

Ψηφιακή Επεξεργασία Σημάτων

2η Εργαστηριακή Άσκηση

Παπαρρηγόπουλος Θοδωρής el18040 paparrigopoulosthodoris@gmail.com

Φιλιππόπουλος Ορφέας el18082 orfeasfil2000@gmail.com

Η εργασία συντάχθηκε σε python3.6 και σε ubuntu 16.04. **Τα ηχητικά δεν ακούγονταν σωστά σε windows.**

Μέρος 1ο

Υπολογίζουμε και σχεδιάζουμε το μέτρο του delay-and-sum beam pattern μέσω του τύπου:

$$B(\omega, \theta) = \frac{1}{N} \frac{\sin[\frac{N}{2} \frac{\omega}{c} d (\cos \theta - \cos \theta_s)]}{\sin[\frac{1}{2} \frac{\omega}{c} d (\cos \theta - \cos \theta_s)]}$$

$f=2\text{kHz}$

και $\theta_s = \pi/2 \text{ rad}$

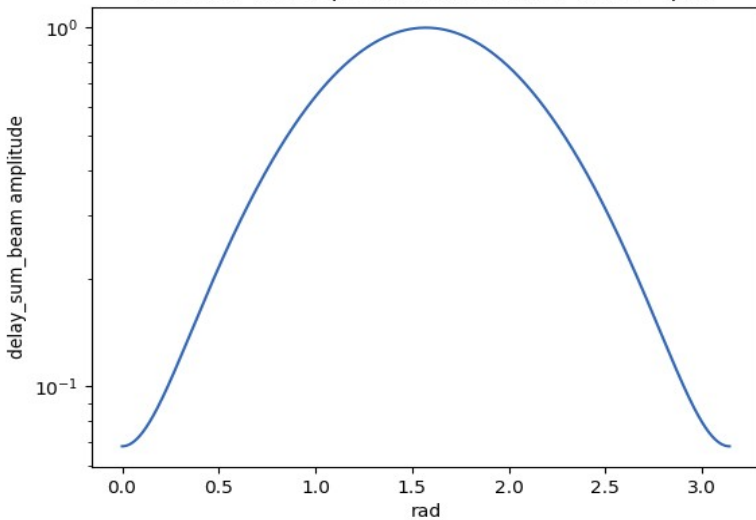
$\lambda = 340/2000 = 0.17\text{m}$

1)

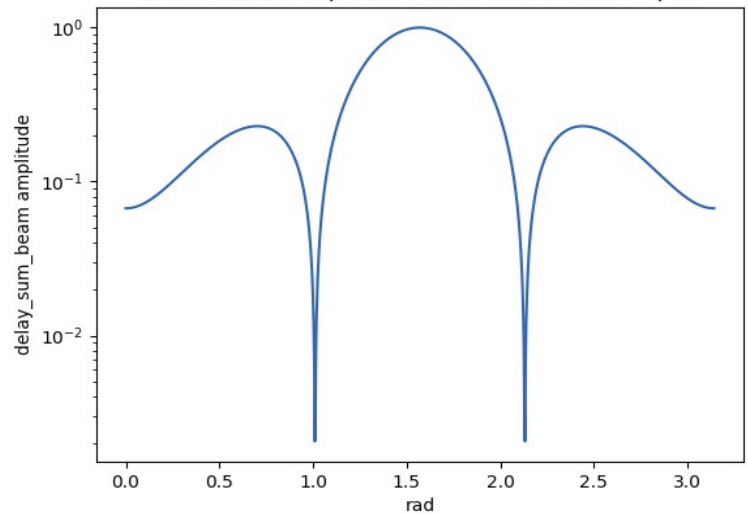
Για $d = 4\text{cm}$ (ικανοποιείται πάντα το θεώρημα της χωρικής δειγματοληψίας) και

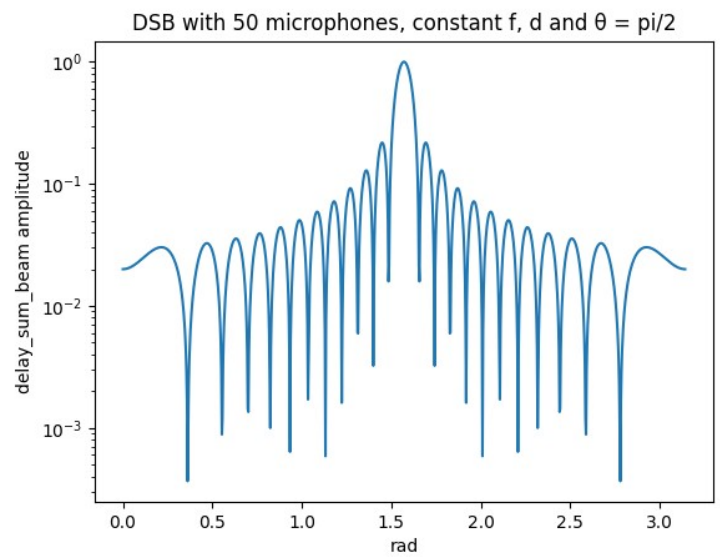
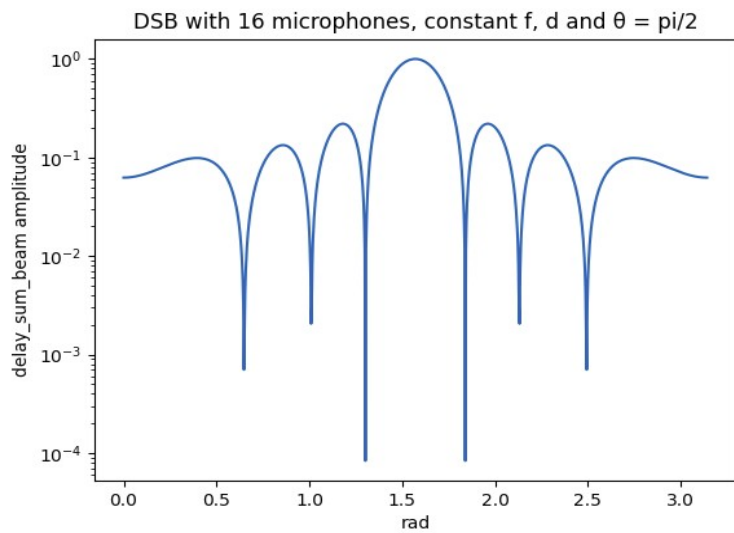
DSB = Delay-and-Sum Beamformer

DSB with 4 microphones, constant f, d and $\theta = \pi/2$



DSB with 8 microphones, constant f, d and $\theta = \pi/2$

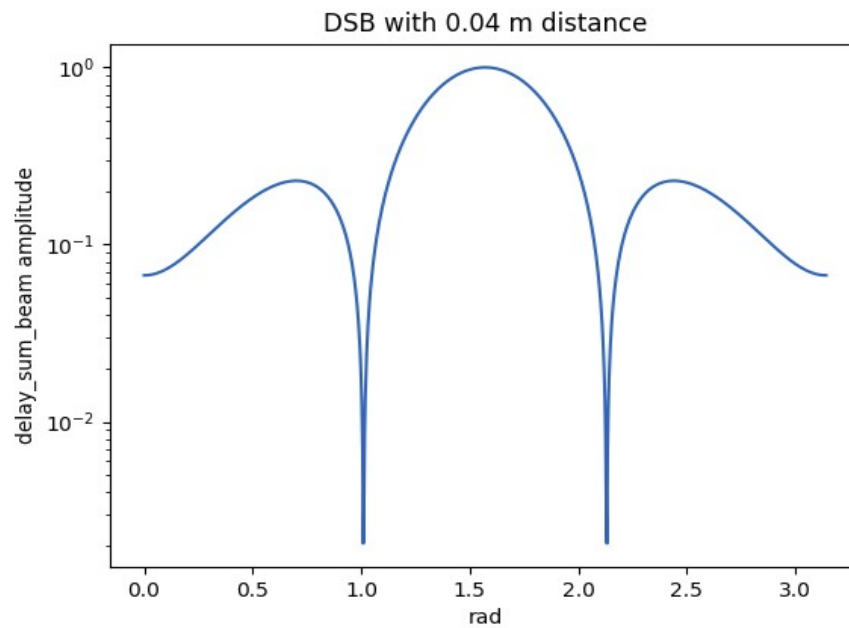


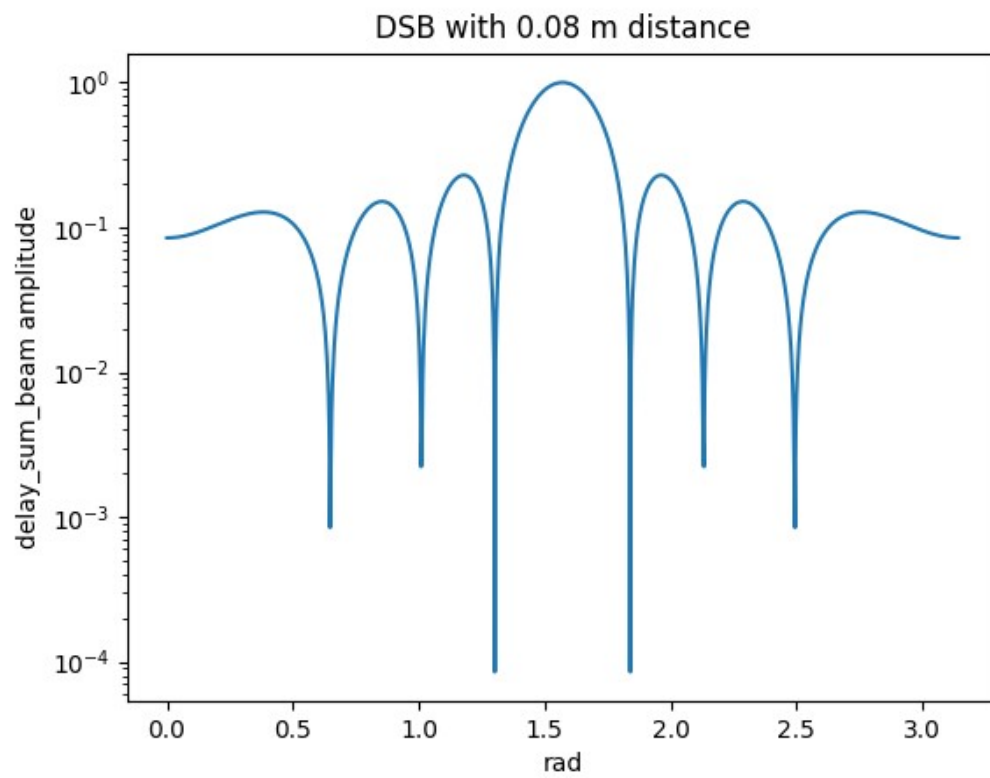
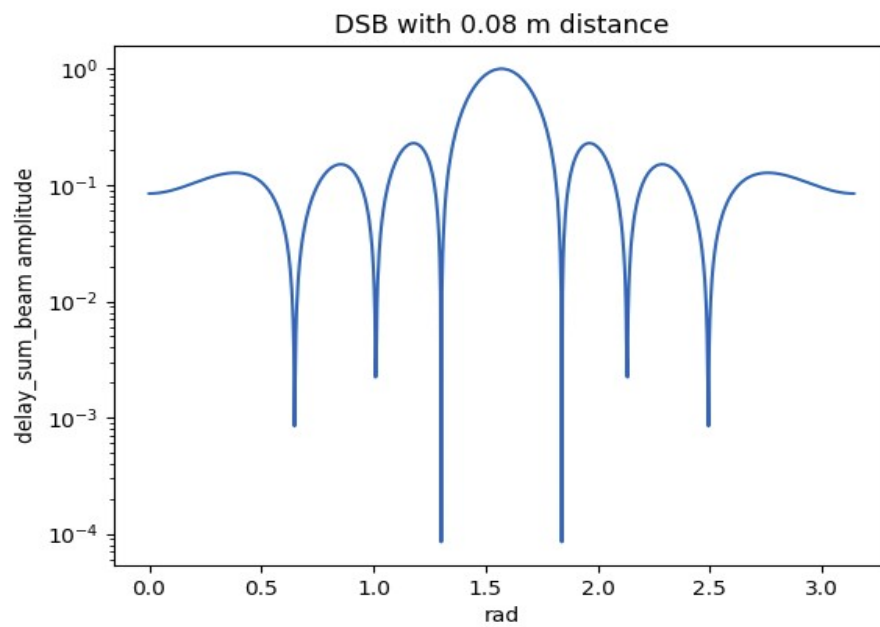


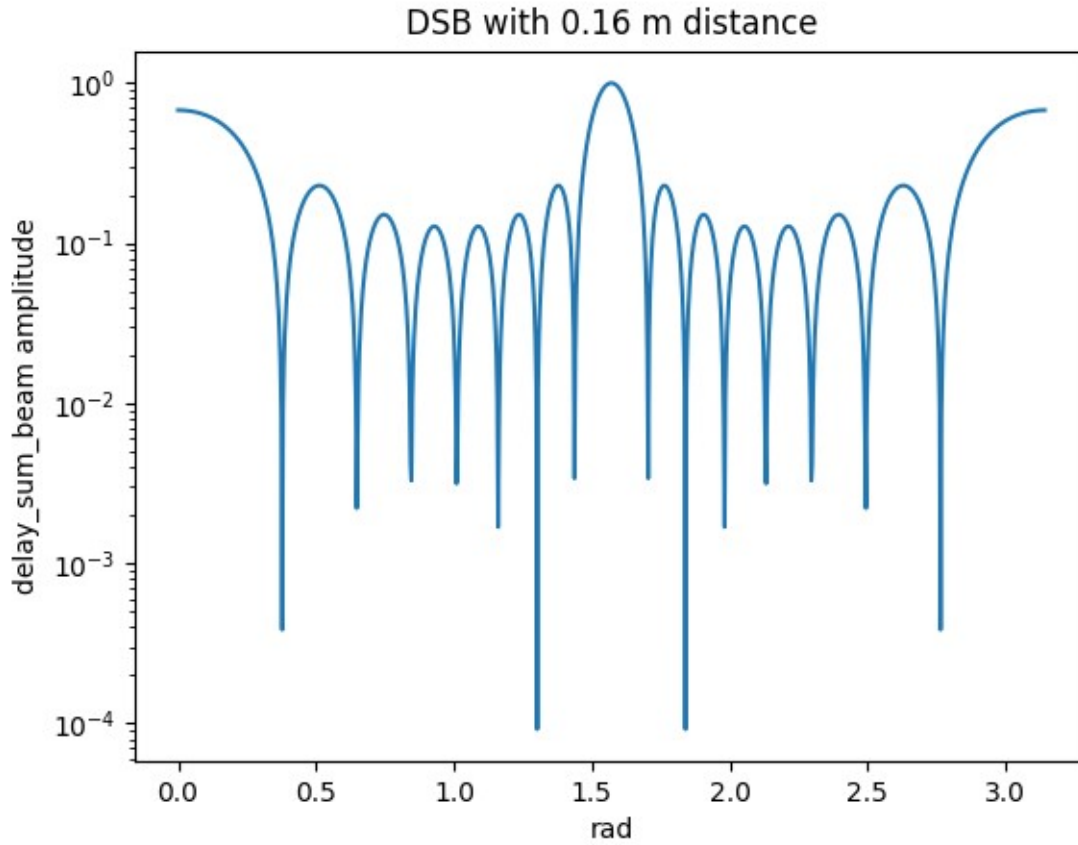
Παρατηρούμε ότι για μεγαλύτερες τιμές του N αυξάνονται και οι πλευρικοί λοβοί του μέτρου του delay-and-sum -beamformer pattern.

2)

Για $N = 8$







Αρχικά παρατηρούμε ότι αυξάνεται το d αυξάνονται και οι πλευρικοί λοβοί.

Επιπροσθέτως, πρέπει να ικανοποιείται η:
(χωρικό θεώρημα δειγματοληψίας)
προκειμένου να λειτουργήσει ορθά το
delay-and-sum beamformer pattern.

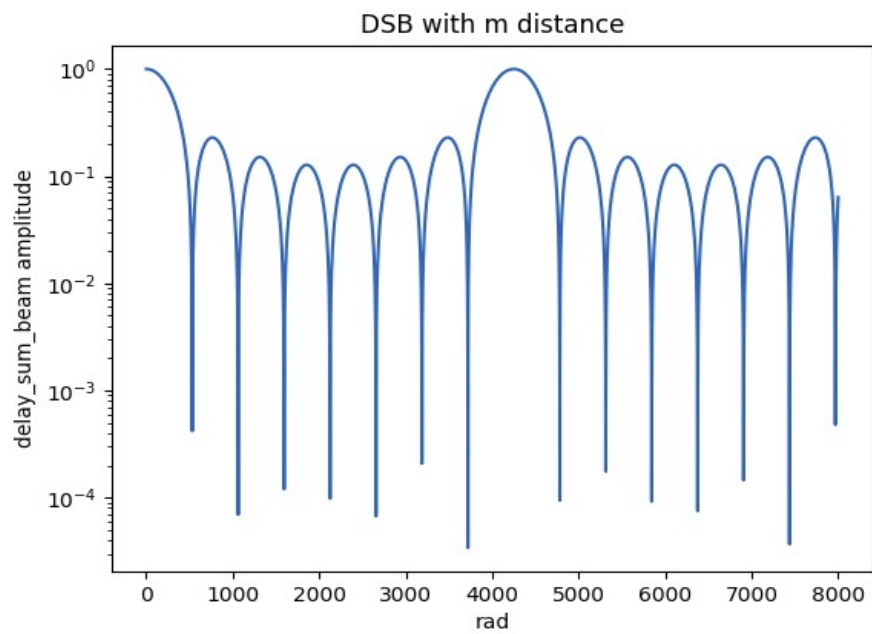
$$d < \frac{\lambda}{1 + |\cos \theta_s|}$$

Εδώ για $\theta_s = \pi/2$ rad πρέπει $d < \lambda \Rightarrow d < 0.17\text{m}$ που ικανοποιείται σε κάθε περίπτωση. Βέβαια στη τελευταία ικανοποιείται οριακά (για 0.01m διαφορά) με αποτέλεσμα να παρουσιάζονται 2 αρκετά υψηλή (σε σχέση με τον κεντρικό) λοβοί, γεγονός που δεν είναι επιθυμητό.

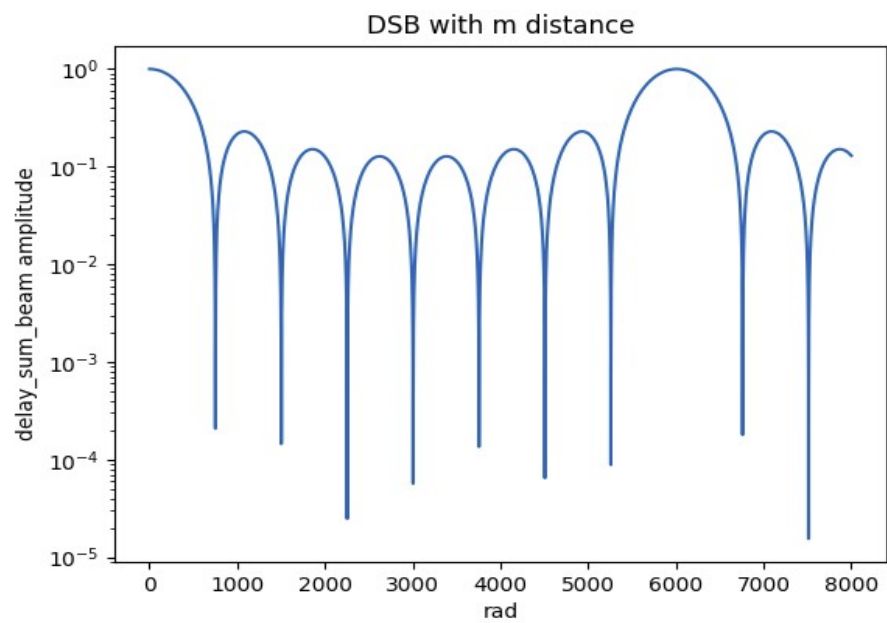
3)

Για $N = 8$, $d = 0.08\text{m}$, $\theta_s = \pi/2$ rad

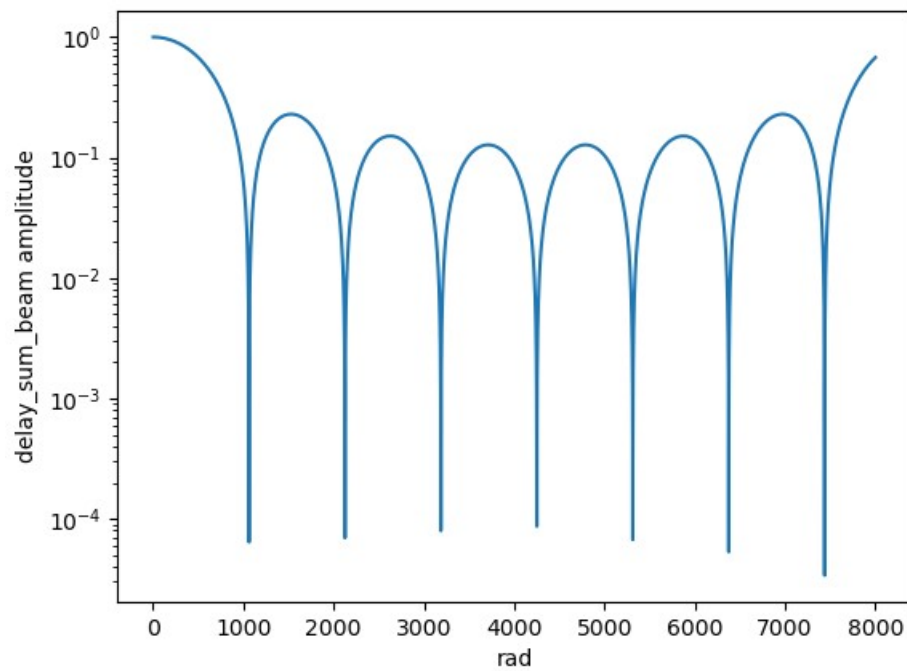
και για $\theta = 0$ rad:



για $\theta = \pi/4$ rad:



και για $\theta = \pi/2$ rad:



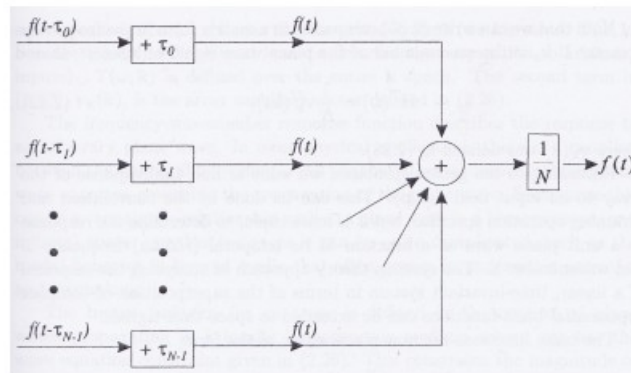
Παρατηρούμε ότι όσο αυξάνεται η γωνία θ αυξάνεται το scale του μέτρου του delay-and-sum beamformer pattern καθώς τόσο “απλώνονται” οι λοβοί όσο και το ότι το μέγιστο που έχουμε αρχικά στα 4100Hz μετατοπίζεται προς τα δεξιά.

Μέρος 2ο

2.1.A)

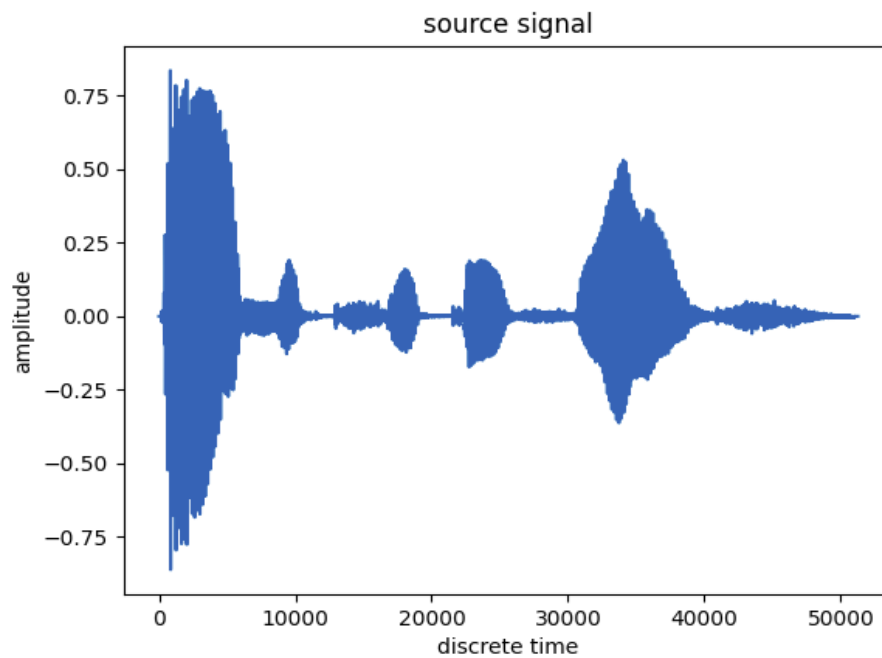
1)

Ακολουθούμε (σε επίπεδο κώδικα) τη διαδικασία που απεικονίζεται στο παρακάτω σχήμα:



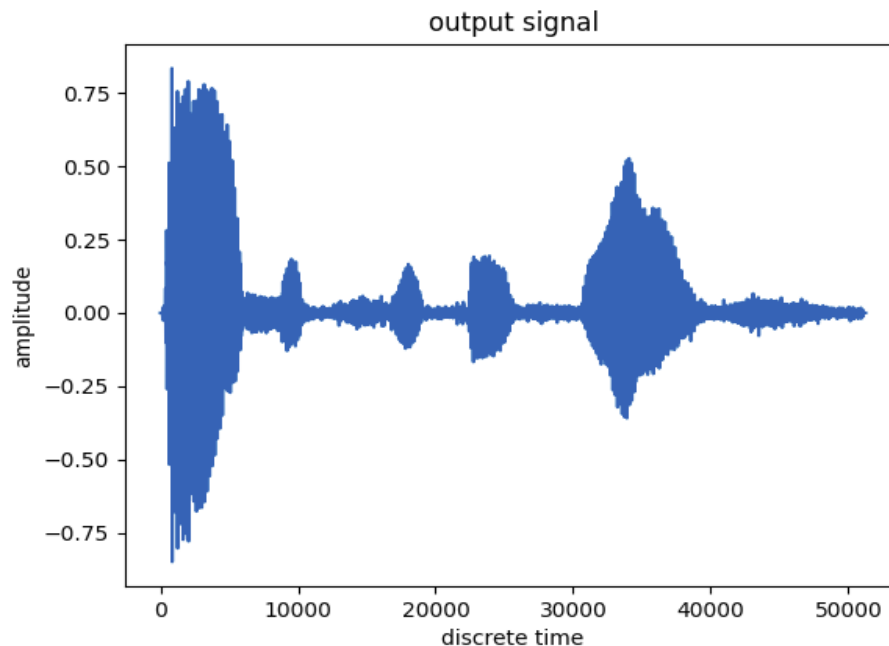
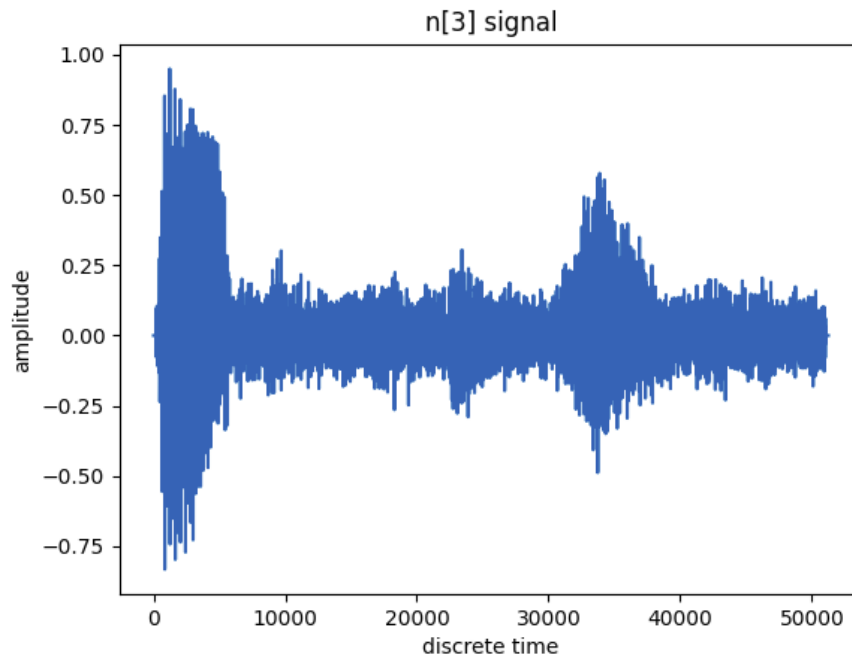
Σχήμα 3: Delay-and-sum beamformer

Το source signal είναι:



Το σήμα που λαμβάνει το κεντρικό μικρόφωνο είναι:

Η έξοδος του delay-and-sum beamformer είναι:



Παρατηρούμε ότι είναι αρκετά αποθρομβωμένο!

Επιπροσθέτως όταν αφαιρούμε από το output signal το source ακούγεται μόνο θόρυβος και άρα λειτουργεί ορθά, δηλαδή ευθυγραμμίστηκαν σωστά τα σήματα των μικροφώνων (βέβαια αυτό οφείλεται στο γεγονός ότι τόσο ικανοποιείται το χωρικό θεώρημα της δειγματοληψίας

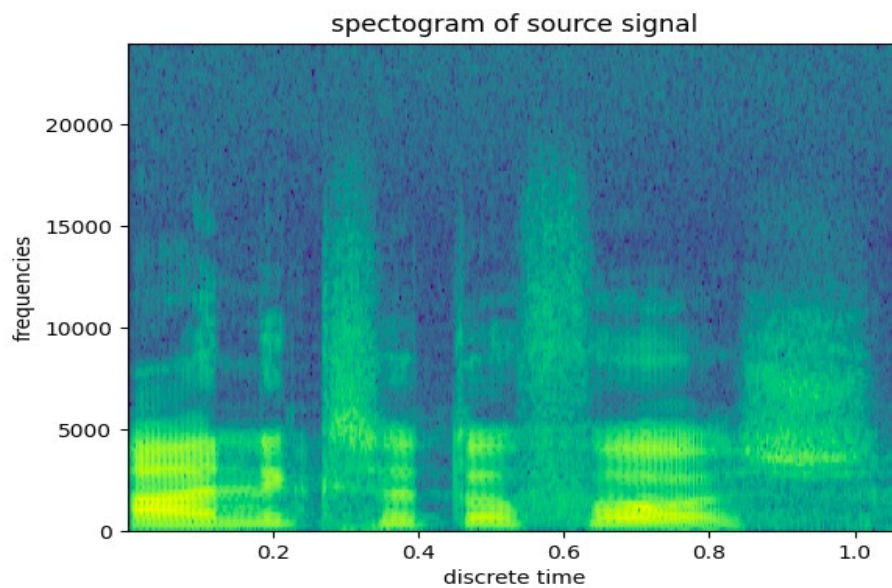
όσο ότι ο θόρυβος προέρχεται από μία συγκεκριμένη κατεύθυνση)!

2)

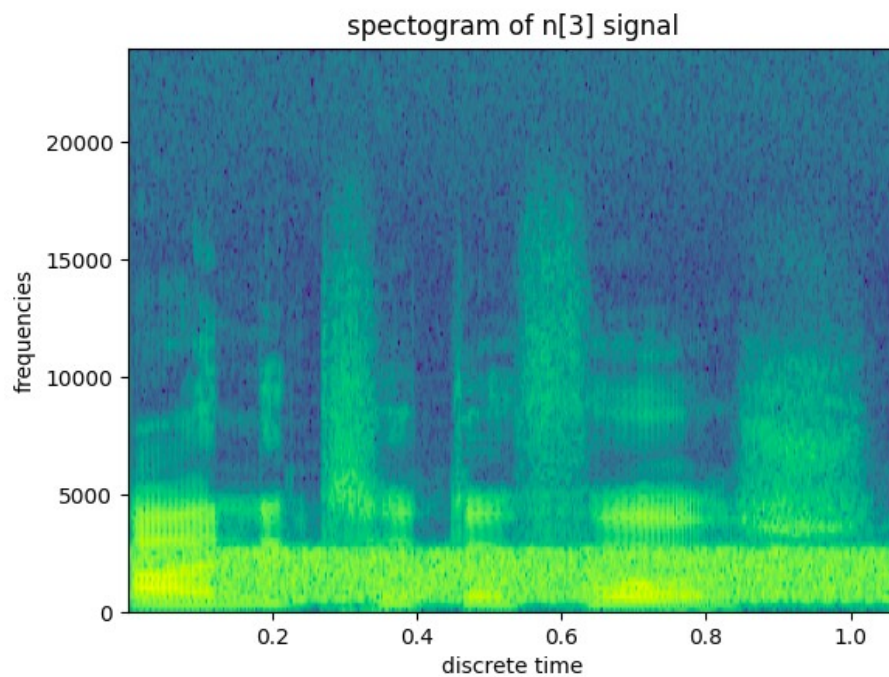
Οι κυματομορφές των σημάτων συγκρίθηκαν στο προηγούμενο ερώτημα.

Όσον αφορά τα σπεκτρογράφημα:

Το σπεκτρογράφημα του source signal είναι:

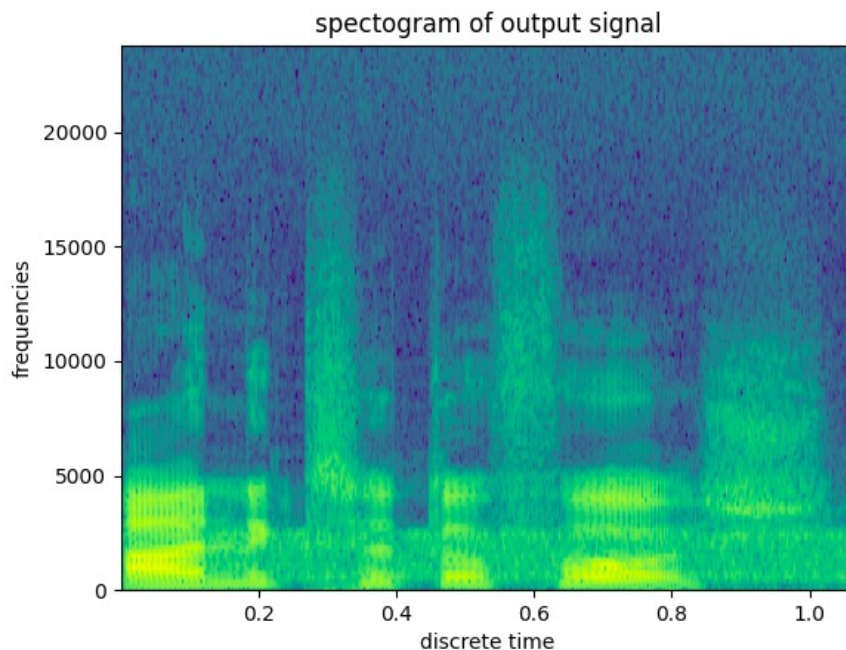


Το σπεκτρογράφημα του σήματος που λαμβάνει το 3ο μικροφώνο είναι:



Το

σπεκτρογράφημα της εξόδου του delay-and-sum beamformer είναι:



Είναι εμφανές ότι, σε σύγκριση με το σπεκτρογράφημα του αρχικού σήματος, στο σπεκτρογράφημα του μικροφώνου 3 εμφανίζονται σε όλο το χρονικό εύρος του συχνότητες που δεν υπάρχουν στο αρχικό που οφείλονται στο θόρυβο. Μάλιστα έχουν αρκετά έντονα πλάτη. Από την άλλη το σπεκτρογράφημα του σήματος της εξόδου του delay-and-sum beamformer έχει και αυτό συχνότητες που δεν εμφανίζονται στο αρχικό αλλά παρατηρούμε πως είναι αρκετά πιο ασθενή τα πλάτη τους γεγονός που μας υποδεικνύει την ορθότητα του κώδικα μας αλλά και την αποτελεσματικότητα του delay-and-sum beamformer.

3)

Όσον αφορά τα SNR's, το SNR του θορυβώδους σήματος βγαίνει: **2.9999999814792315**

Όσον αφορά τα SNR's, το SNR της εξόδου του delay-and-sum beamformer βγαίνει:

20.265981656481472

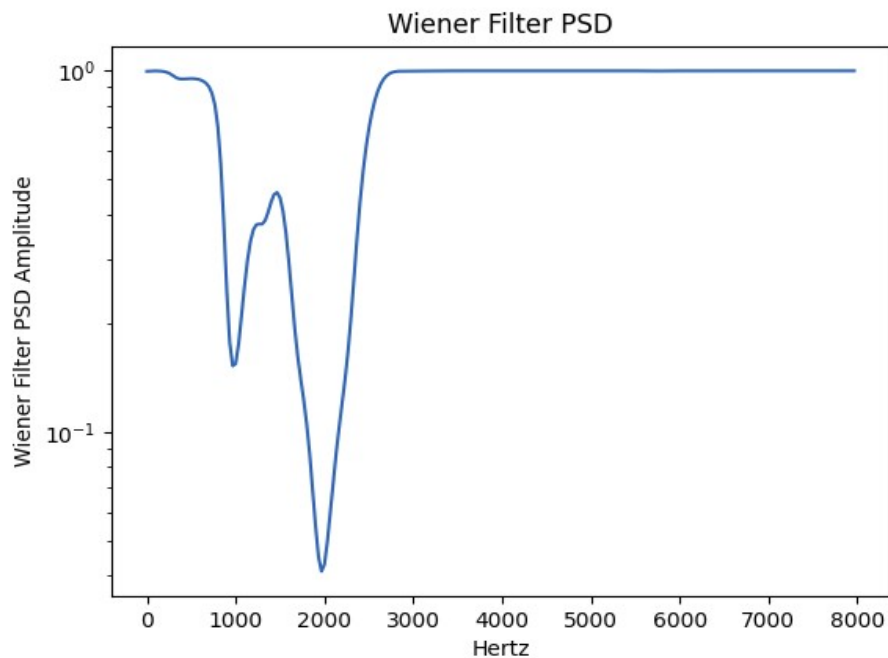
Για ακόμα μία φορά βλέπουμε πως επιτυγχάνεται αποθρομβοποίηση του σήματος και αυτή τη φορά με τα SNR's καθώς το SNR της εξόδου είναι μεγαλύτερο της εισόδου (μικρόφωνο 3).

2.1.B)

Αρχικά παίρνουμε τα ζητούμενα πλαίσια του source signal και του σήματος στο μικρόφωνο 3.

1) Με τη βοήθεια της welch υπολογίζουμε τα φάσματα ισχύος του θορύβου (= $n[3] - \text{original}$) και της σήματος εισόδου (μικρόφωνο 3) στη συνέχεια την απόκριση του Wiener φίλτρου από το δοσμένο τύπο.

Στη συνέχεια plotaroume (για συχνότητες από 0 έως 8000) την απόκριση σε λογαριθμική κλίμακα:



Παρατηρούμε ότι το φίλτρο παίρνει τιμές διάφορες του ενός για συχνότητες περίπου από 500-2.5kHz, γεγονός αναμενόμενο καθώς το φίλτρο Wiener προσπαθεί να φιλτράρει το θόρυβο ο οποίος είναι ζωνοπερατός στις συχνότητες.

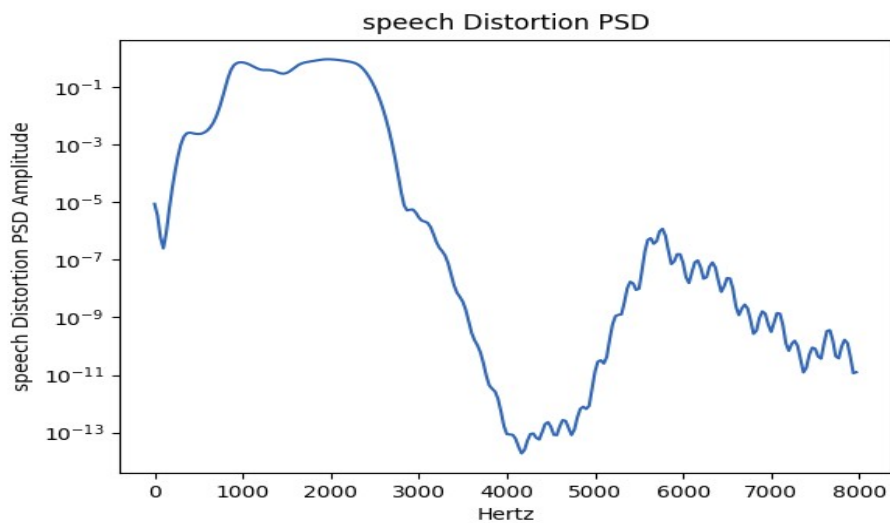
2)

Plotaroume σε λογαριθμική κλίμακα το speech distortion index με που το υπολογίζουμε με βάση τον τύπο:

$$n_{sd}(\omega) = \frac{E[|S(\omega) - H_W(\omega)S(\omega)|^2]}{P_x(\omega)} = \boxed{|1 - H_W(\omega)|^2}$$

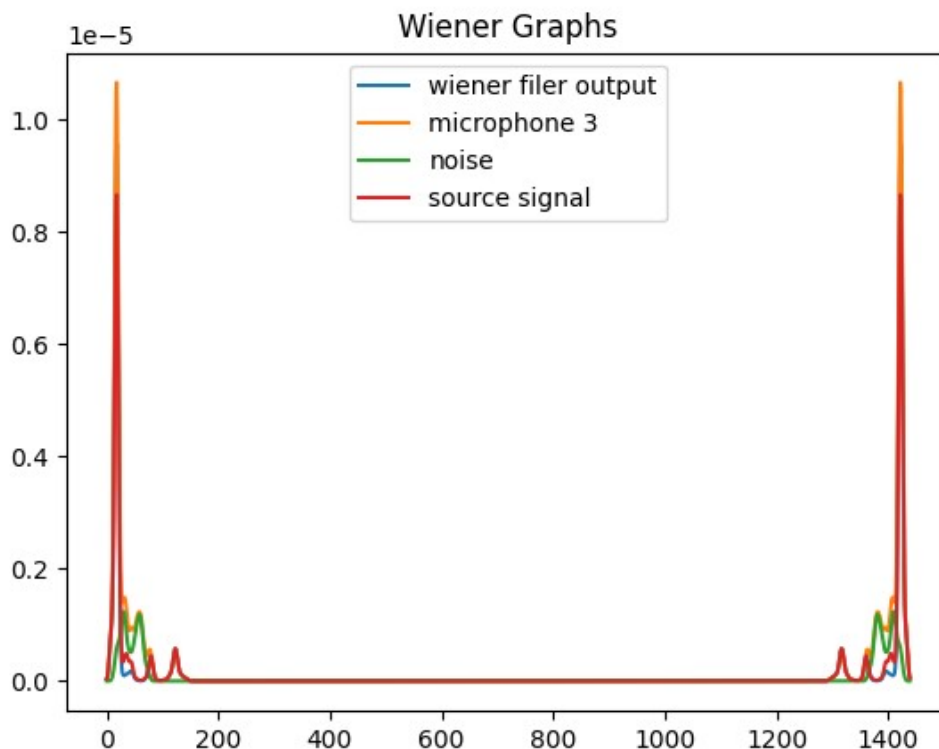
και παίρνουμε ως

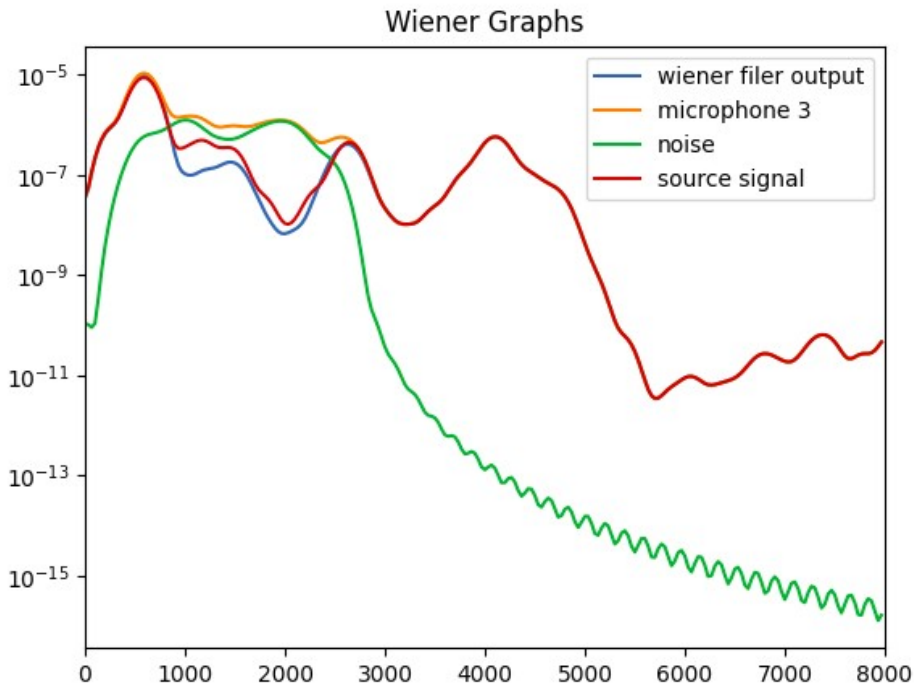
αποτέλεσμα:



Παρατηρούμε ότι το speech distortion index παίρνει μεγάλες τιμές για τις τιμές των συχνοτήτων περίπου 500-2.5kHz. Αυτό συμβαίνει διότι σε εκείνες τις συχνότητες έχουμε θόρυβο με αποτέλεσμα το φίλτρο Wiener να τις φιλτράρει (δηλαδή εκεί “επιδρά” το φίλτρο) γεγονός που προκαλεί παραμόρφωση και στο αρχικό σήμα φωνής.

3) Τα φάσματα ισχύος για τα ζητούμενα σήματα είναι:





Λογαριθμική
Κλίμακα

Για ακόμη μία φορά αναδεικνύεται η επιτυχία του φιλτραρίσματος καθώς το φάσμα ισχύος της εξόδου του wiener φίλτρου προσεγγίζει σε πολύ μεγαλύτερο βαθμό το φάσμα ισχύος του source signal.

4)

Τα ζητούμενα SNR's είναι:

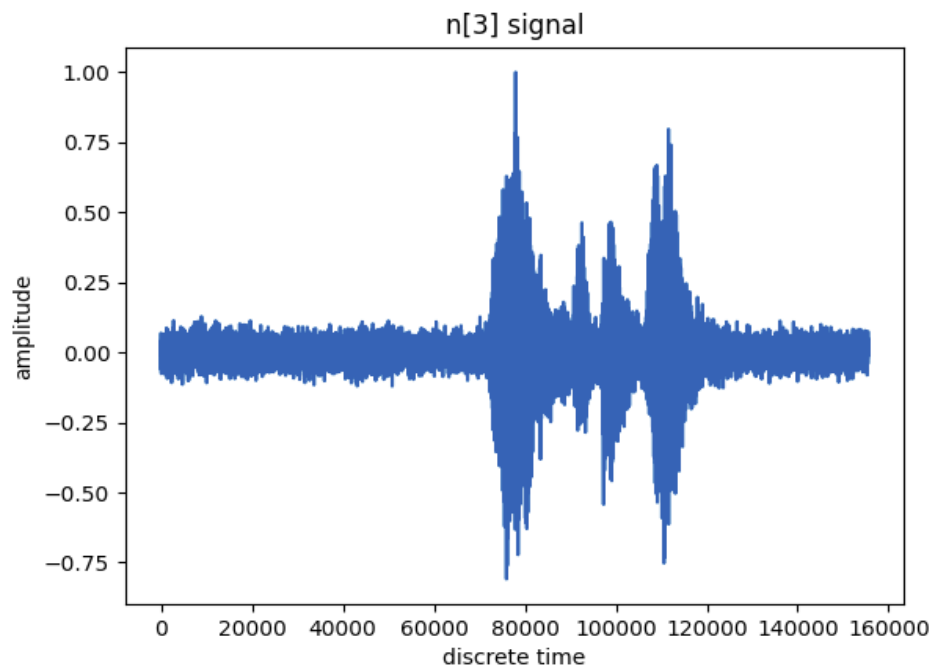
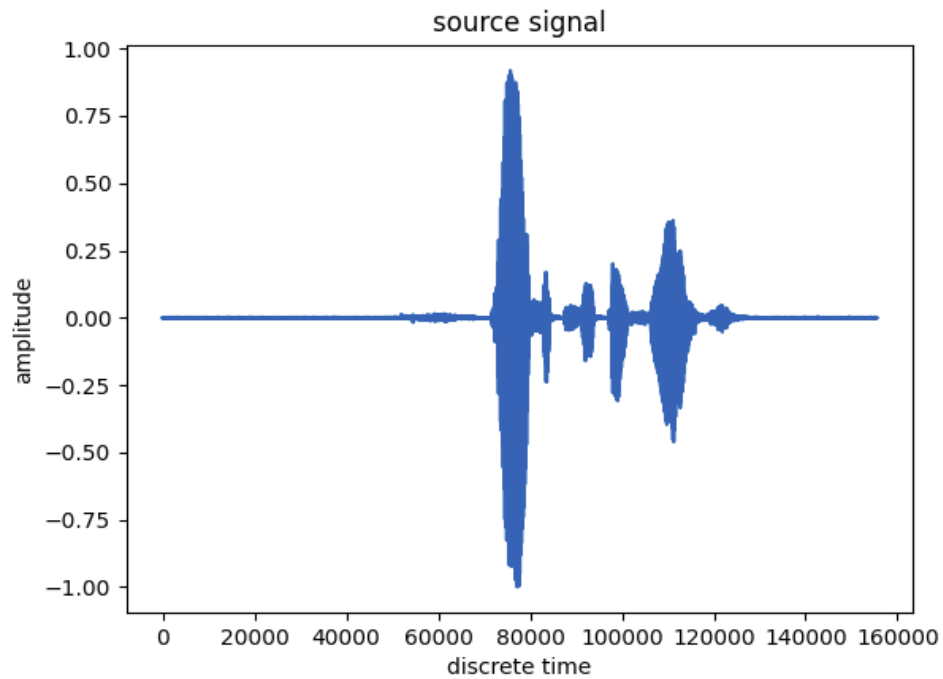
```
SNR_input:  3.1423070739391563
SNR_beam:   19.567487474433896
SNR_wiener_output: 9.301037052489049
```

Παρατηρούμε ότι ο beam_former στο συγκεκριμένο πρόβλημα (δηλαδή τα σήματα να έρχονται από συγκεκριμένες κατευθύνσεις) έχει το μεγαλύτερο SNR (=19.5dB). Το output του wiener φίλτρου και αυτό επιτυγχάνει αποθορυβοποίηση του σήματος αλλά με χαμηλότερο SNR(=9.3dB).

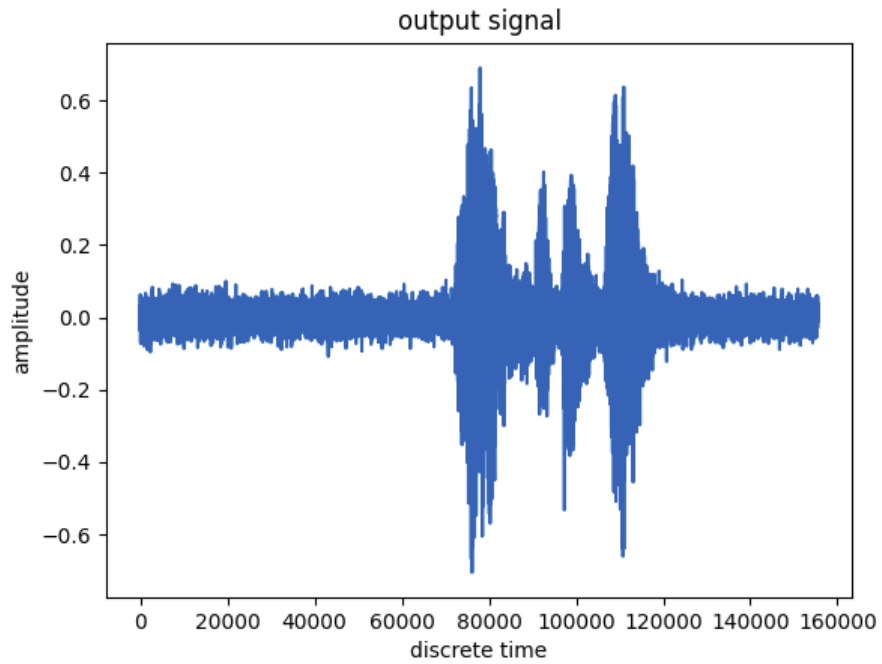
2.2.A)

1) Υπολογίζουμε, όπως και στο 2.1.A.1) ερώτημα, τα βάρη του delay-and-sum beamformer

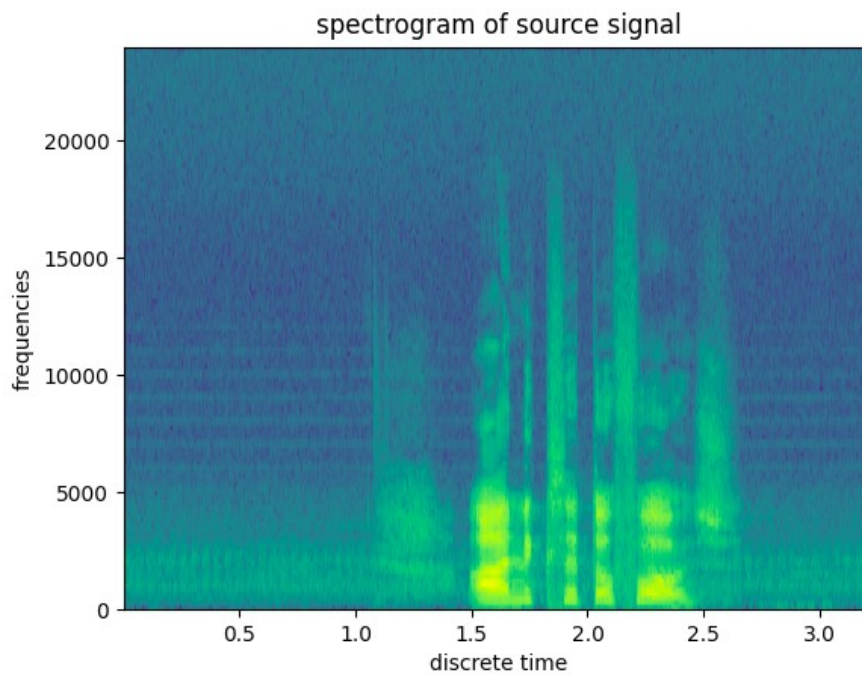
2) Οι κυματομορφές των σημάτων είναι:

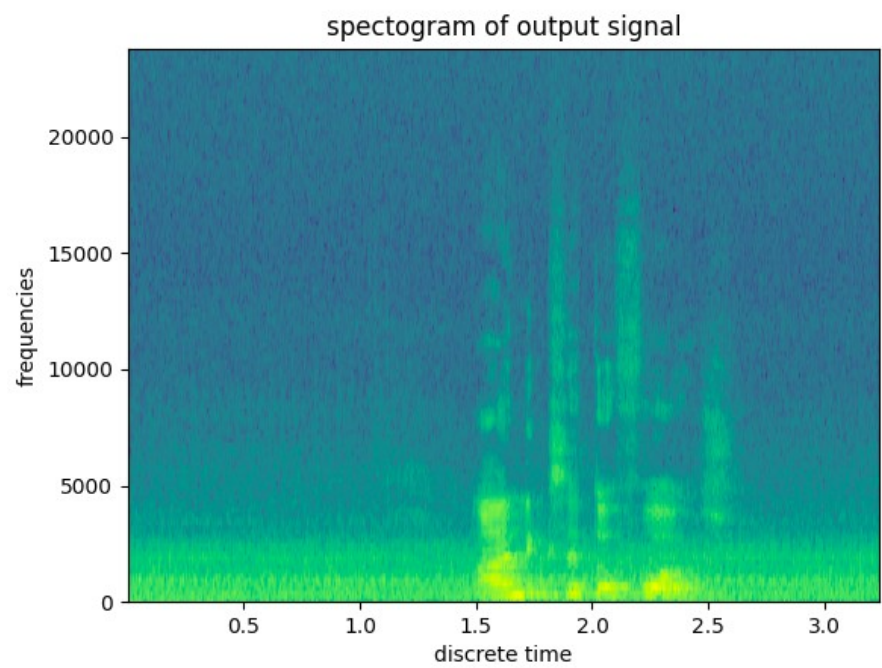
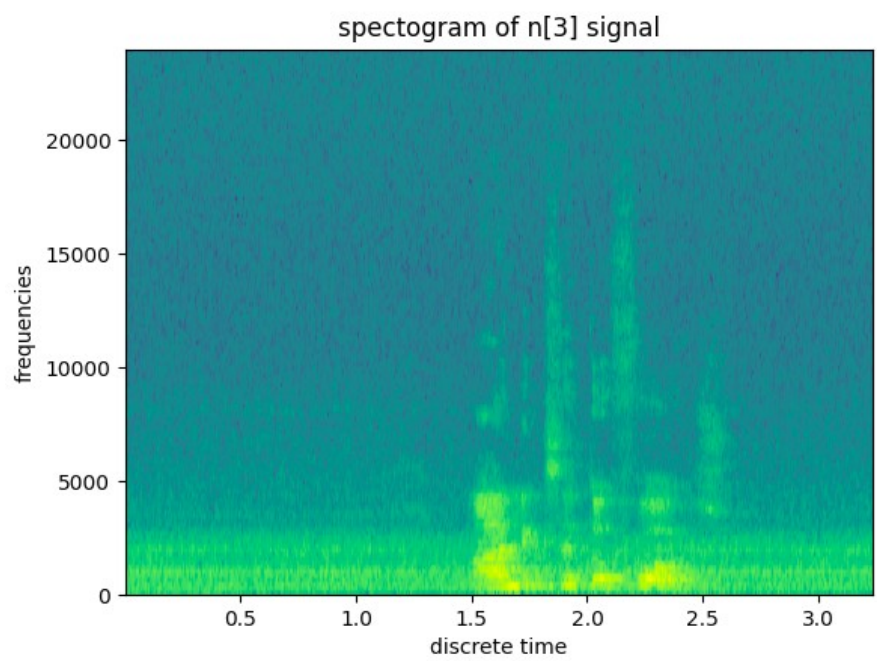


και



και τα αντίστοιχα σπεκτρογράφημα:





Παρατηρούμε ότι εάν και ο beamformer μειώνει στοιχειωδώς τον θόρυβο, δεν μπορεί να το πετύχει με μεγάλη αποτελεσματικότητα καθώς ο θόρυβος δεν προέρχεται από μία συγκεκριμένη πηγή με μία συγκεκριμένη κατεύθυνση (που τότε γίνεται αποτελεσματικός). Επιπροσθέτως η αδυναμία του beamformer στο να εξαλείψει το θόρυβο φαίνεται και από το γεγονός ότι όταν αφαιρούμε από την έξοδο του το original σήμα (και αποθηκεύσουμε το νέο αυτό σήμα σε .wav αρχείο) τότε ακούει κανείς πέρα από θόρυβο και φωνή και μάλιστα αρκετά έντονα.

3)

Τα SNR's της εξόδου του beamformer και της εξόδου (για $L = 1440$ (30ms) και κατώφλι -20dB) είναι:

```
SSNR_n_3: 1.869223064861179
SSNR_output: 2.0975179084721005
```

Παρατηρούμε ότι έχουμε πολύ μικρή βελτίωση του SSNR στην έξοδο του beamformer γεγονός που αναδεικνύει την αδυναμία του στο συγκεκριμένο πρόβλημα.

Για λόγους πληρότητας, για $L = 5000$ και κατώφλι -10dB έχουμε:

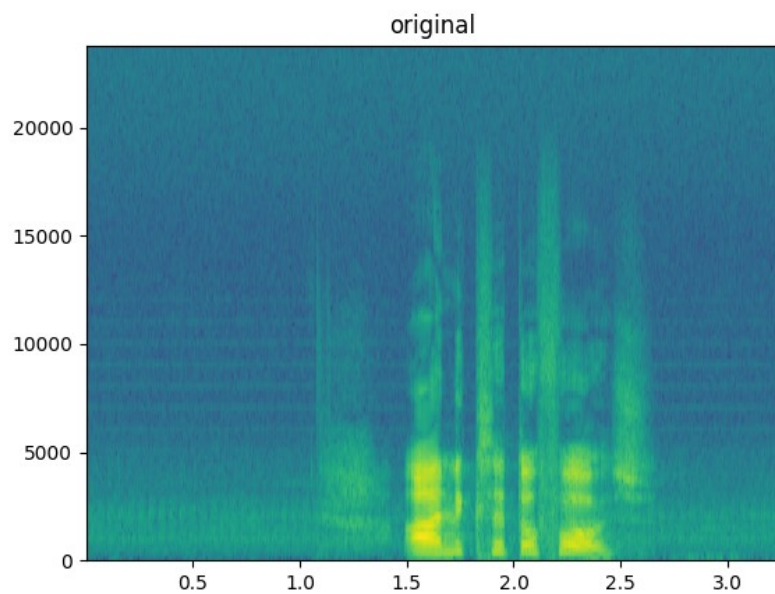
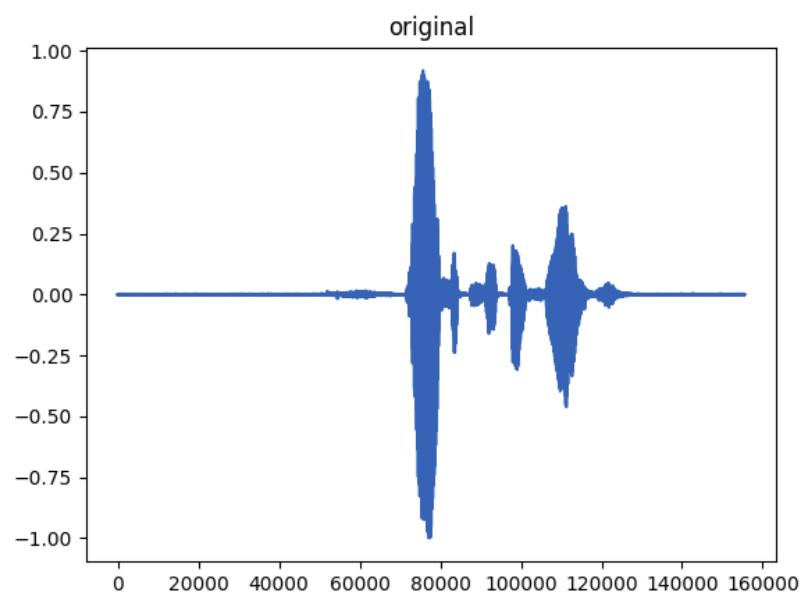
```
SSNR_n_3: 3.453458458705439
SSNR_output: 3.6861107536514215|
```

Και εδώ φαίνεται ότι έχουμε πολύ μικρή βελτίωση του SSNR αναδεικνύοντας και εδώ την αδυναμία του beamformer.

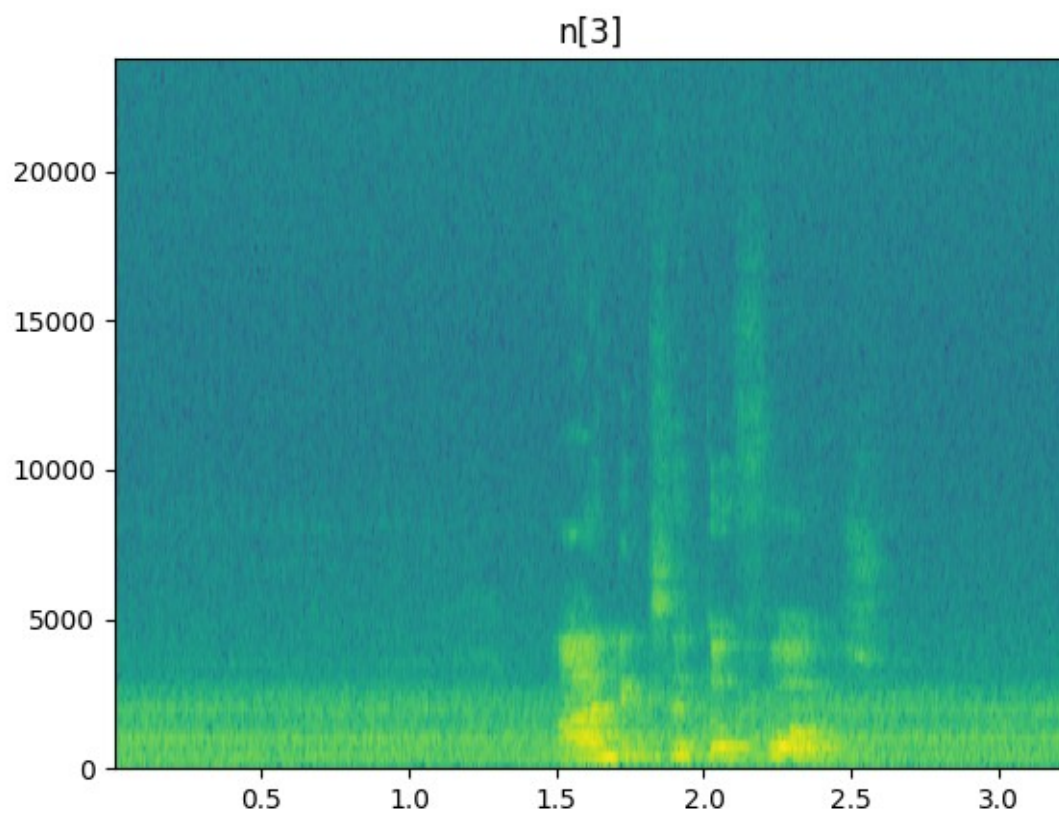
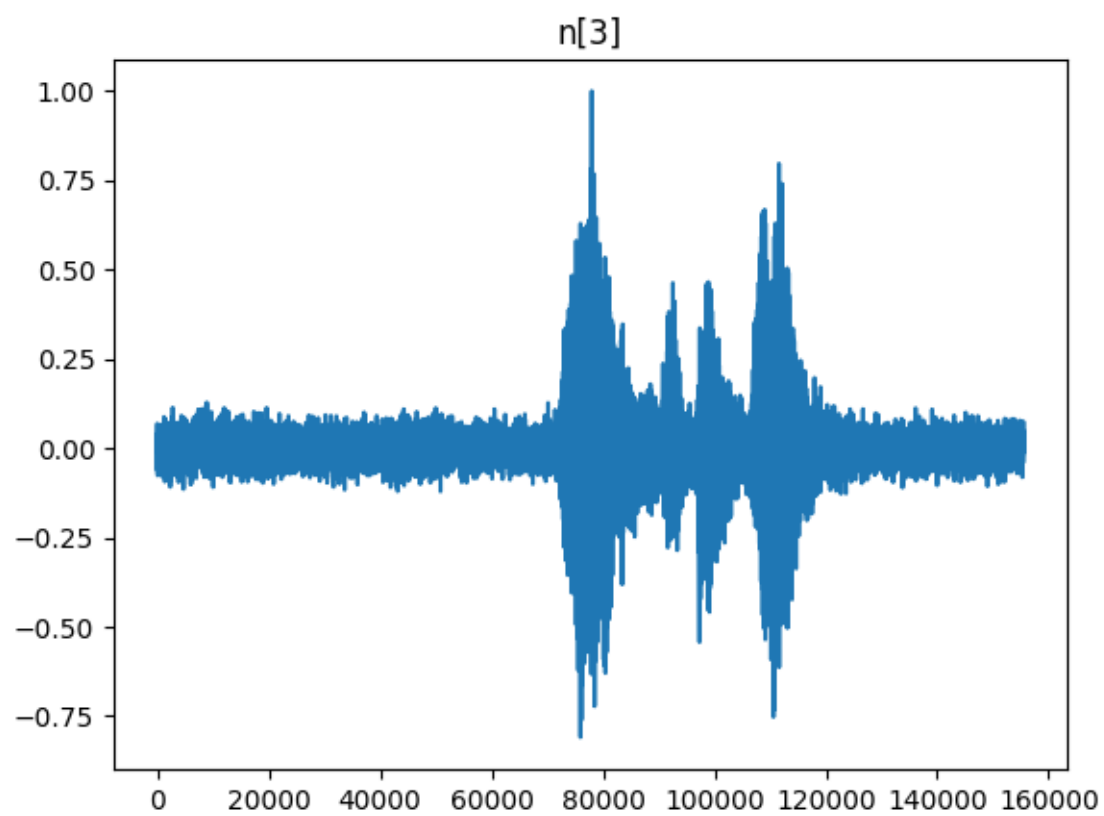
2.2.B)

2)

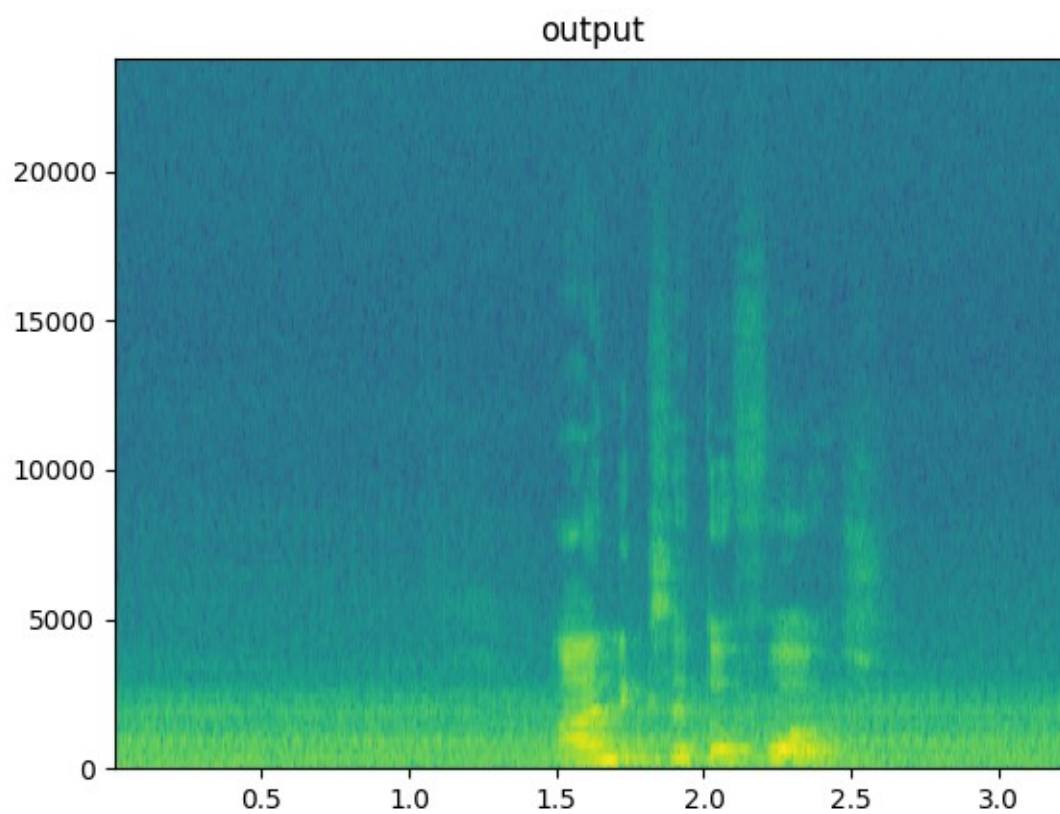
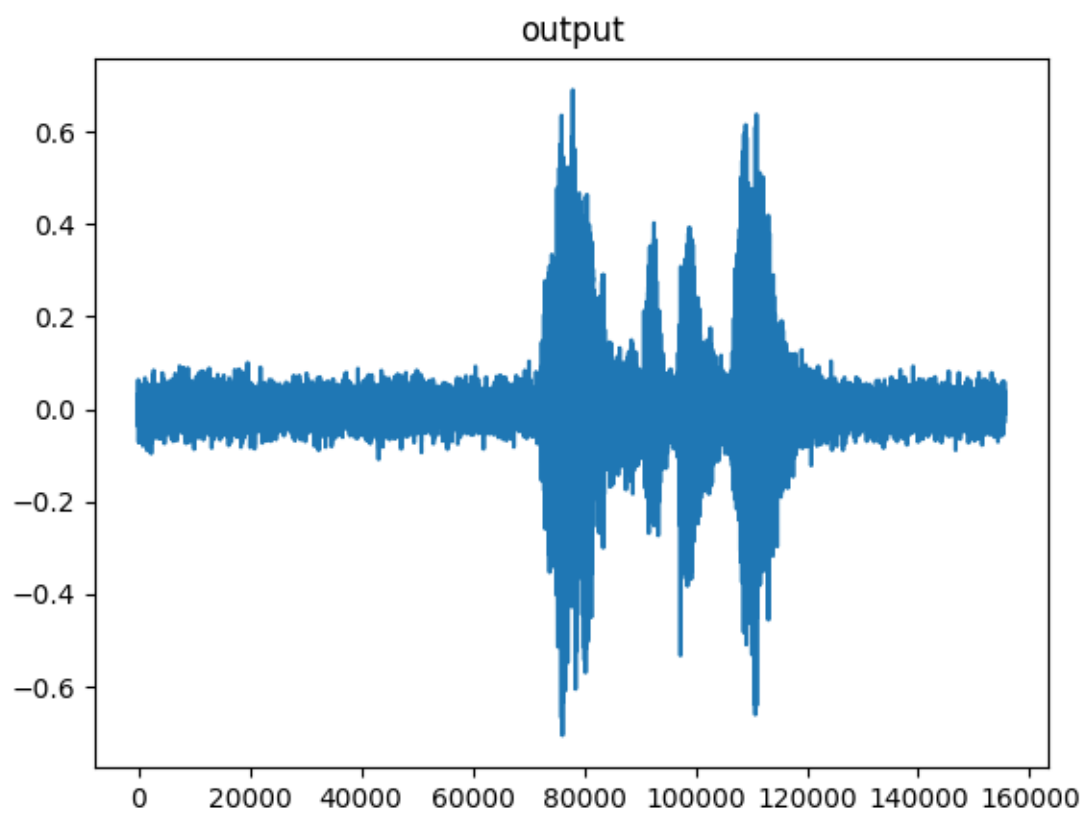
Το αρχικό μας σήμα είναι



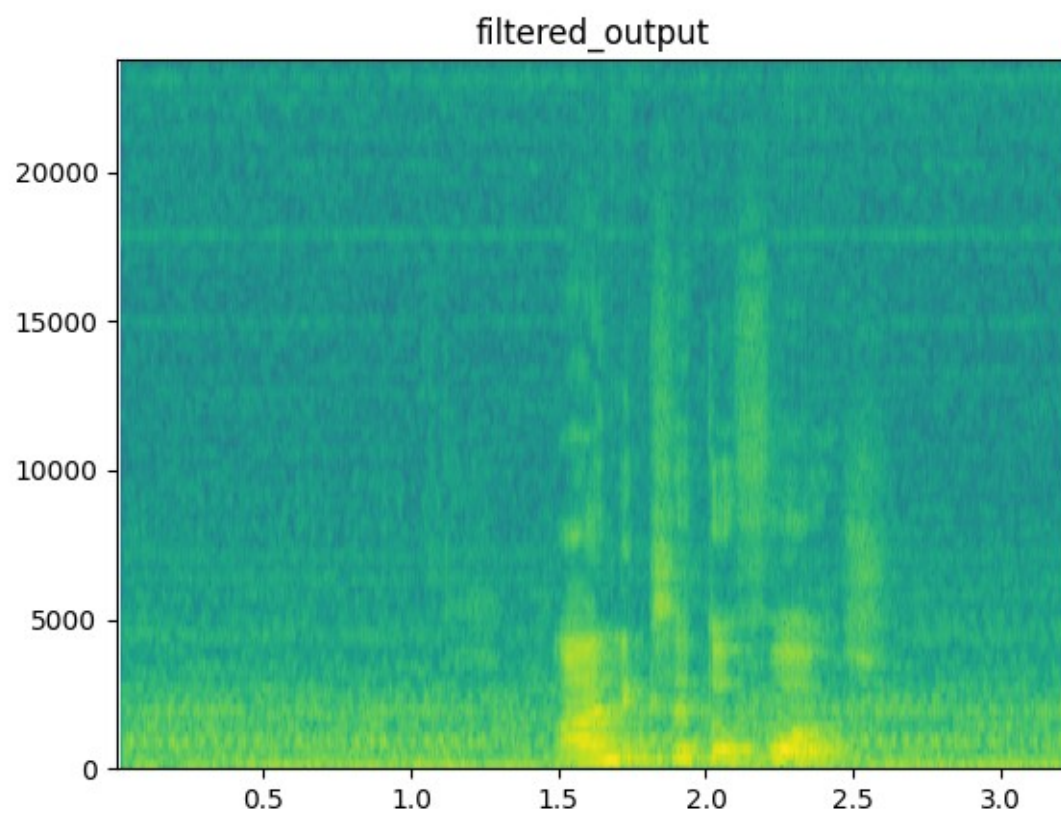
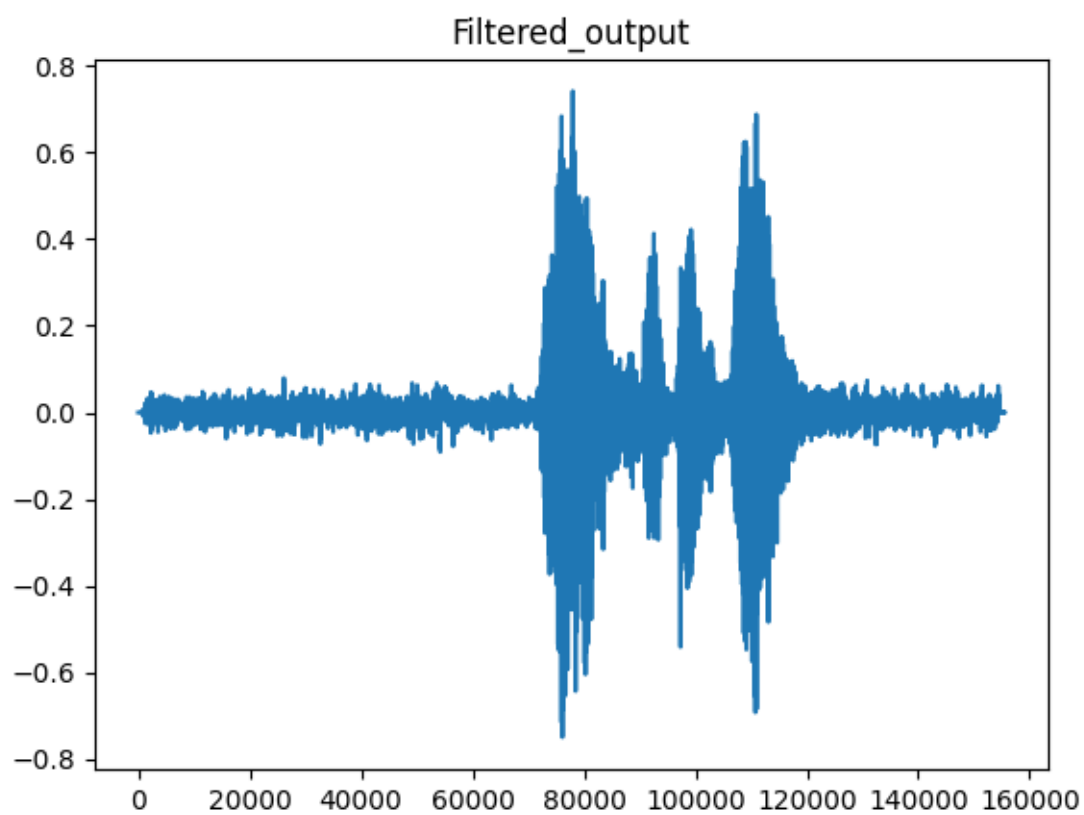
Το σήμα που διαβάζουμε είναι



Από την έξοδο του beamformer παίρνουμε



Τελικά έχουμε



Παρατηρούμε πως από τα διάγραμμα έχει μειωθεί ο θόρυβος.

3) Τα SSNR που υπολογίζουμε είναι:

SSNR_n_3: 1.869223064861179

SSNR_output: 2.097517908472103

SSNR_filtered_output: 3.4702657901401683

Παρατηρούμε πως σε κάθε βήμα της επεξεργασίας μας έχουμε μια βελτίωση του ήχο μας.

4)

Average SSNR's: 1.617092903679724

Παρατηρούμε πως σε σχέση με το αρχικό σήμα μας από το μικρόφωνο 3, το SSNR έχει σχεδόν διπλασιαστεί.