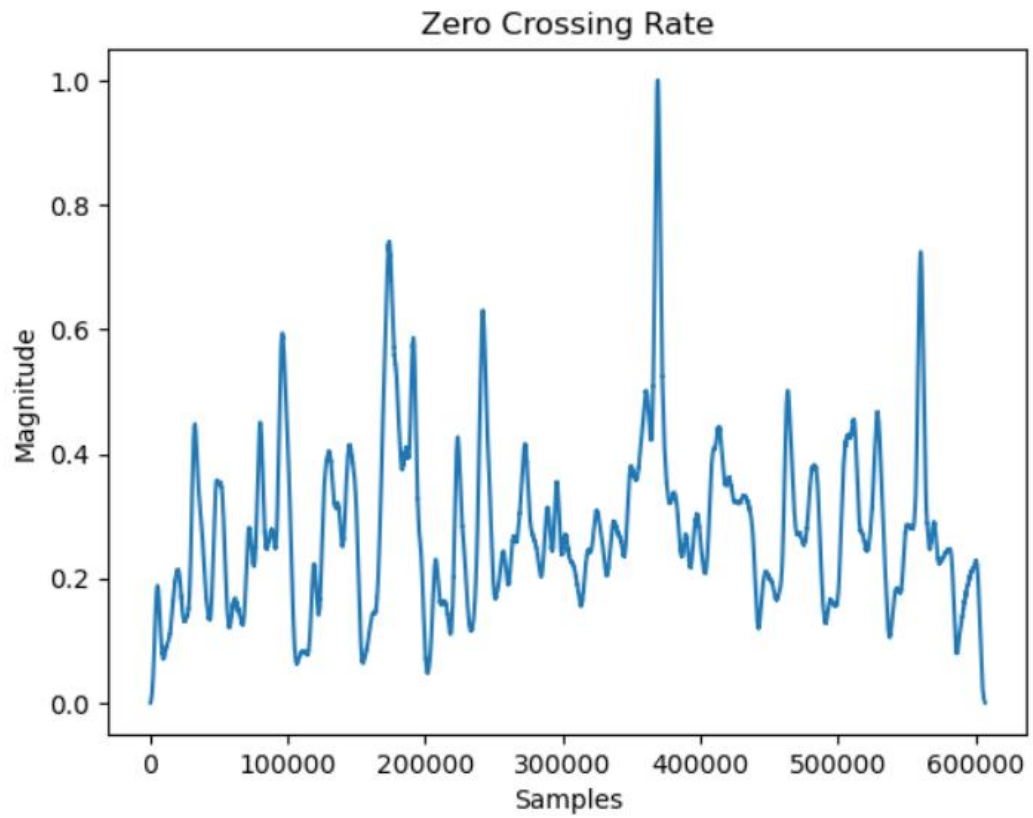


Παρατηρούμε εξομάλυνση των γραφικών παραστάσεων με την αύξηση του μήκους του παραθύρου. Η εξομάλυνση αυτή δεν επηρεάζει την διάκριση των τόνων. Παρατηρούμε ακόμα πως σημεία με μέσο φασματικό κέντρο αντιστοιχούν σε σιωπή, με χαμηλό φασματικό κέντρο σε φωνηέντα, ενώ με φασματικό κέντρο που κορυφώνεται σε σύμφωνα. Ακόμα, η χρονική περιοχή που εμφανίζει μεγάλες αυξομειώσεις στην φασματική ροή είναι και η περιοχή κατά την οποία έχουμε ομιλία, λόγω της γρήγορης εναλλαγής συμφώνων και φωνηέντων.

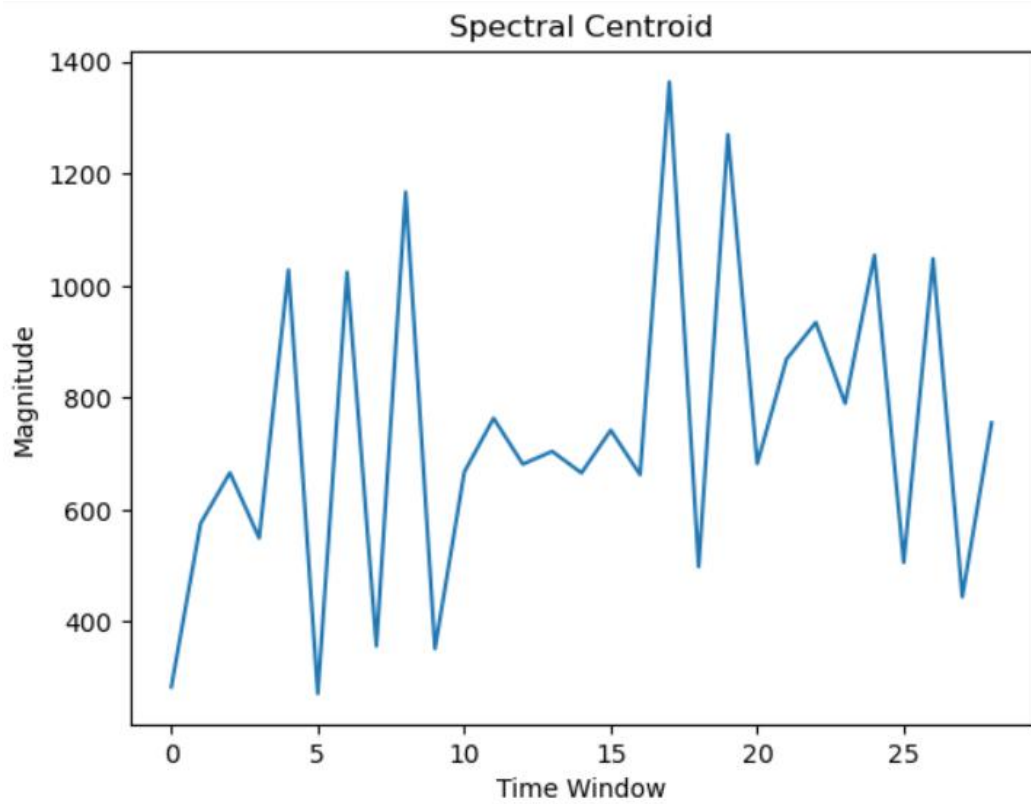
2.4 Για παράθυρο μήκους 150ms

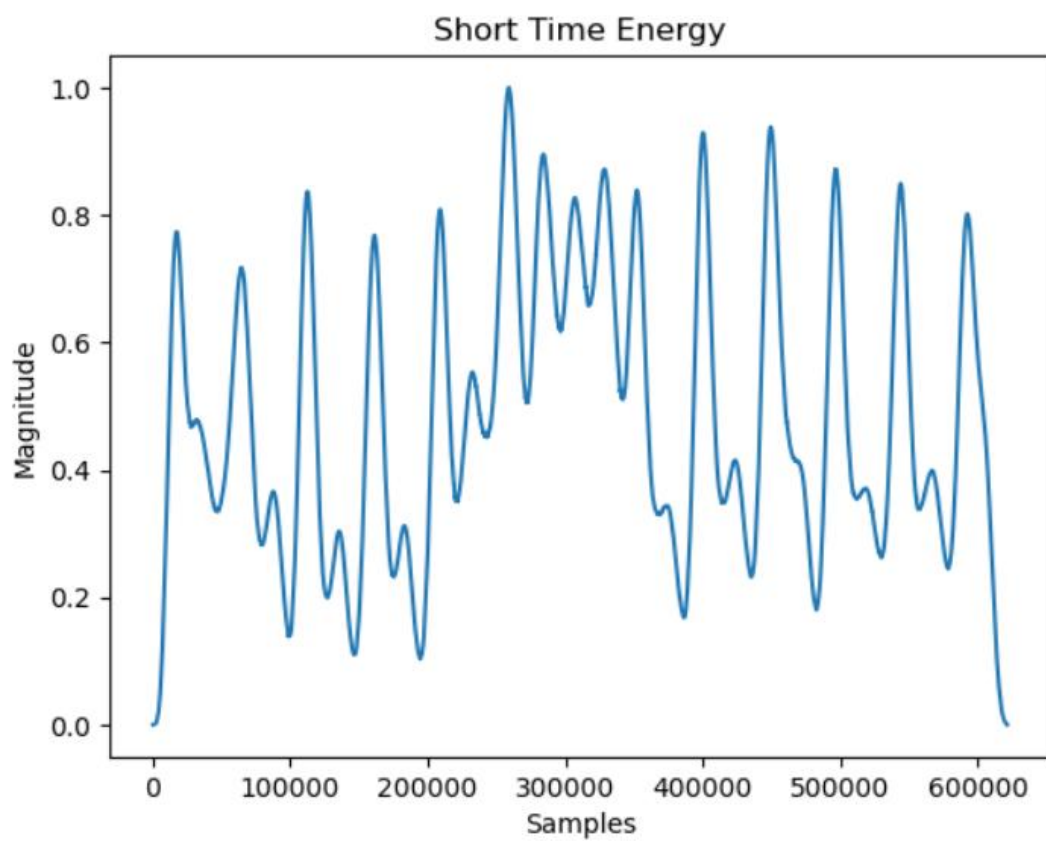
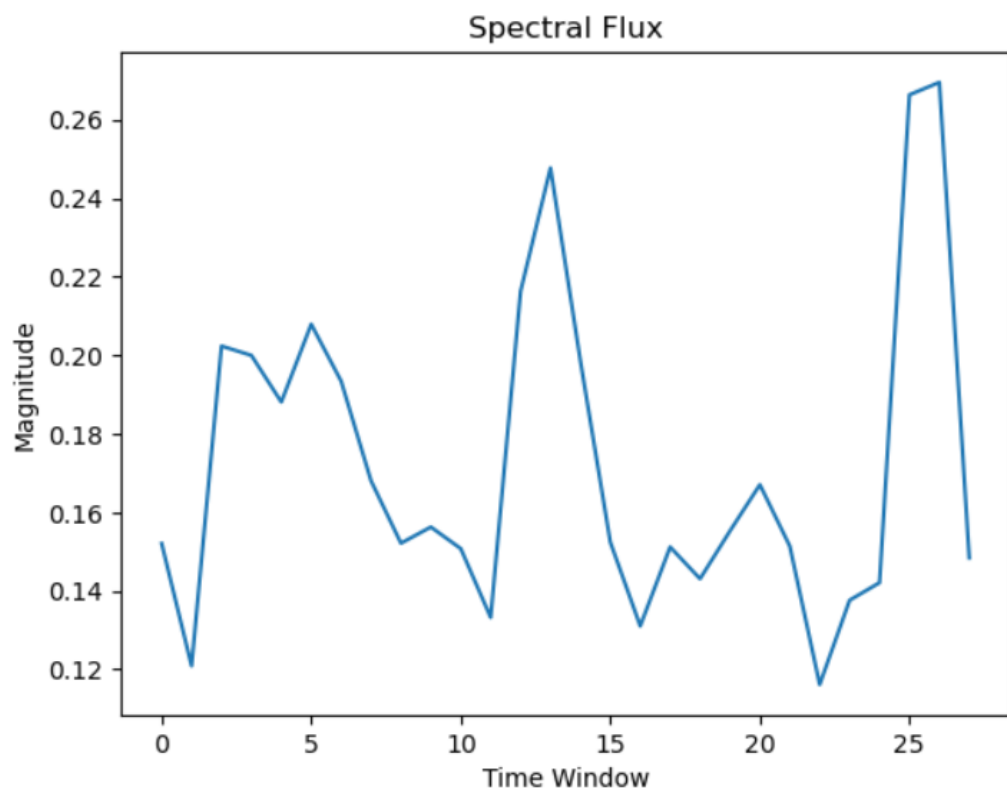


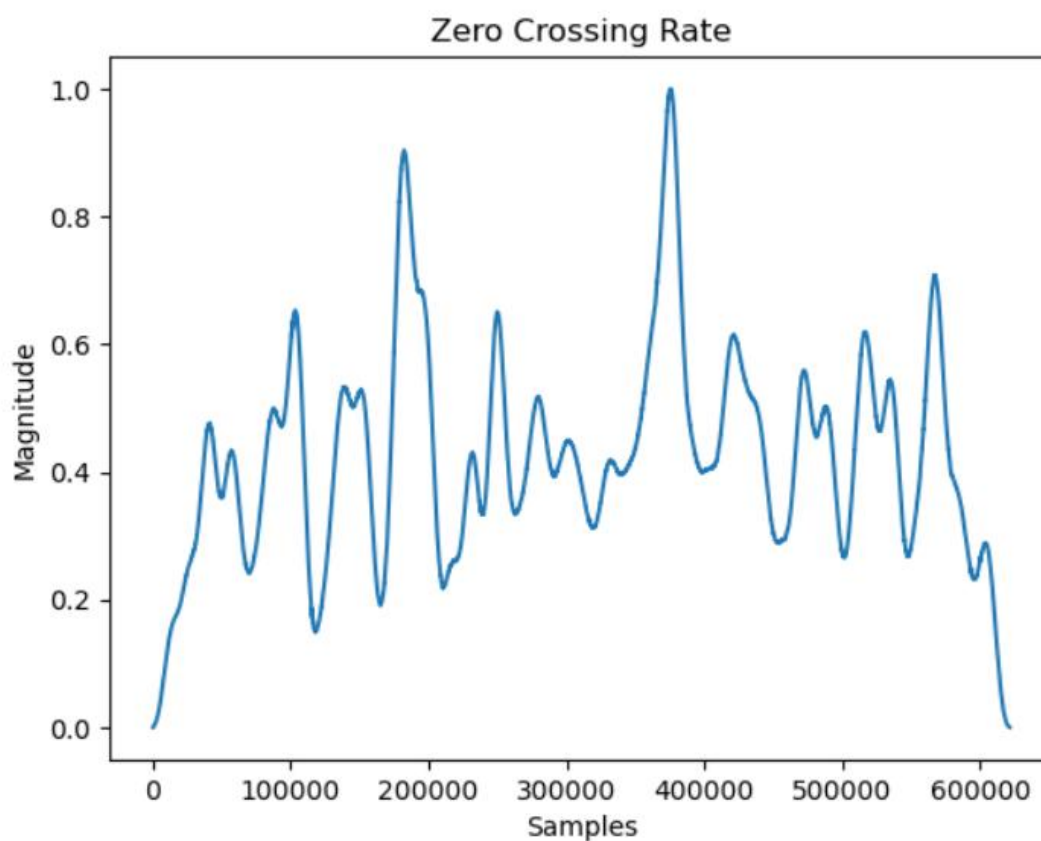




Για παράθυρο μήκους 450ms



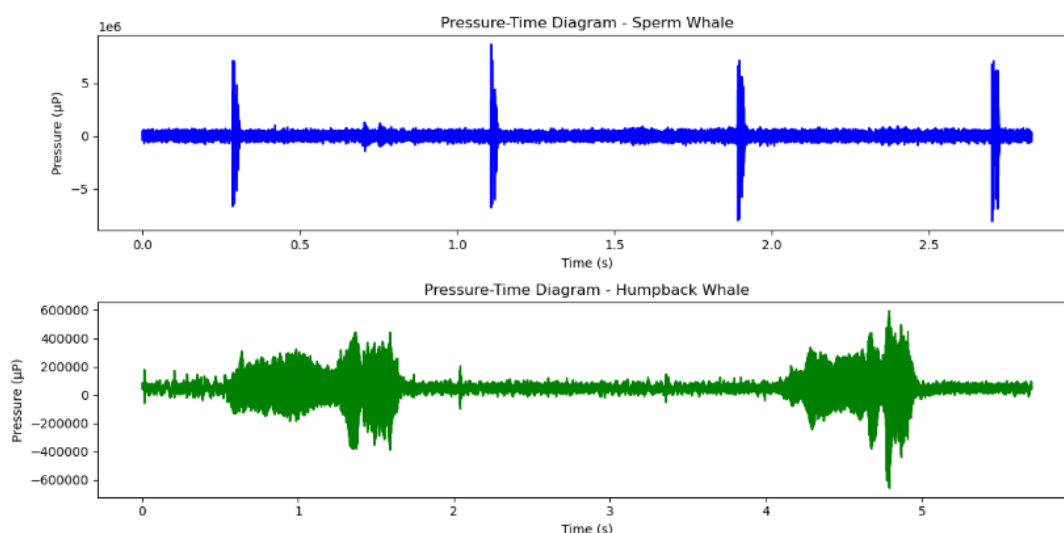




Παρατηρούμε πως για μεγάλο παράθυρο, χάνουμε σημαντική πληροφορία για την φασματική ροή. Η μεταβολή των τιμών των χαρακτηριστικών ανά το χρόνο δίνει πληροφορίες για τη δομή του σήματος και τις εναλλαγές στο σήμα. Στο φασματικό κέντρο, αύξηση των τιμών υποδηλώνει μεγαλύτερη ενέργεια στις υψηλότερες συχνότητες του σήματος κατά τη διάρκεια μιας περιόδου ή την εκκίνηση μιας ομιλίας. Αντίστοιχα, οι αλλαγές στη φασματική ροή υποδεικνύουν την αλλαγή στο ρυθμό των γεγονότων ή την επιβράδυνση ή επιτάχυνση του ρυθμού της ομιλίας.

Μέρος 3^ο -Ψηφιακή Επεξεργασία και Φασματική Ανάλυση Βιοσημάτων Κητωδών

3.1 Μετατρέπουμε το σήμα τάσης σε σήμα πίεσης χρησιμοποιώντας τα χαρακτηριστικά λειτουργίας του υδρόφωνου, δηλαδή τον συντελεστή ενίσχυσης $GF=0.16$ και την ευαισθησία του οργάνου $Sensitivity=-155$ dB relative to 1V/mPa. Σχεδιάζουμε το διάγραμμα Πίεσης-Χρόνου για τα δύο θηλαστικά



Παρατηρούμε ότι στην περίπτωση της Humpback Whale οι μεταβολές της πίεση είναι πιο σταδιακές και συνεχείς στο χρόνο, ενώ στην Sperm Whale παρουσιάζονται απότομες αυξήσεις.

3.2 Υπολογίζουμε τη μέση τετραγωνική ρίζα (rms) της πίεσης με βάση τη σχέση:

$$p_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N pressure_i^2},$$

Και το επίπεδο ηχητικής πίεσης (Sound Pressure Level – SPL) με βάση τη σχέση:

$$SPL_{RMS} = 20 \log_{10} \frac{p_{rms}}{p_{ref}},$$

Προκύπτουν τα εξής:

Sperm Whale:

RMS Pressure: 336416.18 μPa

SPL: 230.54 dB re 1μPa

Humpback Whale:

RMS Pressure: 83914.95 μPa

SPL: 218.48 dB re 1μPa

3.3 Υπολογίζουμε ξεχωριστά τα δύο σκέλη της παρακάτω εξίσωσης που περιγράφει το θεώρημα του Parseval για την Sperm Whale:

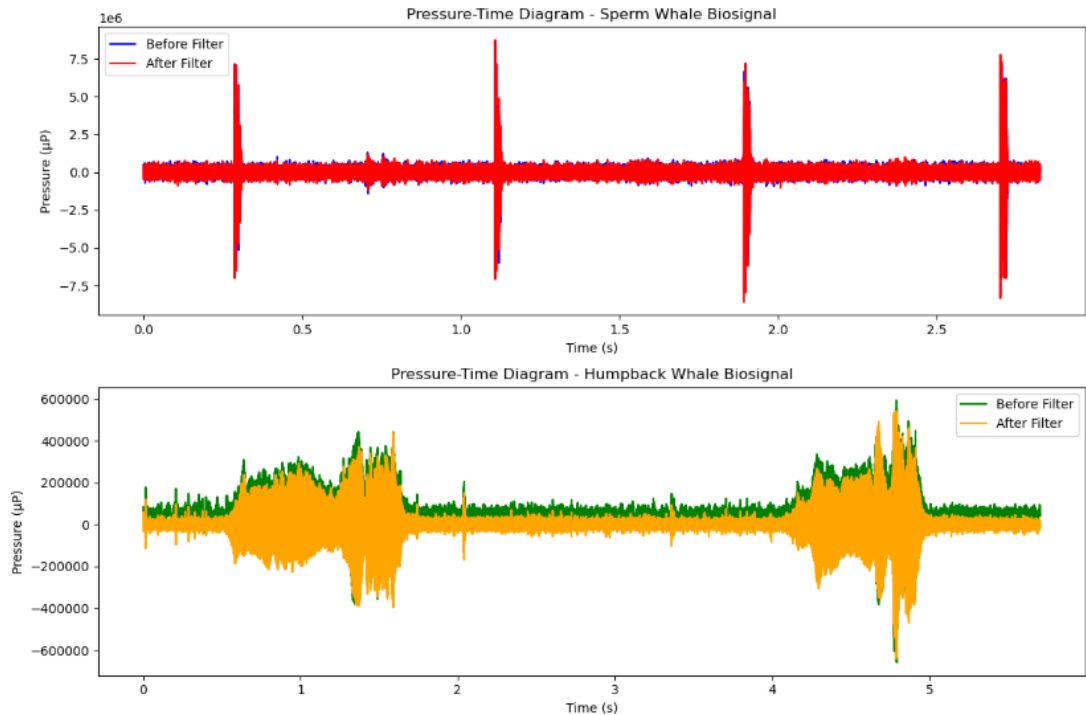
$$\sum_{n=0}^{N-1} |p[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |P[k]|^2,$$

Παρατηρούμε ότι οι δύο τιμές είναι πολύ κοντά και άρα το θεώρημα επιβεβαιώνεται.

Energy in the time domain: 2.6206093e+16

Energy in the frequency domain: 2.6206095797354132e+16

3.4 Εφαρμόζουμε στα δύο βιοσήματα ένα ζωνοπερατό Butterworth φίλτρο, τάξης N=3 με το οποίο να απομονώνουμε τις συχνότητες μεγαλύτερες από 200 Hz. Σχεδιάστε για κάθε βιοσήμα σε κοινό διάγραμμα το γράφημα Πίεσης-Χρόνου και το γράφημα του σήματος Πίεσης-Χρόνου μετά την εφαρμογή του Butterworth φίλτρου και προκύπτουν τα εξής:

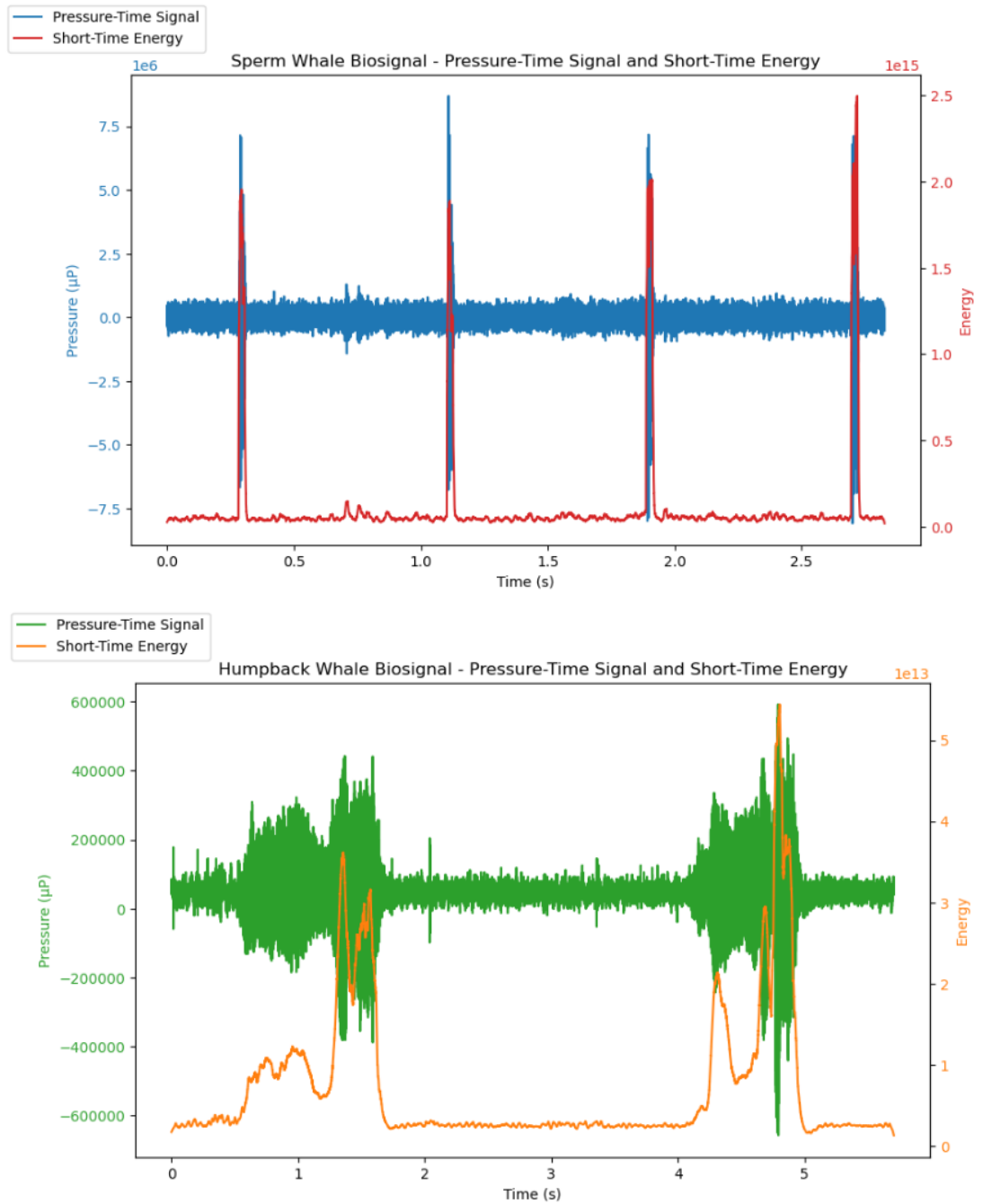


Παρατηρούμε ότι στην περίπτωση της Sperm Whale το σήμα μετά την εφαρμογή του φίλτρου δεν παρουσιάζει πολύ μεγάλες αποκλίσεις σε σχέση με το αρχικό, ενώ στην Humpback Whale φαίνεται να έχουν αποκοπεί περισσότερες συχνότητες.

3.5 Υπολογίζουμε το επίπεδο ηχητικής πίεσης SPL του αρχείου Pile driving.wav και βρίσκουμε ότι υπερβαίνει τα 100dB και άρα επηρεάζει τη συμπεριφορά των κητωδών.

The Sound Pressure Level of the 'Pile driving.wav' file (235.15 dB) exceeds the threshold level (100 dB) and may affect cetacean behavior.

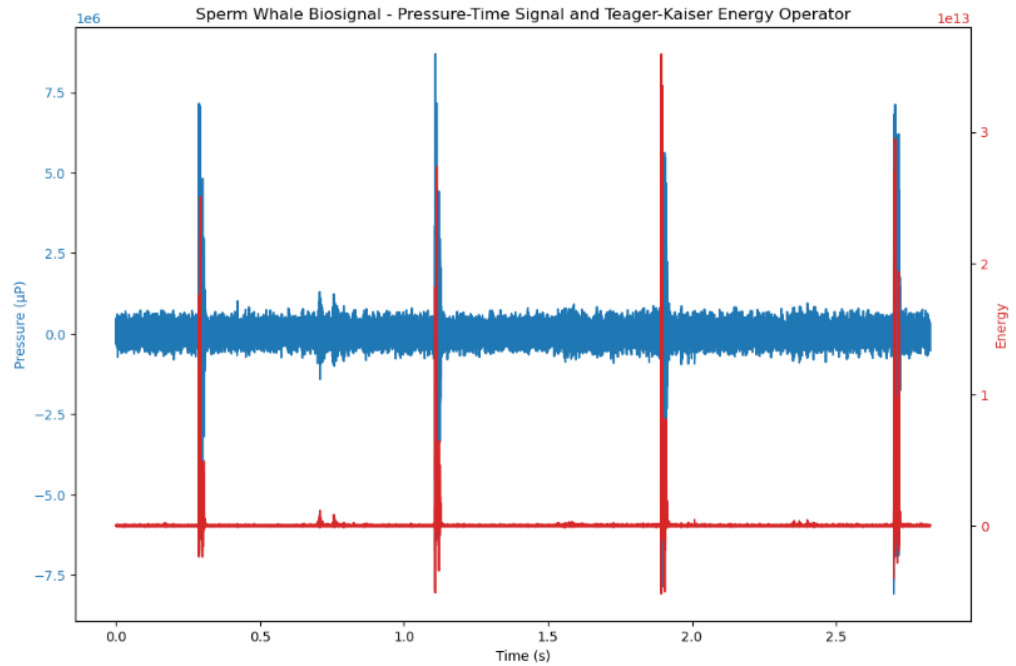
3.6 Υπολογίζουμε την ενέργεια βραχέος χρόνου του αρχικού σήματος ΠίεσηςΧρόνου και σχεδιάζουμε σε κοινό διάγραμμα το αρχικό σήμα και το σήμα ενέργειας και για τα δύο βιοσήματα.



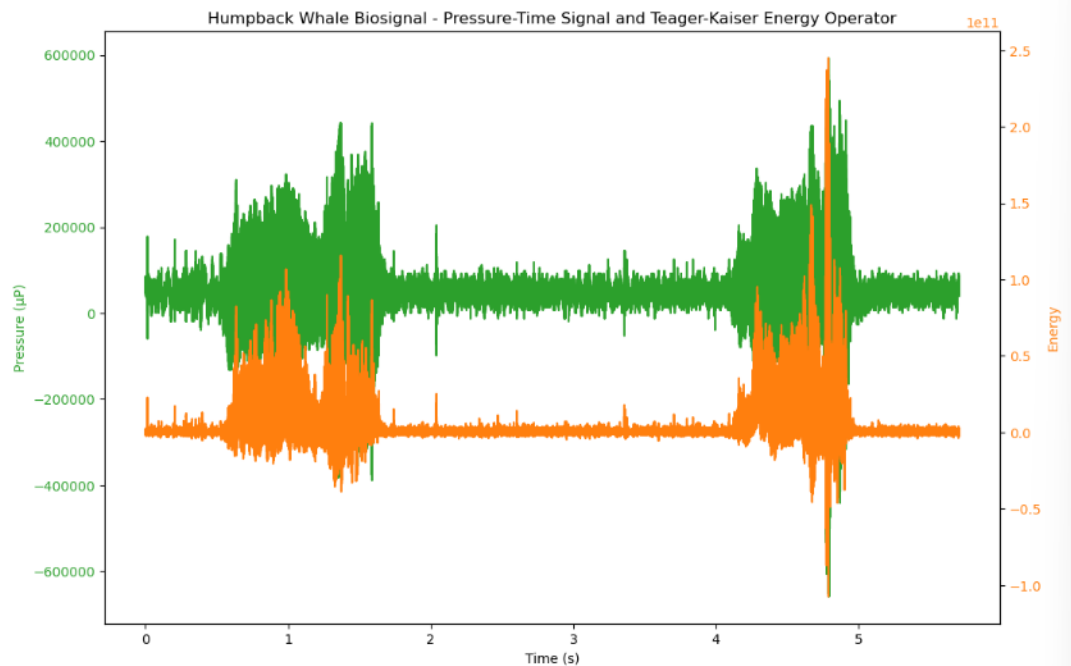
Παρατηρούμε ότι στα σημεία υψηλής πίεσης αντιστοιχεί και υψηλή ενέργεια. Γενικά οι διακυμάνσεις του σήματος ενέργειας ακολουθούν τις διακυμάνσεις του αρχικού σήματος πίεσης-χρόνου.

3.7 Σχεδιάζουμε σε κοινό διάγραμμα το αρχικό σήμα πίεσης χρόνου και την ενέργεια Teager για κάθε ένα από τα δύο βιοσήματα. Παρατηρούμε ότι η ενέργεια Teager έχει παρόμοια μορφή με την ενέργεια βραχέος χρόνου και εμφανίζουν αντίστοιχα peaks και αυξομειώσεις. Συμπεριλαμβάνει περισσότερες μη γραμμικές διακυμάνσεις ενώ το σήμα της ενέργειας βραχέος χρόνου είναι πιο ομαλό.

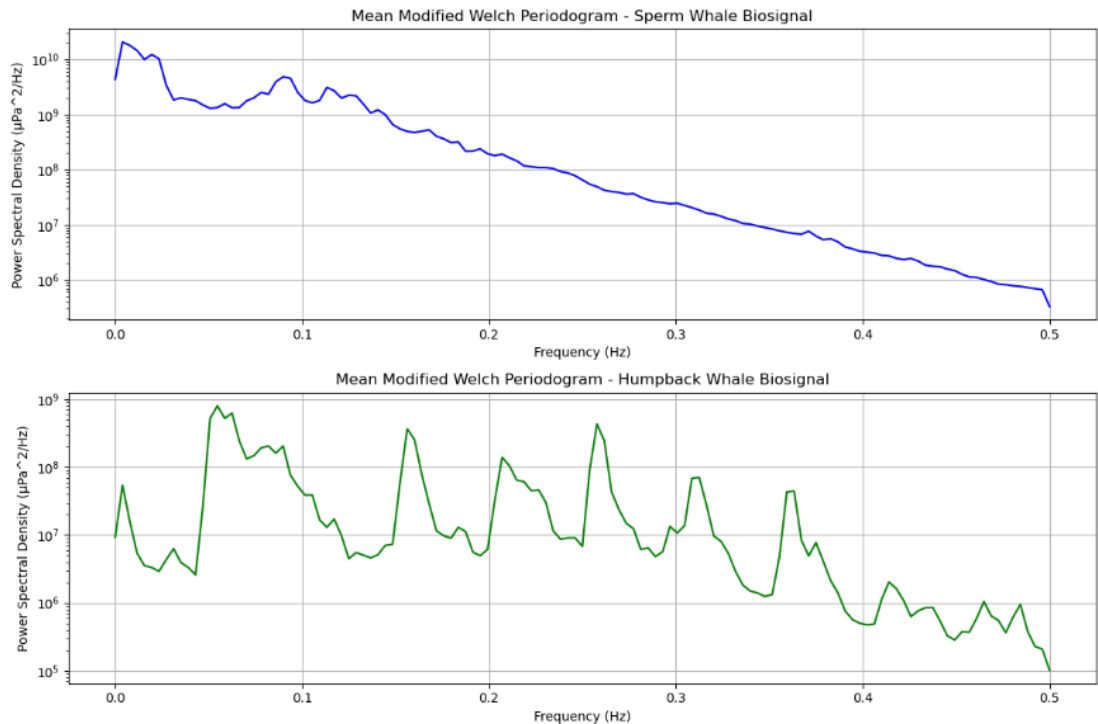
— Pressure-Time Signal
— Teager-Kaiser Energy Operator



— Pressure-Time Signal
— Teager-Kaiser Energy Operator

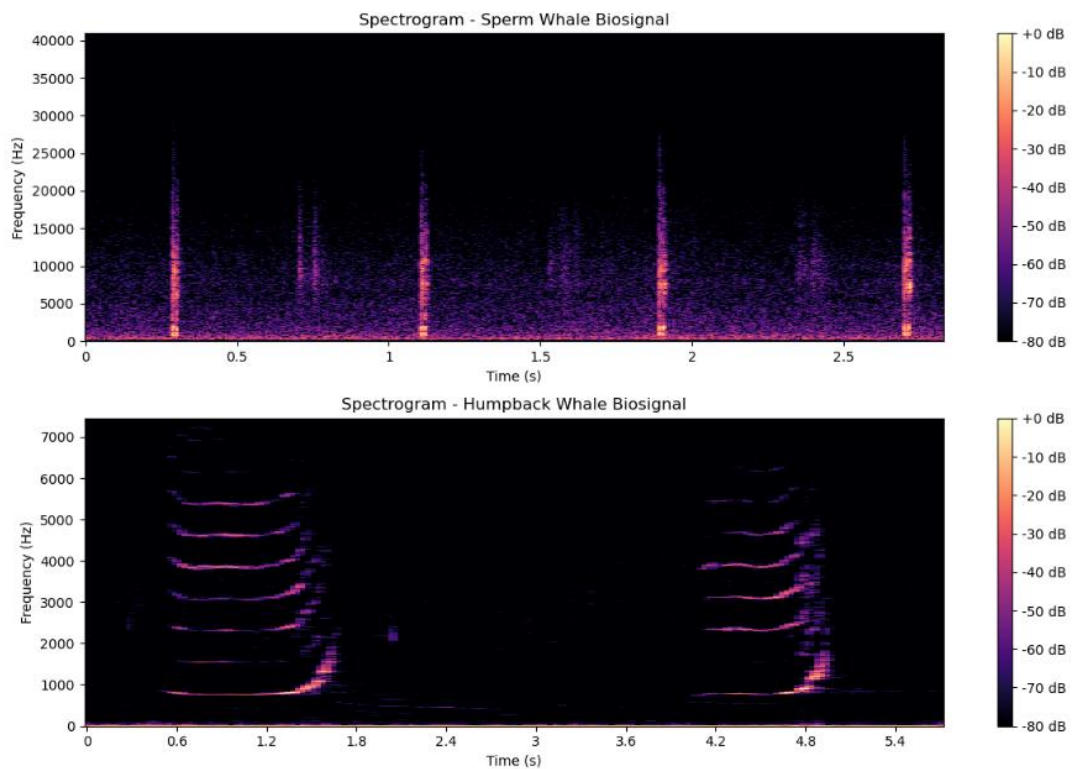


3.8 Σχεδιάζουμε το περιοδόγραμμα Welch σε λογαριθμική κλίμακα:



Η φασματική πυκνότητα ισχύος (PSD) αντιπροσωπεύει τον τρόπο με τον οποίο η ενέργεια του σήματος κατανέμεται σε όλο το φάσμα συχνοτήτων. Οι υψηλότερες τιμές PSD υποδεικνύουν μεγαλύτερη ενέργεια σε αυτές τις συχνότητες. Επίσης υψηλότερα επίπεδα ενέργειας σε ορισμένες συχνότητες υποδηλώνουν ισχυρότερα στοιχεία σήματος σε αυτές τις συχνότητες.

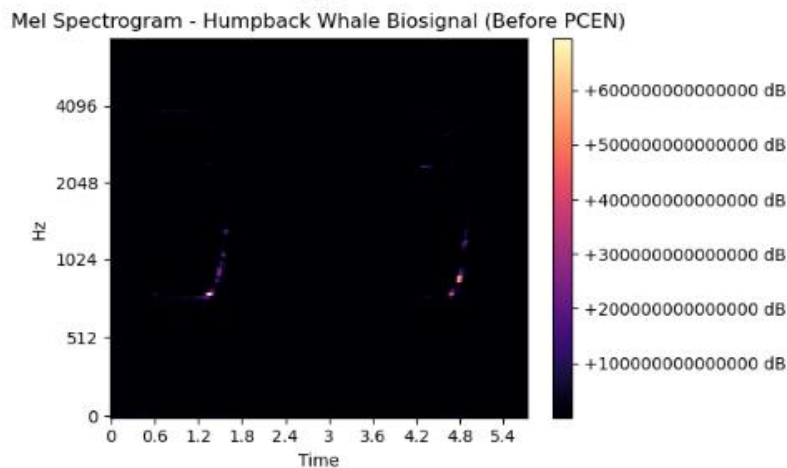
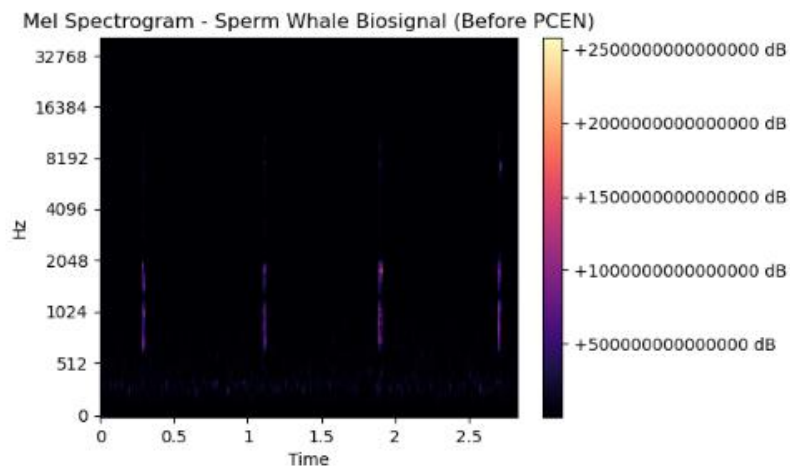
3.9 Σχεδιάζουμε το φασματογράφημα των δύο σημάτων:



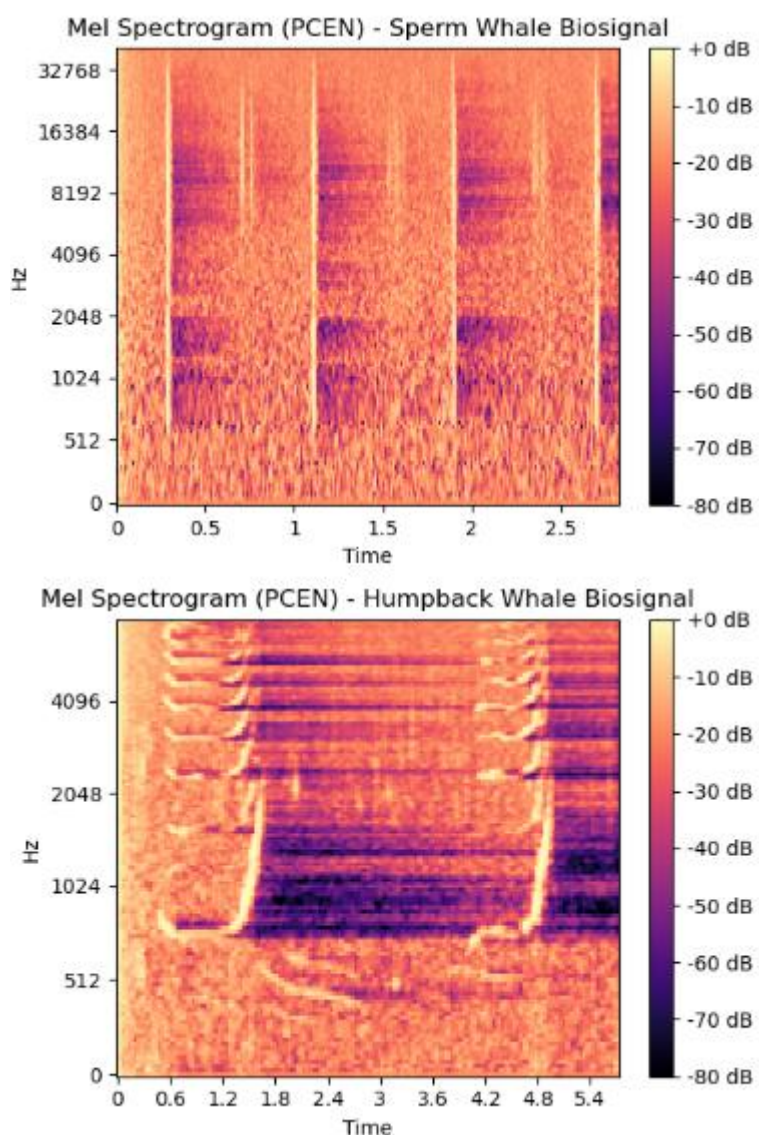
Τα φασματογράμματα αποτελούν μια απεικόνιση του περιεχομένου συχνότητας των βιοσημάτων με την πάροδο του χρόνου. Η ένταση χρώματος σε διαφορετικές συχνότητες υποδεικνύει τη σχετική ενέργεια αυτών των συχνοτήτων σε κάθε χρονικό σημείο. Οι περιοχές με χρώματα υψηλότερης έντασης αντιπροσωπεύουν υψηλότερα επίπεδα ενέργειας σε αυτές τις συχνότητες. Παρατηρούμε ότι οι περιοχές στις οποίες τα παραπάνω σχήματα έχουν πιο έντονα χρώματα και απεικονίζουν άρα μεγαλύτερη ενέργεια, ταυτίζονται χρονικά με τις αντίστοιχες περιοχές στις οποίες παρουσιάζεται υψηλή ενέργεια στα παραπάνω σχήματα των ερωτημάτων 3.7,3.8.

3.10 Σχεδιάζουμε τα mel-φασματογραφήματα για τα δύο σήματα.

Πριν την εφαρμογή της μεθόδου PCEN:



Μετά την εφαρμογή της μεθόδου PCEN:



Μετά την εφαρμογή του PCEN, η αντίθεση στα φασματογράμματα βελτιώνεται σημαντικά. Οι περιοχές με χαμηλή ενέργεια, οι οποίες μπορεί να εμφανίζονται αδιάκριτες ή καλυμμένες στα φασματογράμματα πριν από την εφαρμογή της μεθόδου PCEN, γίνονται πιο έντονες και πιο ευδιάκριτες μετά το PCEN. Το PCEN βοηθά επίσης στην ομαλοποίηση των επιπέδων ενέργειας σε διαφορετικές ζώνες συχνότητας και χρονικά πλαίσια.