

Ψηφιακή Επεξεργασία Σημάτων 3η Εργαστηριακή Άσκηση

Θέμα: Συστοιχίες Μικροφώνων (Microphone Arrays) και Πολυκαναλική Επεξεργασία Σημάτων (Multichannel Signal Processing)

Μαρία Παρέλλη Γεώργιος-Ορέστης Χαρδούβελης
03115155 03115100
6ο Εξάμηνο 6ο Εξάμηνο

Μέρος 1. Συστοιχίες Μικροφώνων και Χωρικό Φιλτράρισμα (Spatial Filtering)

Η εφαρμογή που θα μελετηθεί στην παρούσα άσκηση είναι το *speech enhancement*. Με τη χρήση συστοιχιών μπορεί να γίνει χωρικό φιλτράρισμα των ακουστικών σημάτων προκειμένου να ενισχυθούν ή να απορριφθούν σήματα που καταφθάνουν στη συστοιχία από συγκεκριμένη κατεύθυνση. Αυτό επιτυγχάνεται με κατάλληλο συνδυασμό των σημάτων που καταγράφονται από τα διάφορα μικρόφωνα ώστε το επιθυμητό σήμα που καταφθάνει από συγκεκριμένη κατεύθυνση να ενισχυθεί με ενισχυτική συμβολή, ενώ ο θόρυβος από τις υπόλοιπες κατευθύνσεις να εξασθενηθεί με αποσβεστική συμβολή.

Συγκεκριμένα,

Στην παρούσα εφαρμογή χρησιμοποιούμε συστοιχία 7 μικροφώνων, τα οποία τοποθετούνται σε ισαπέχοντα σημεία στον z άξονα. Αυτά δειγματοληπτούν το ακουστικό πεδίο, καταγράφοντας τα σήματα

$$f(t, p) = [f(t, p_0), f(t, p_1), f(t, p_2), f(t, p_3), f(t, p_4), f(t, p_5), f(t, p_6), f(t, p_7)]$$

Κάθε καταγεγραμμένο σήμα φιλτράρεται από ένα γραμμικό, χρονικά αναλλοίωτο φίλτρο με κρουστική απόκριση $h_n(t)$ και στη συνέχεια τα σήματα αθροίζονται δίνοντας την τελική έξοδο. Το σύστημα των φίλτρων και του αθροιστή λέγεται *beamformer*.

$$y(t) = \sum_{n=0}^{N-1} \int_{-\infty}^{\infty} h_n(t - \tau) f(\tau, \mathbf{p}_n) d\tau = \int_{-\infty}^{\infty} \mathbf{h}^T(t - \tau) \mathbf{f}(\tau, \mathbf{p}_n) d\tau,$$

Κάθε μικρόφωνο καταγράφει το ηχητικό σήμα με μία χρονική μετατόπιση σε σχέση με τα υπόλοιπα, λόγω της καθυστέρησης διάδοσης του ηχητικού κύματος.

Συνεπώς στο πεδίο της συχνότητας το ηχητικό σήμα που καταγράφεται στην αρχή των αξόνων είναι :

$$\mathbf{F}(\omega, \mathbf{a}) = \mathbf{d}(\mathbf{k})F(\omega), \quad \mathbf{d}(\mathbf{k}) = [e^{-j\mathbf{k}^T \mathbf{p}_0}, e^{-j\mathbf{k}^T \mathbf{p}_1}, \dots, e^{-j\mathbf{k}^T \mathbf{p}_{N-1}}]^T,$$

Όπου $k = \omega/c \cdot a$.

Ένας απλός beamformer είναι ο delay and sum beamformer. Για το delay and sum beamforming επιλέγεται:

$$\mathbf{H}_{DS}^T(\omega) = \frac{1}{N} \mathbf{d}^H(\mathbf{k}_s), \quad \mathbf{k}_s = \frac{\omega}{c} \mathbf{a}_s,$$

Το \mathbf{a}_s ονομάζεται *steering direction* και η εισαγωγή χρονικών ολισθήσεων ονομάζεται *array steering*.

Με το παραπάνω το επιθυμητό σήμα διέρχεται αναλλοίωτο από το beamformer και τα παρεμβαλλόμενα σήματα από άλλες κατευθύνσεις εξασθενούν. Ουσιαστικά ο beamformer θα μπορούσε να θεωρηθεί ως ένας beamformer με ομοιόμορφα βάρη, *steered* στην κατεύθυνση \mathbf{a}_s .

1.4

Θεωρούμε γραμμική ομοιόμορφη συστοιχία μικροφώνων κατα μήκος του άξονα z, τα οποία ισπαέχουν μεταξύ τους.

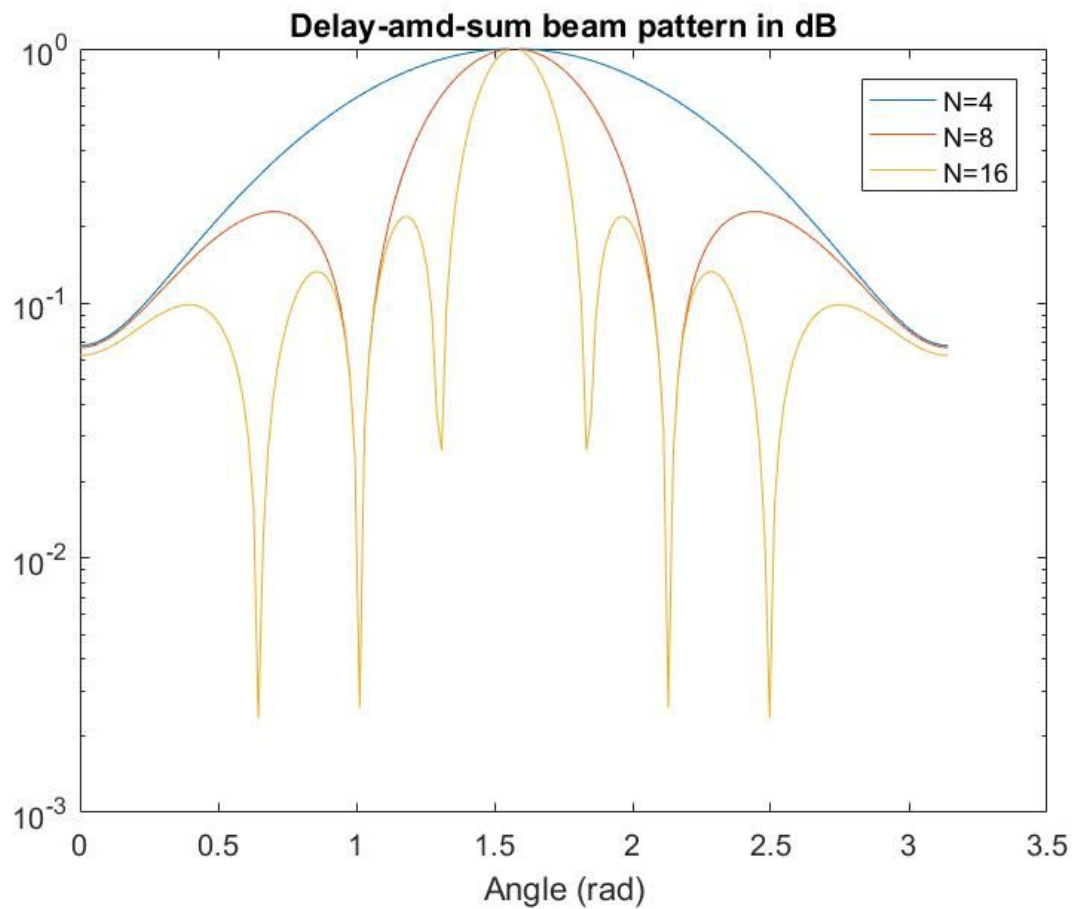
Συνεπώς, το beam pattern γίνεται :

$$B(\omega, \mathbf{a}) = \frac{1}{N} \mathbf{d}^H(\mathbf{k}_s) \mathbf{d}(\mathbf{k}) = \frac{1}{N} e^{-j \frac{N-1}{2} \frac{\omega}{c} d (\cos \theta - \cos \theta_s)} \sum_{n=0}^{N-1} e^{j \frac{\omega}{c} n d (\cos \theta - \cos \theta_s)}$$

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

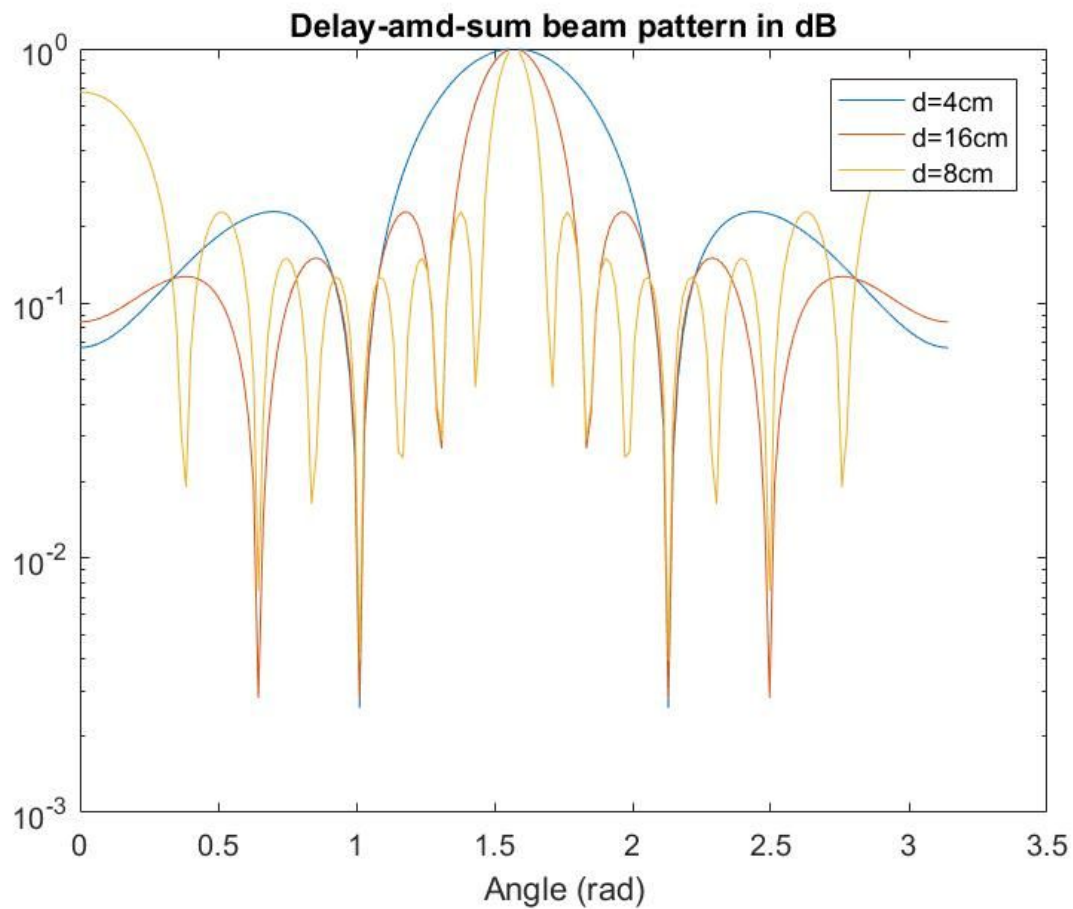
Θεωρούμε ότι το επιθυμητό σήμα φθάνει με γωνία $\theta_s = 90^\circ$

1. Για απόσταση $d=4$ cm και αριθμό μικροφώνων $N=4,8,16$ λαμβάνουμε τις εξής γραφικές παραστάσεις:



Είναι φανερό πως καθώς αυξάνεται ο αριθμός των μικροφώνων αυξάνεται και το πλήθος των λοβών ενώ το κέρδος μειώνεται.

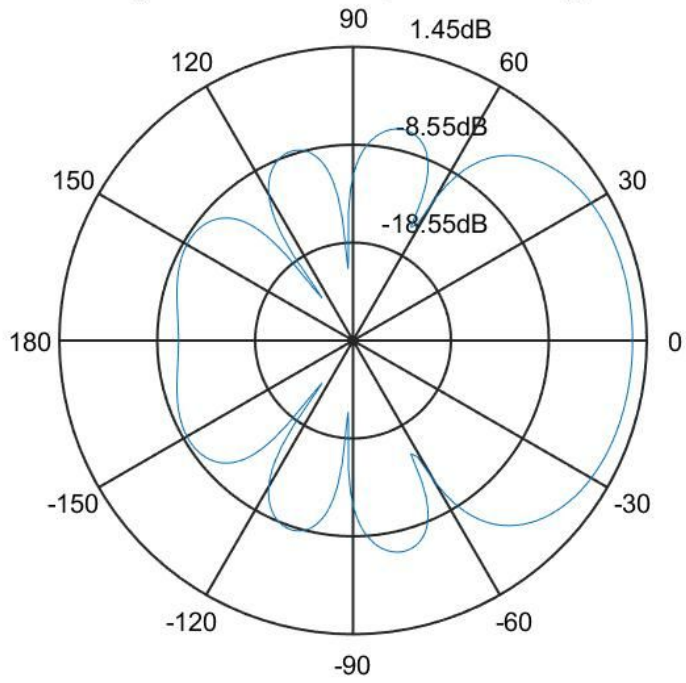
2. Για $N=8$ και απόσταση μικροφώνων $d = 4,8,16$ λαμβάνουμε το εξής:



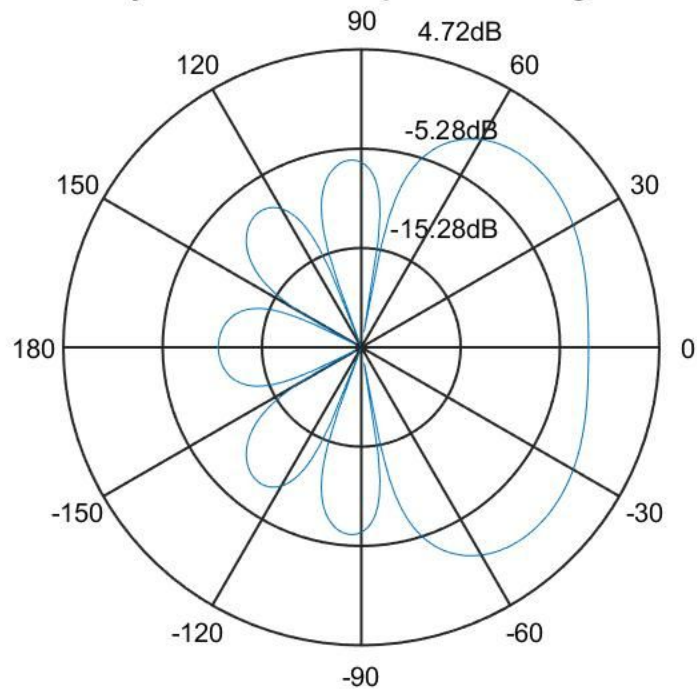
Εδώ φαίνεται ότι όσο αυξάνεται η απόσταση μεταξύ των μικροφώνων έχουμε περισσότερους και πυκνότερους λοβούς (το ύψος των οποίων αναλογικά είναι περίπου ίδιο).

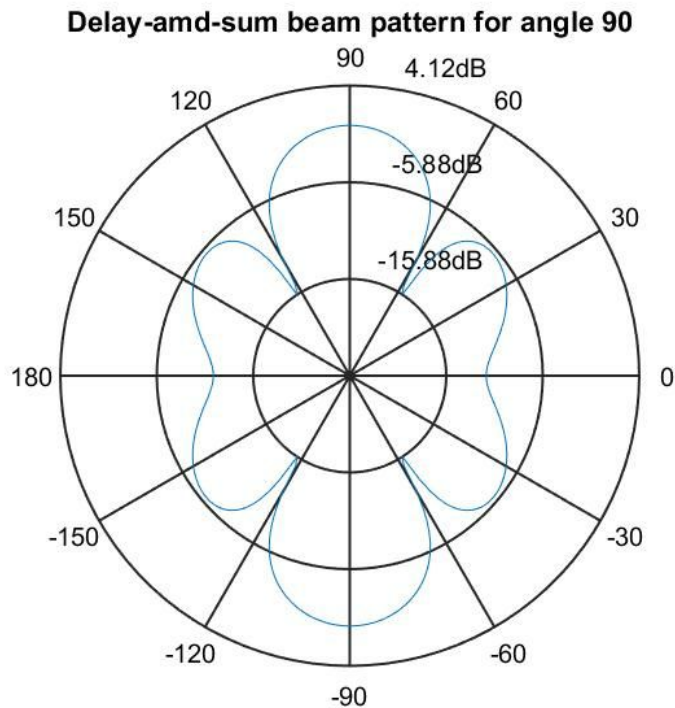
3. Για $\theta_s = 0$, 45 και 90 μοίρες, σχεδιάζουμε το μέτρο του delay-and-sum beam pattern σε λογαριθμική κλίμακα (dB) συναρτήσει της γωνίας $\theta \in [-180^\circ, 180^\circ]$ σε πολικό διάγραμμα. Για την δημιουργία του διαγράμματος χρησιμοποιούμε την συνάρτηση `semilogr_polar`.

Delay-and-sum beam pattern for angle 0



Delay-and-sum beam pattern for angle 45





Όπως περιμέναμε, έχουμε μέτρο 1 στο steering direction, το οποίο συντελεί σε special aliasing και στο μέτρο του beam pattern εμφανίζονται grating lobes. Καθώς μεταβάλλεται η θ_s οι λοβοί μετακινούνται σε διαφορετικές γωνίες, ανάλογα με το ποιές γωνίες ενισχύει το σύστημά μας. Για παράδειγμα για θ_s ίσο με 0 μοίρες δεν έχουμε τόσους λοβούς αλλά υπάρχει ένας κυρίαρχος, για θ_s ίσο με 45 μοίρες έχουμε έναν μεγάλο κύριο λοβό και πολλούς παραπλήσιους λοβούς (δηλαδή ενισχύουμε σήματα από πολλές παραπλήσιες γωνίες, κάτι που δεν το θέλουμε) ενώ για θ_s ίσο με 90 μοίρες οι λοβοί που θα υπήρχαν έχουν ενωθεί σε έναν.

Μέρος 2. Εφαρμογή Beamforming για Speech Enhancement

2.1 Beamforming σε προσομοιωμένα σήματα

Εδώ θεωρούμε beamforming σε προσομοιωμένα δείγματα. Συγκεκριμένα, έχουμε γραμμική συστοιχία μικροφώνων $N = 7$ στοιχείων με απόσταση $d = 4\text{cm}$, η οποία καταγράφει σήματα που παράγονται από δύο σημειακές πηγές. Η μία πηγή παράγει ένα σήμα φωνής και βρίσκεται σε γωνία $\theta = 45$ μοίρες σε σχέση με τη συστοιχία. Η δεύτερη πηγή παράγει ένα σήμα θορύβου και βρίσκεται σε γωνία $\theta = 135$ μοίρες σε σχέση με τη συστοιχία. Ο θόρυβος είναι ζωνοπερατός και η ενέργειά του είναι συγκεντρωμένη στις συχνότητες $f \in [500\text{Hz}, 2.5\text{kHz}]$. Τα σήματα πηγής και θορύβου είναι ασυσχέτιστα.

Στον φάκελο MicArraySimulatedSignals έχουμε το καθαρό σήμα φωνής source.wav (όπως καταγράφεται από το κεντρικό μικρόφωνο της συστοιχίας) ενώ τα σήματα sensor {0,...,6}.wav είναι τα θορυβώδη σήματα όπως καταγράφονται από τα αντίστοιχα μικρόφωνα. Η συχνότητα δειγματοληψίας για όλα τα παραπάνω είναι 48kHz.

A) Delay-and-sum beamforming

1. Αρχικά υπολογίζουμε τα βάρη για τον delay-and-sum-beamformer με βάση την:

$$\mathbf{H}_{DS}^T(\omega) = \frac{1}{N} \mathbf{d}^H(\mathbf{k}_s), \quad \mathbf{k}_s = \frac{\omega}{c} \mathbf{a}_s$$

Τα $\mathbf{d}(\mathbf{k}_s)$ υπολογίζονται για κάθε μικρόφωνο / πηγή ξεχωριστά με βάση την:

$$\mathbf{d}(\mathbf{k}) = [e^{-j\mathbf{k}^T \mathbf{p}_0}, e^{-j\mathbf{k}^T \mathbf{p}_1}, \dots, e^{-j\mathbf{k}^T \mathbf{p}_{N-1}}]^T, \quad \text{με } \mathbf{p} \text{ να είναι τα σημεία των}$$

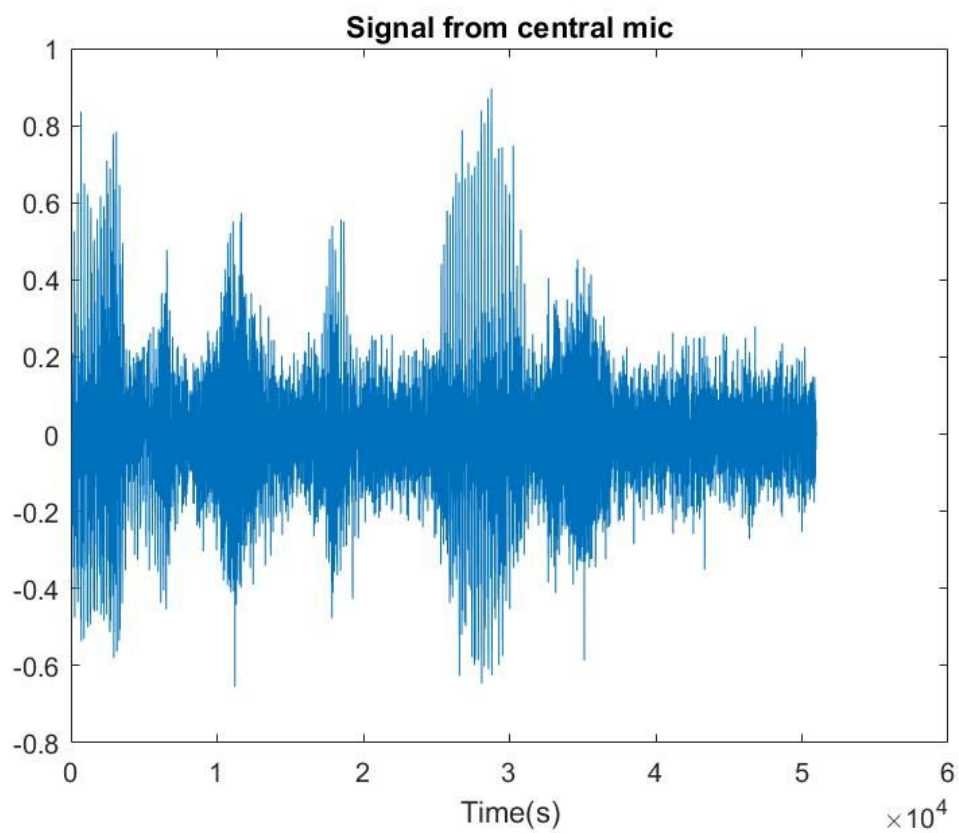
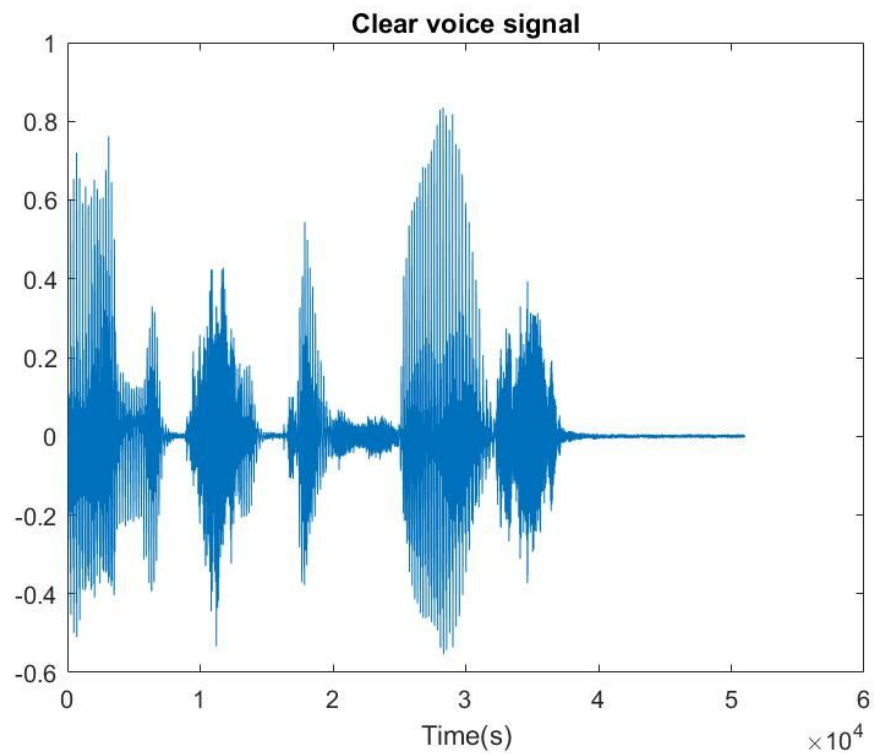
μικροφώνων στο χώρο.

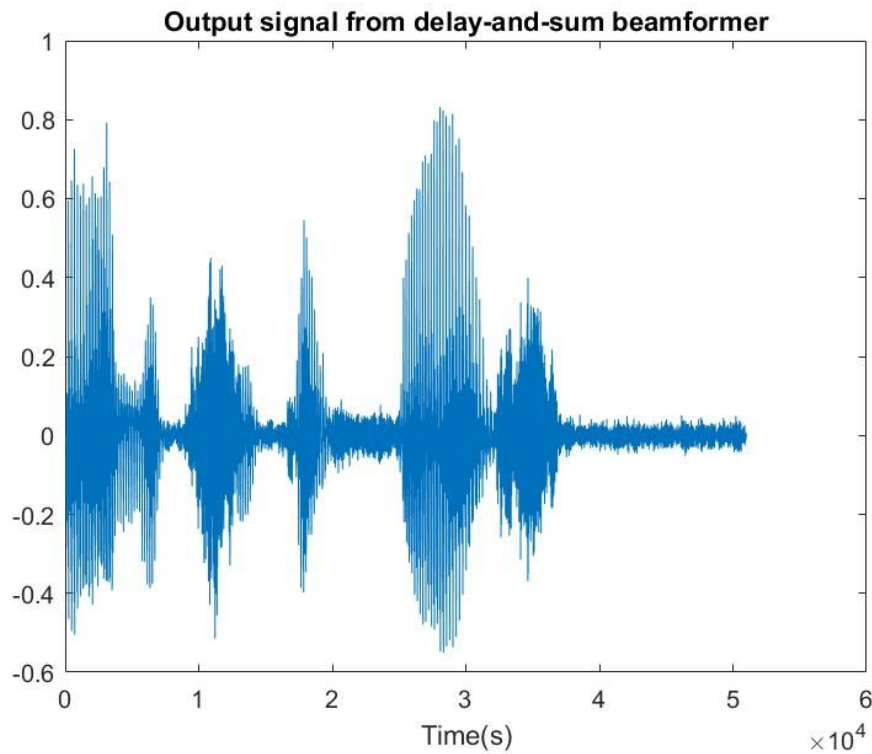
Στη συνέχεια χρησιμοποιούμε τον ερμιτιανό ανάστροφο του \mathbf{d} .

Η κατεύθυνση υπολογίζεται με βάση την γωνία στην οποία βρίσκεται το σήμα φωνής.

Κατ'αυτόν τον τρόπο, δημιουργούμε ένα time-shifting (delay) φίλτρο, από το οποίο περνάνε τα σήματα και ύστερα αθροίζονται ώστε να προκύψει το τελικό σήμα, που αποθηκεύεται σε αρχείο με όνομα **sim_ds.wav**.

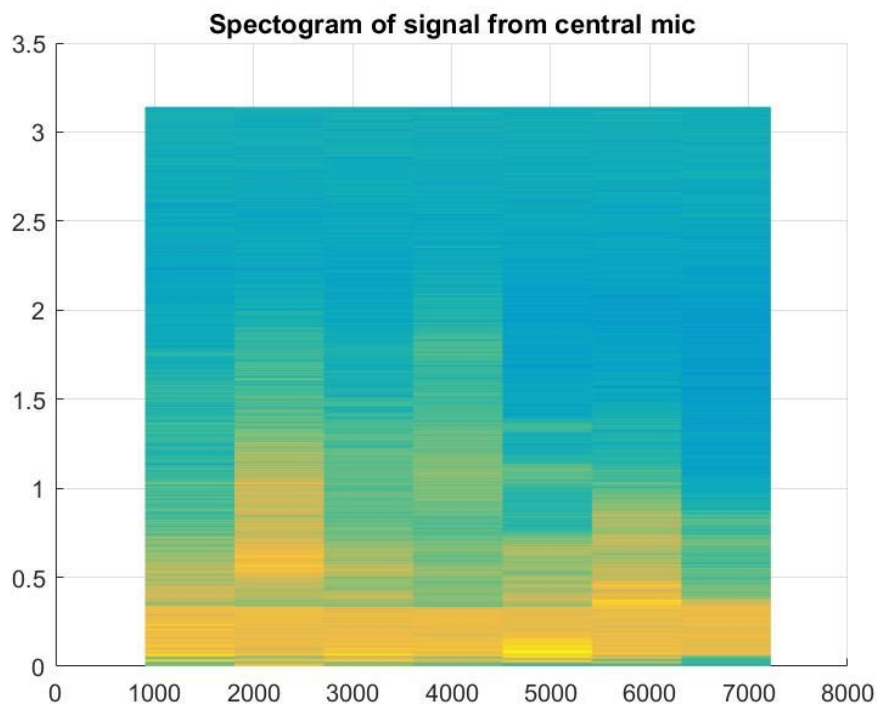
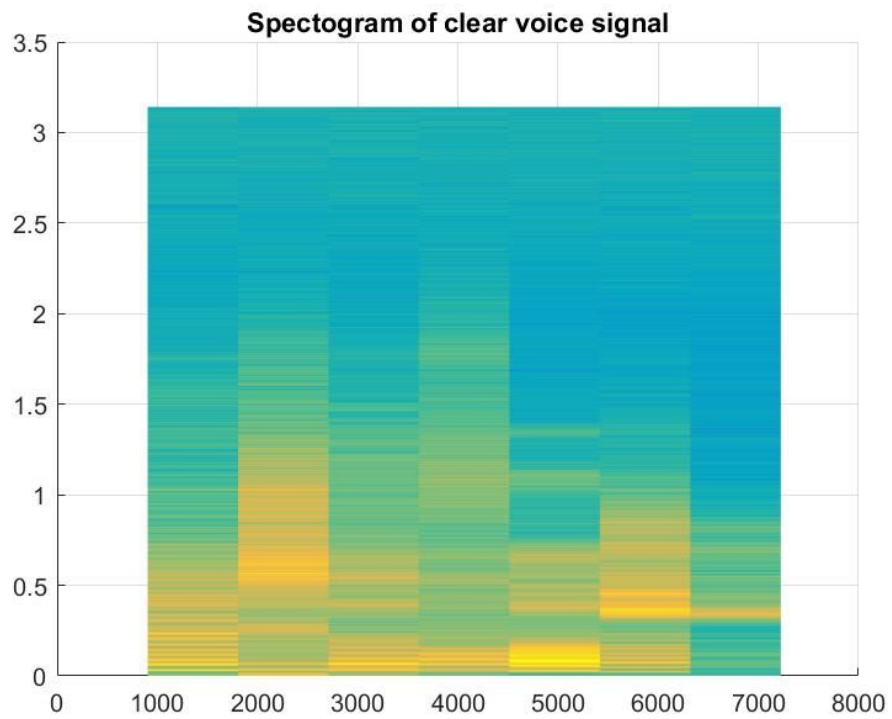
2. Οι ζητούμενες κυματομορφές έχουν ως εξής:

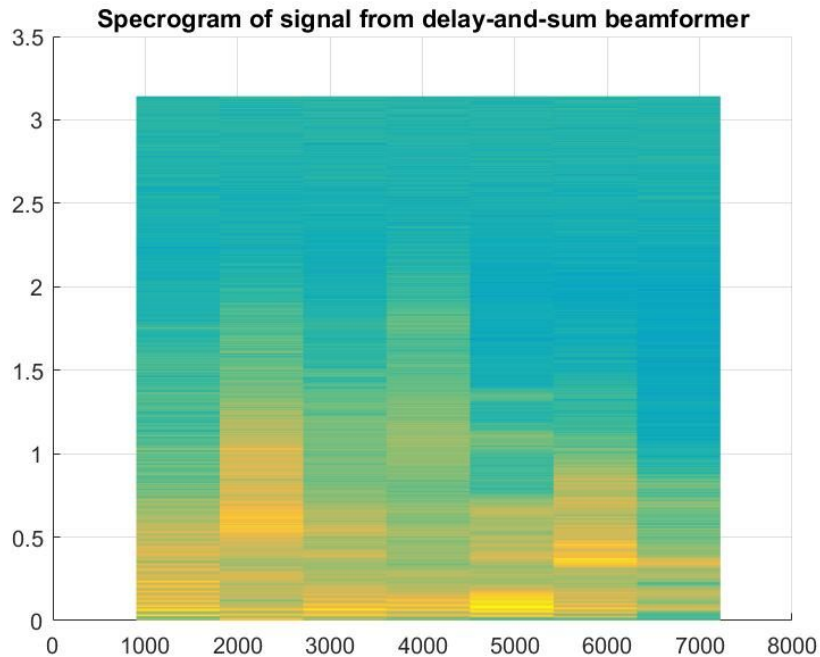




Βλέπουμε πως το τελικό μας σήμα είναι σχεδόν πανομοιότυπο, ειδικότερα στα σημεία όπου υπάρχει ομιλία. Δηλαδή η αποθορυβοποίηση με delay-and-sum beamformer ήταν επιτυχής.

Τα σπεκτρογραφήματα έχουν ως εξής:





Ομοίως και εδώ βλέπουμε πως τα σπεκτρογραφήματα από το αποθρομβοποιημένο σήμα και από το το σήμα της πηγής είναι πρακτικά πανομοιότυπα.

3. Υπολογίζουμε τα ζητούμενα SNR. Ο θόρυβος που χρησιμοποιείται στον τύπο σε κάθε περίπτωση υπολογίζεται αν αφαιρέσουμε από το σήμα το καθαρό σήμα φωνής. Βρίσκουμε πως το SNR του θορυβώδους σήματος στο κεντρικό μικρόφωνο είναι ακριβώς ίσο με **3dB** και το SNR της εξόδου $y(t)$ του delay-and-sum beamformer είναι ίσο με **17.4685dB**. Συμπεραίνουμε λοιπόν πως στην δεύτερη περίπτωση έχουμε σαφώς περισσότερο καθαρό σήμα (σε σχέση με τον θόρυβο) του κεντρικού μικροφώνου.

B) Μονοκαναλικό Wiener φιλτράρισμα

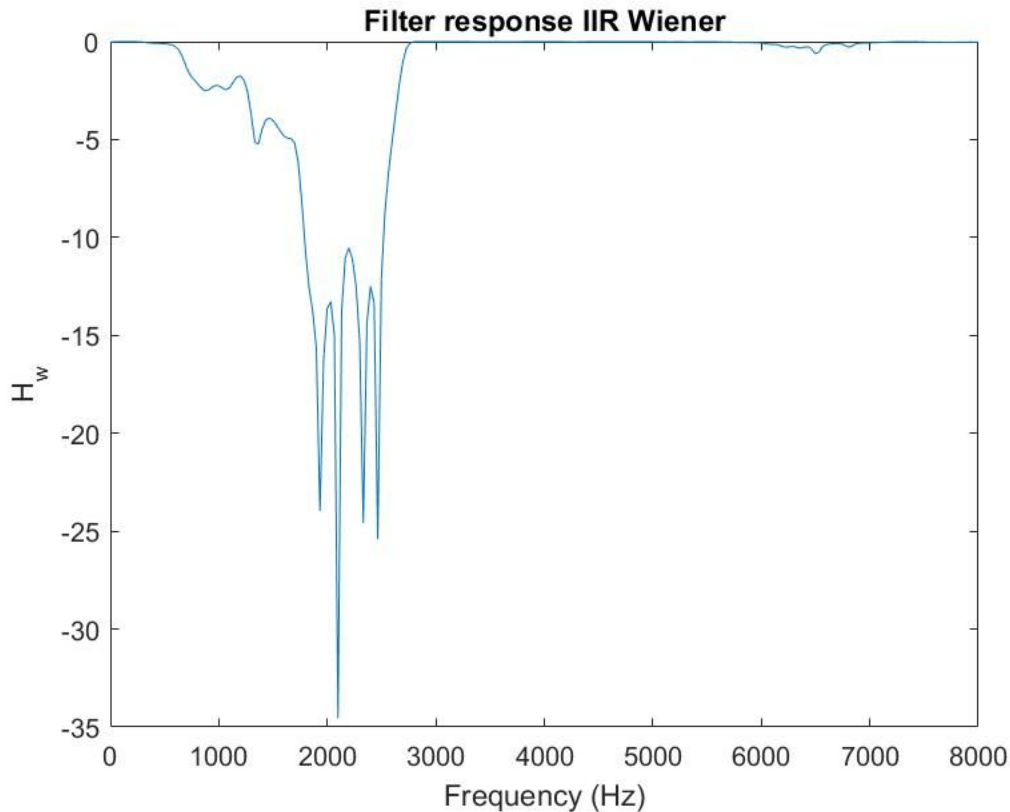
Σε αυτό το μέρος χρησιμοποιούμε την μονοκαναλική μέθοδο Wiener filtering. Θεωρούμε το πλαίσιο $f(t, p3)$, $t \in [0.36s, 0.39s]$ διάρκειας 30ms

1. Υπολογίζουμε αρχικά την απόκριση συχνότητας του IIR Wiener φίλτρου που, εφόσον τα σήματα φωνής και θορύβου είναι -όπως μας δίνεται- ασυσχέτιστα, ίση με:

$$H_W(\omega) = \frac{P_s(\omega)}{P_x(\omega)} = 1 - \frac{P_v(\omega)}{P_x(\omega)}$$

Όπου x το σήμα όπως λαμβάνεται από το μικρόφωνο, s το καθαρό σήμα φωνής, v ο θόρυβος και $P_{x,s,u}$ τα αντίστοιχα φάσματα ισχύος που υπολογίζονται με την μέθοδο Welch.

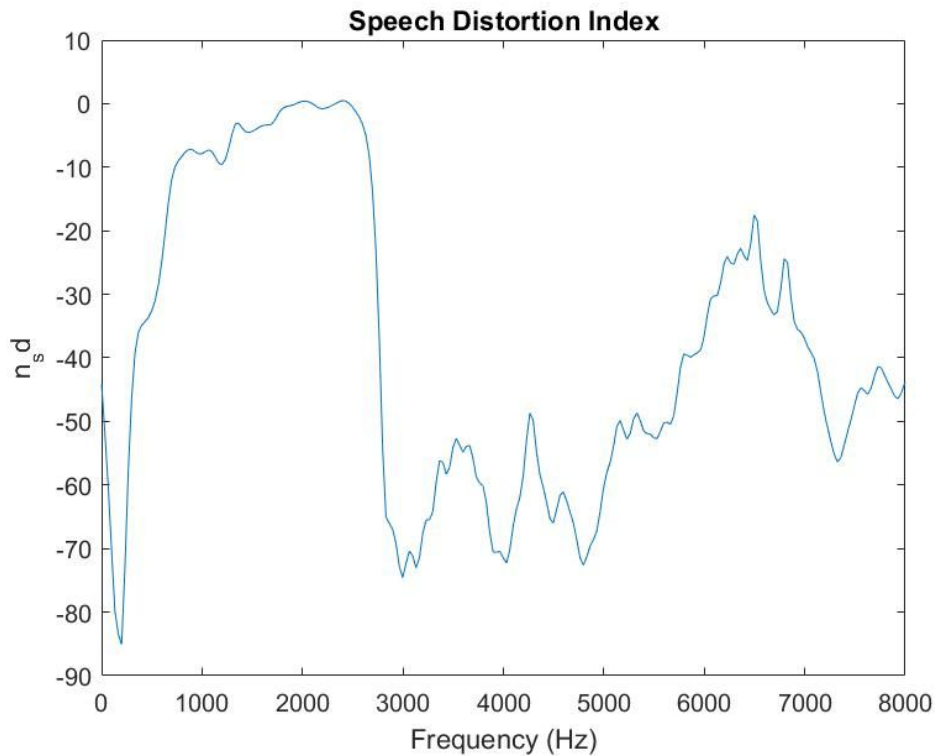
Για συχνότητες $f \in [0, 8\text{kHz}]$ βλέπουμε εδώ το H_w σε λογαριθμική κλίμακα.



2. Το Wiener φίλτρο προκαλεί παραμόρφωση (distortion) στο σήμα φωνής $s(t)$, η οποία ισούται με $s(t) - h_w(t) * s(t)$. Ένας τρόπος να μετρηθεί η παραμόρφωση αυτή είναι το speech distortion index, n_{sd} , το οποίο ορίζεται ως ο λόγος του φάσματος ισχύος της παραμόρφωσης προς το φάσμα ισχύος του σήματος φωνής.

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

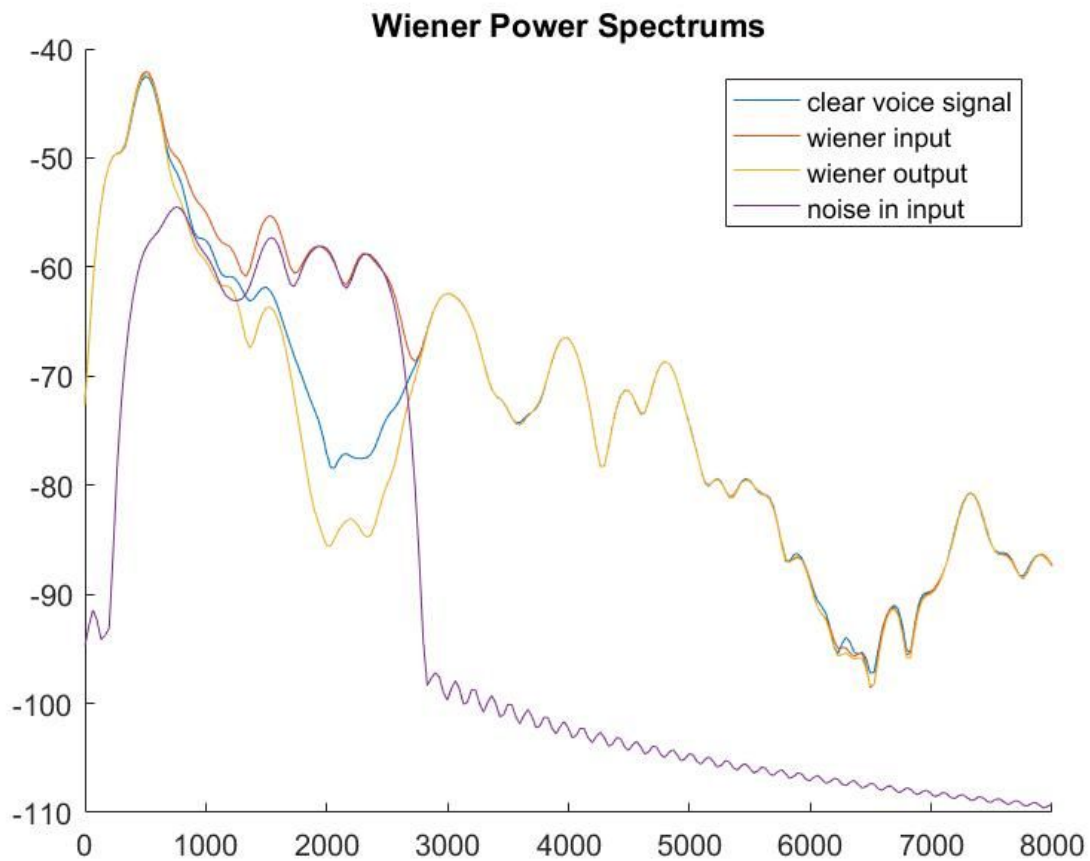
Παρακάτω βλέπουμε το speech distortion index σε λογαριθμική κλίμακα (dB) για συχνότητες $f \in [0, 8\text{kHz}]$.



Παρατηρούμε ότι στις συχνότητες θορύβου (υψηλές συχνότητες), έχουμε μεγάλη παραμόρφωση (μεγάλο distortion), ενώ έχουμε καλή απόκριση στις συχνότητες ανθρώπινης ομιλίας(20-250Hz).

3. Σε αυτό το βήμα εφαρμόζουμε το wiener φιλτράρισμα.

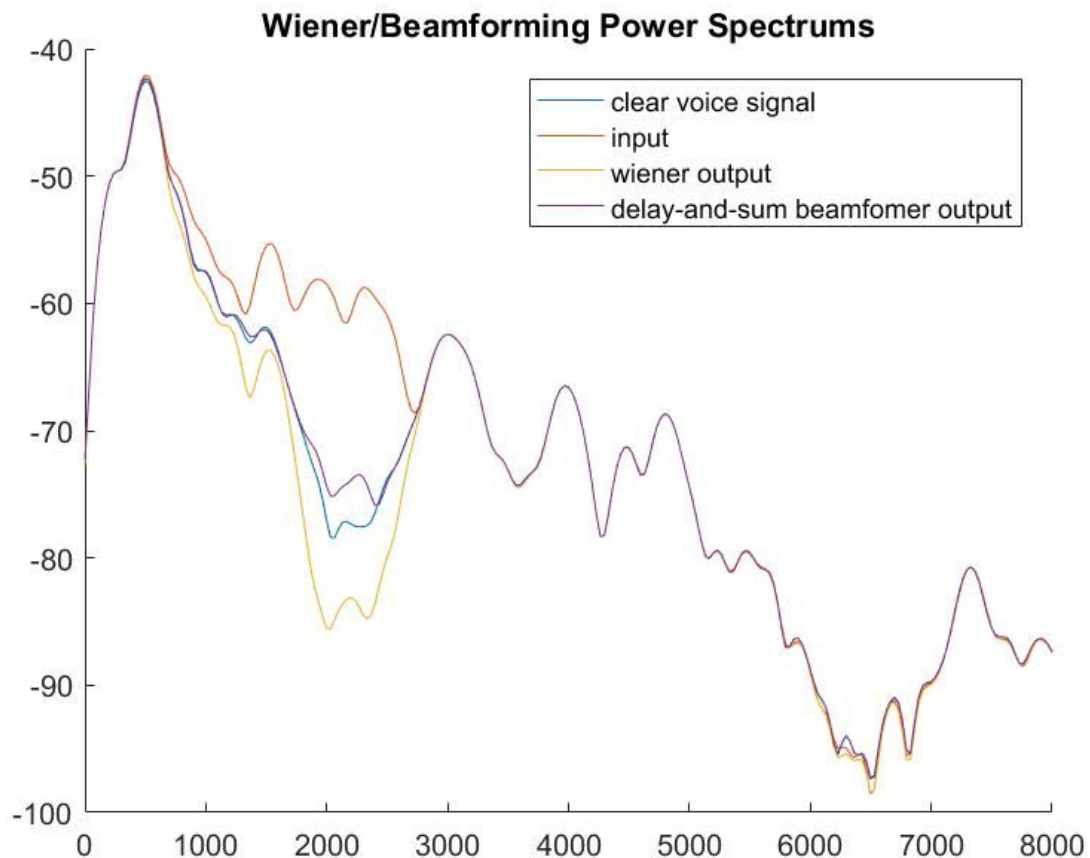
Σχεδιάζουμε ύστερα στην ίδια γραφική παράσταση σε λογαριθμική κλίμακα (dB) και για συχνότητες $f \in [0, 8\text{kHz}]$ τα φάσματα ισχύος για τα εξής σήματα: (α) το καθαρό σήμα φωνής στην είσοδο του Wiener φίλτρου $s(t)$, (β) τη θορυβώδη είσοδο του Wiener φίλτρου $x(t)$, (γ) την έξοδο του Wiener φίλτρου και (δ) το σήμα θορύβου $v(t)$ στην είσοδο του Wiener φίλτρου.



Βλέπουμε πως στο κομμάτι που είχαμε μεγάλη απόκλιση εισόδου (μέχρι 2800Hz περίπου) και όπως βλέπουμε έχουμε σημαντικό θόρυβο, η έξοδος του φίλτρου φτάνει πολύ κοντά στο καθαρό σήμα και άρα είχαμε σε μεγάλο βαθμό πετυχημένη αποθορυβοποίηση.

4. Το SNR στην έξοδο του Wiener φίλτρου είναι ίσο με **11.2395dB**, ενώ είδαμε προηγουμένως πως με την πολυκαναλική μέθοδο είχαμε **17.4685dB**, επομένως έχουμε μια βελτίωση της τάξης του 55.42%.

Στην συνέχεια σχεδιάζουμε στην ίδια γραφική παράσταση σε λογαριθμική κλίμακα (dB) και για συχνότητες $f \in [0, 8\text{kHz}]$ τα φάσματα ισχύος για τα εξής σήματα: (α) το καθαρό σήμα φωνής για το πλαίσιο υπό μελέτη $s(t)$, (β) το θορυβώδες πλαίσιο $x(t)$, (γ) την έξοδο του Wiener φίλτρου και (δ) την έξοδο του delay-and-sum beamformer για το πλαίσιο υπό μελέτη.



Βλέπουμε πως το σήμα από την έξοδο του beamformer σχεδόν ταυτίζεται με το καθαρό σήμα και το προσεγγίζει σαφώς καλύτερα από το σήμα στην έξοδο του wiener φίλτρου.

2.2 Beamforming σε πραγματικά σήματα

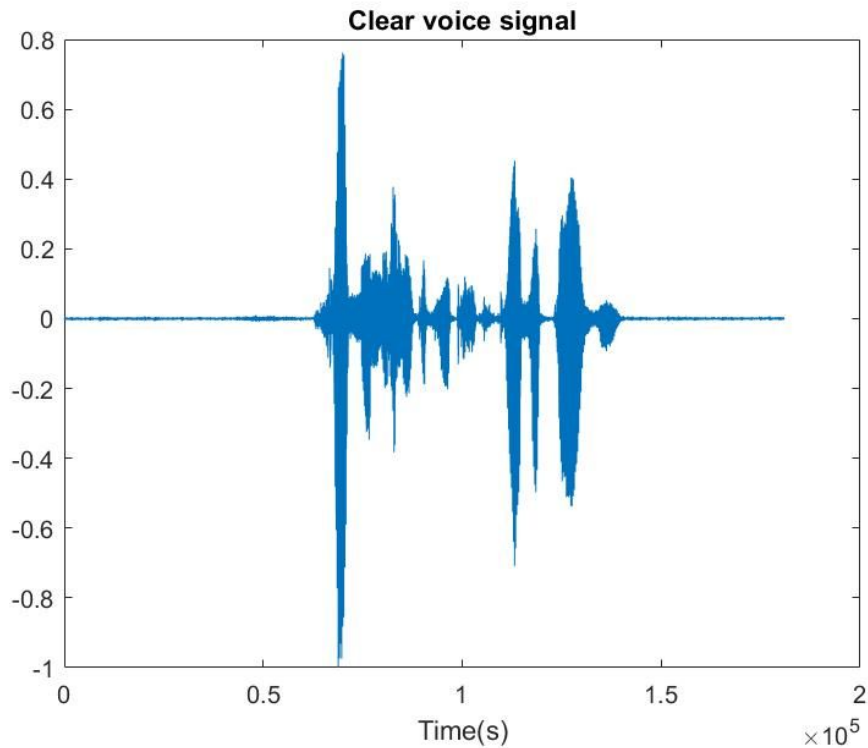
Σε αυτό το μέρος θεωρούμε γραμμική συστοιχία μικροφώνων $N = 7$ στοιχείων με απόσταση $d=4\text{cm}$ βρίσκεται σε ένα θορυβώδες δωμάτιο και καταγράφει το σήμα φωνής που εκφωνεί ένας άνθρωπος σε γωνία $\theta = 45$ μοίρες. Ο θόρυβος δεν προέρχεται από σημειακή πηγή αλλά από διάφορες πηγές, όπως ανεμιστήρες κ.ά., οι οποίες δημιουργούν ένα ιστροπικό και ομογενές πεδίο θορύβου που ονομάζεται *diffuse noise field*. Ο θόρυβος θεωρείται στάσιμος (*stationary random process*).

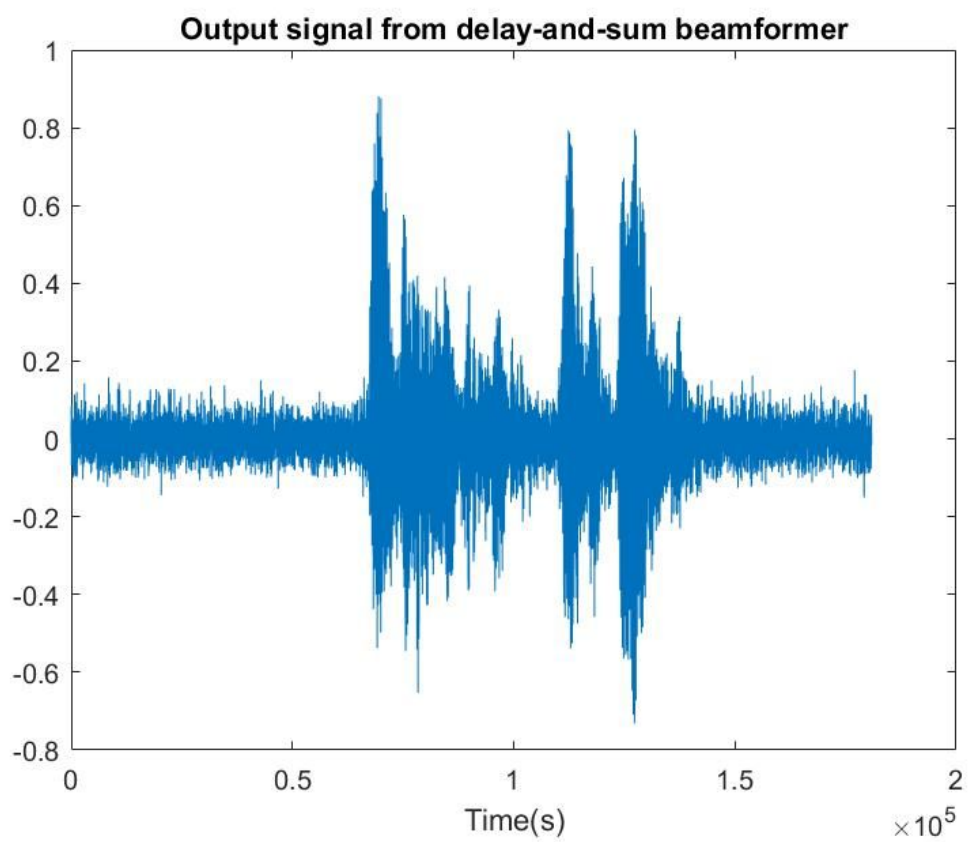
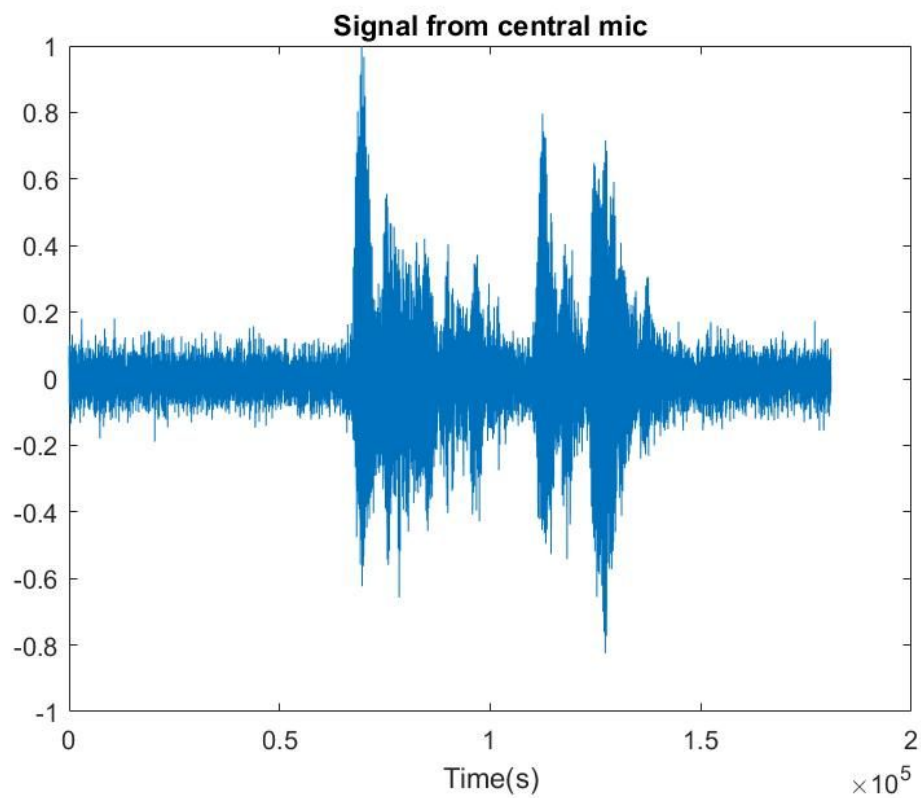
Στον φάκελο `MicArraySignals` έχουμε το καθαρό σήμα φωνής `source.wav` (όπως καταγράφεται από το κεντρικό μικρόφωνο της συστοιχίας) ενώ τα σήματα `sensor {0,...,6}.wav` είναι τα θορυβώδη σήματα όπως καταγράφονται από τα αντίστοιχα μικρόφωνα. Η συχνότητα δειγματοληψίας για όλα τα παραπάνω είναι 48kHz.

A) Delay-and-sum beamforming

1. Υπολογίζουμε ομοίως με το 2.1 τα βάρη του delay-and-sum beamformer και δημιουργούμε ένα time-shifting (delay) φίλτρο, από το οποίο περνάνε τα σήματα και ύστερα αθροίζονται ώστε να προκύψει το τελικό σήμα, που αποθηκεύεται σε αρχείο με όνομα **real_ds.wav**.

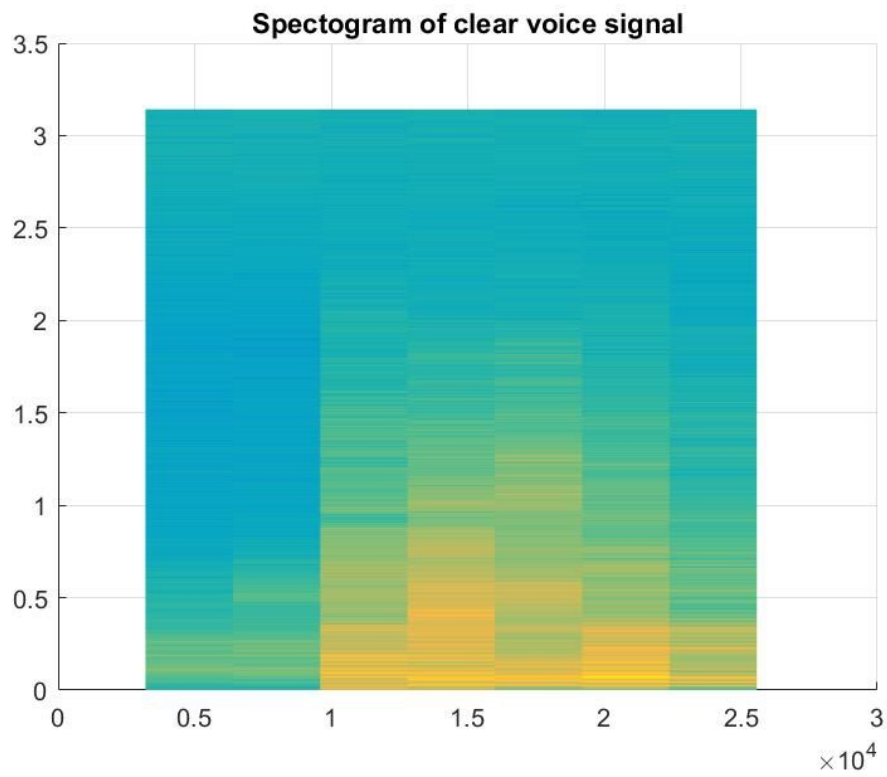
2. Οι ζητούμενες κυματομορφές είναι οι εξής:

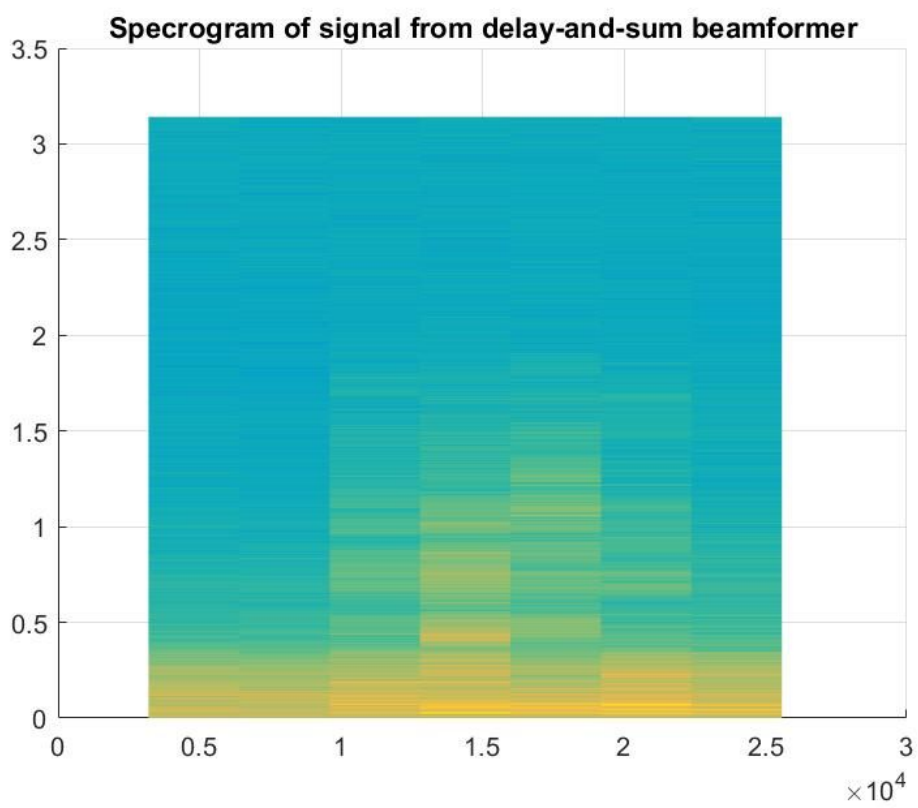
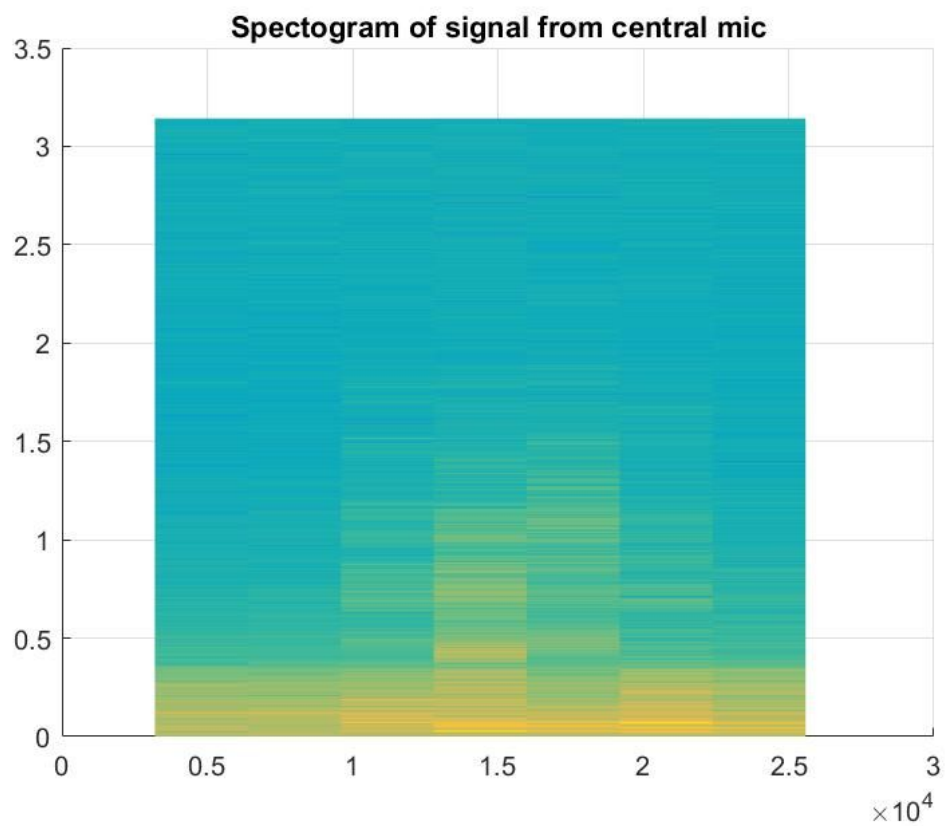




Είναι εμφανές πως το τελικό μας σήμα μοιάζει πολύ με το θορυβώδες (ελάχιστος θόρυβος έχει αφαιρεθεί) και απέχει από το αρχικά. Επομένως η αποθορυβοποίηση δεν ήταν ιδιαίτερα επιτυχής.

Τα αντίστοιχα σπεκτρογράφημα έχουν ως εξής:





Και εδώ είναι εμφανές πως το σπεκτρογράφημα του τελικούτος σήμα μετά το beamforming μοιάζει πιο πολύ με εκείνο του θορυβώδες παρά του αρχικού.

3. Εδώ μας ζητείται να υπολογιστεί το ο SSNR στο κεντρικό μικρόφωνο της συστοιχίας και στην έξοδο του beamformer.

Το SNR ορίζεται ως το μέσο SNR των πλαισίων βραχέος χρόνου του σήματος φωνής.

$$SSNR = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \frac{\sum_{n=Lm}^{Lm+L-1} s^2(n)}{\sum_{n=Lm}^{Lm+L-1} v^2(n)},$$

όπου L το μήκος των πλαισίων, s(n) το σήμα φωνής και u(n) ο θόρυβος. Πλαίσια με SNR μεγαλύτερο των 35dB δεν έχουν σημαντικές διαφορές στην ποιότητα σήματος και τίθενται στα 35dB για την εξαγωγή του μέσου όρου. Σε πλαίσια σιωπής το SNR είναι έντονα αρνητικό, οπότε πλαίσια με SNR μικρότερο μίας τιμής κατωφλίου που θέτει το όριο διαχωρισμού φωνής από σιωπή αγνοούνται κατά τον υπολογισμό του μέσου όρου. Στην παρούσα άσκηση θέσαμε ως κατώφλι τα 0db.

Για τον υπολογισμό των SSNR χρησιμοποιείται η ισχύς του καθαρού σήματος φωνής καθώς και εκείνου που λαμβάνουμε από το κεντρικό μικρόφωνο.

Η ισχύς του καθαρού σήματος υπολογίζεται μέσω της ισχύς θορύβου. Η ισχύς θορύβου υπολογίζεται από ένα κομμάτι του σήματος που περιέχει μόνο θόρυβο. Θεωρώντας ότι ο θόρυβος είναι στάσιμος, το σ_v^2 δε μεταβάλλεται με το χρόνο (είναι το ίδιο για όλα τα πλαίσια). Σε κάθε πλαίσιο $x(t) = s(t) + v(t)$ έχουμε $\sigma_s^2 = \sigma_x^2 - \sigma_v^2$, θεωρώντας ότι ο θόρυβος και το σήμα φωνής είναι ασυσχέτιστα.

Καταλήγουμε λοιπόν πως το SSNR στο κεντρικό μικρόφωνο της συστοιχίας είναι **9.6858** ενώ στην έξοδο του beamformer **10.2680**. Σαφώς στην έξοδο του beamformer είναι μεγαλύτερο, παρόλα αυτά η διαφορά δεν είναι μεγάλη. Συνεπώς συμπεραίνουμε πως η επεξεργασία μόνο με τον beamformer δεν επιφέρει τα επιθυμητά αποτελέσματα σε πραγματικά ακουστικά σήματα.

Αυτό οφείλεται κυρίως στο γεγονός ότι δεν έχουμε συγκεκριμένη γωνία από την οποία κατευθύνεται ο ήχος στα μικρόφωνα και γι' αυτό καθώς ενισχύει μια συγκεκριμένη γωνία υπάρχει θόρυβος και σε αυτήν την γωνία που ενισχύουμε.

B) Post-filtering με Wiener φίλτρο

Σε αυτό το βήμα εφαρμόζουμε και μονοκαναλικό φιλτράρισμα μετά το beamforming του παραπάνω βήματος για περαιτέρω βελτίωση της ποιότητας του σήματος. Η διαδικασία αυτή ονομάζεται post-filtering.

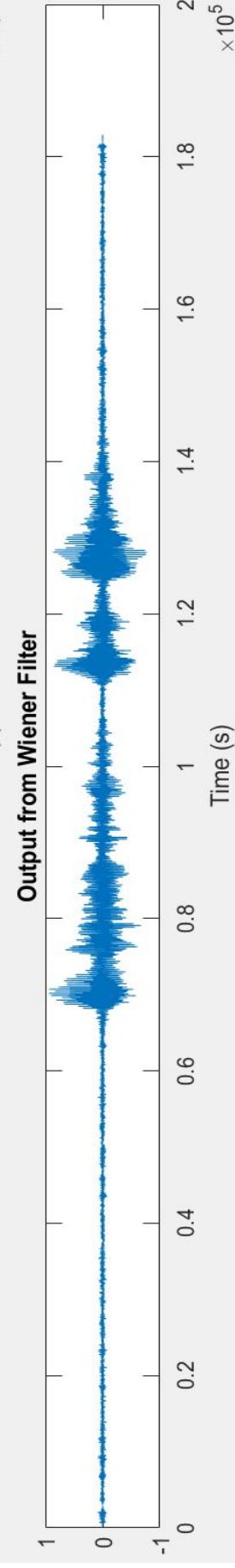
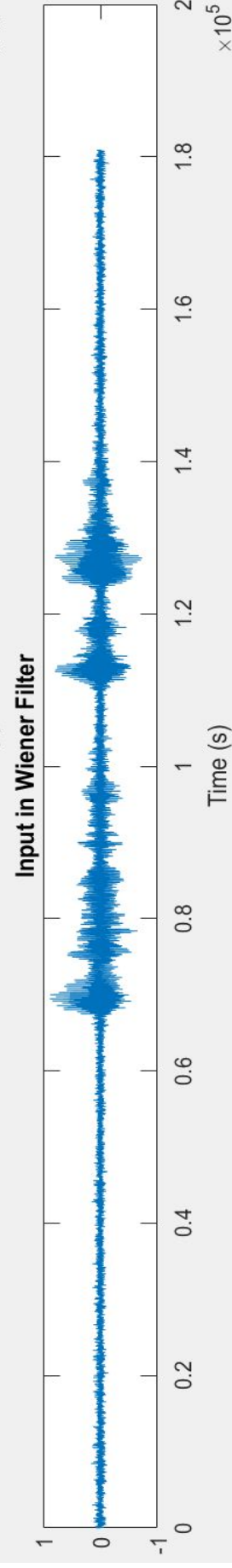
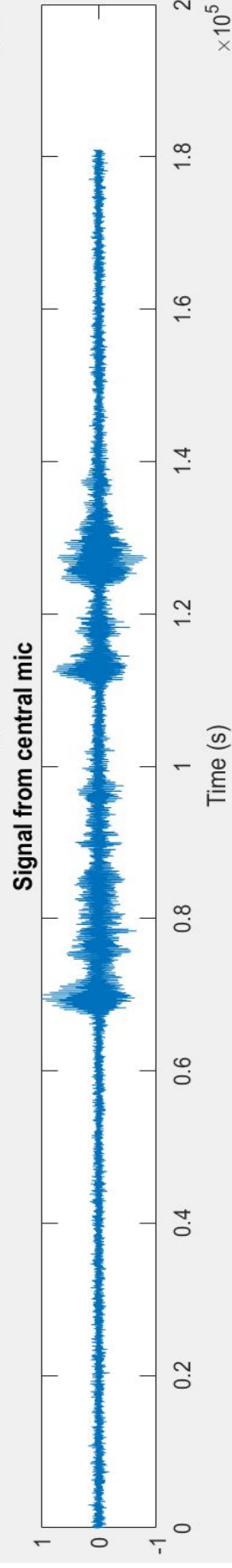
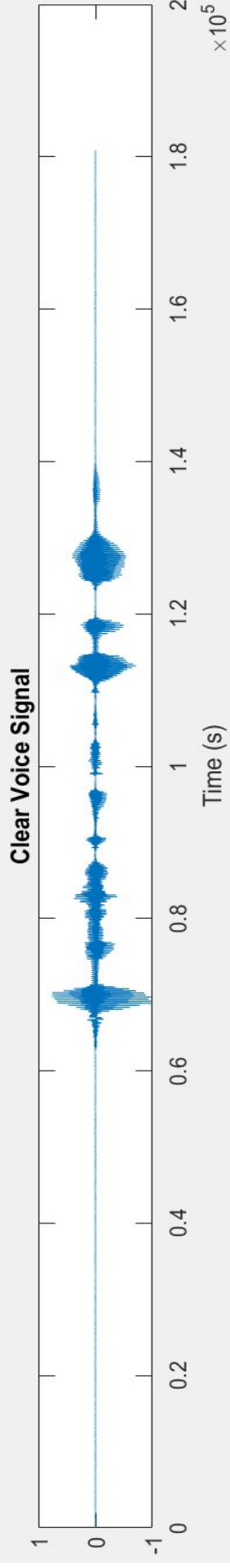
1. Εφαρμόζουμε μονοκαναλικό IIR Wiener φιλτράρισμα στην έξοδο του beamformer.

Επειδή το σήμα φωνής δεν είναι στάσιμο ακολουθούμε ανάλυση βραχέος χρόνου (short-time analysis) χωρίζοντας το σήμα σε επικαλυπτόμενα πλαίσια διάρκειας 30ms. Χρησιμοποιούμε Hamming παράθυρο. Για κάθε πλαίσιο υπολογίζουμε τη συνάρτηση μεταφοράς του Wiener φίλτρου εκτιμώντας το φάσμα ισχύος για το κάθε πλαίσιο του θορυβώδους σήματος με τη μέθοδο Welch. Ύστερα φιλτράρουμε κάθε πλαίσιο και τέλος ανασυνθέτουμε το σήμα στην έξοδο με overlap-add σύνθεση.

Η επικάλυψη έχει επιλεγεί στο 50%.

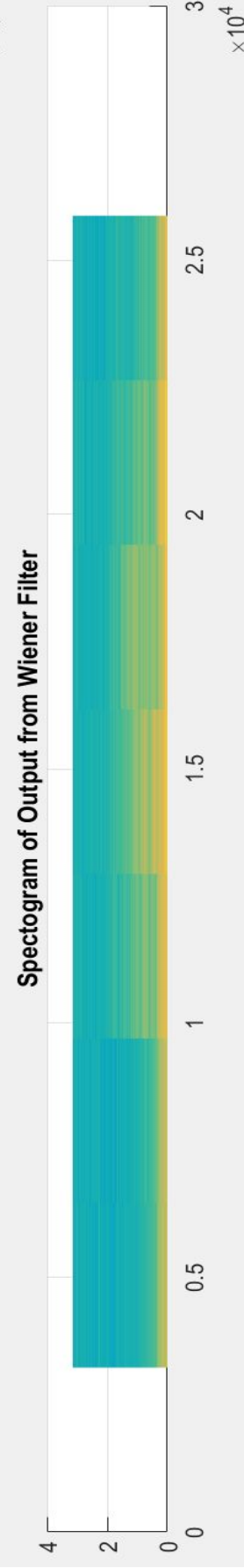
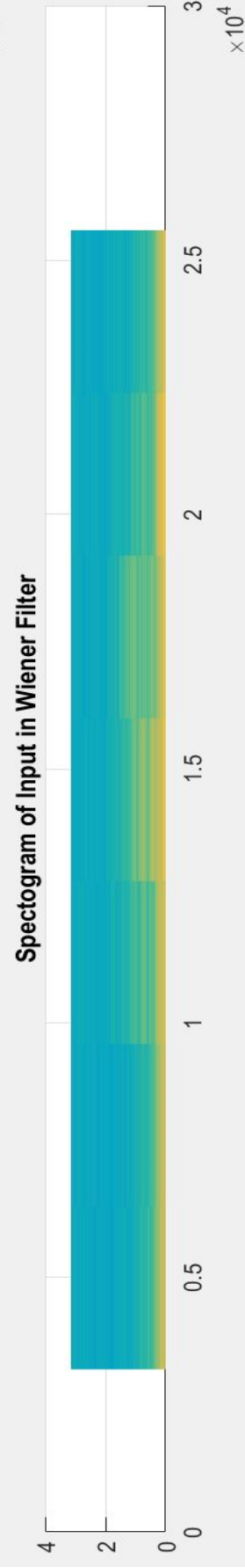
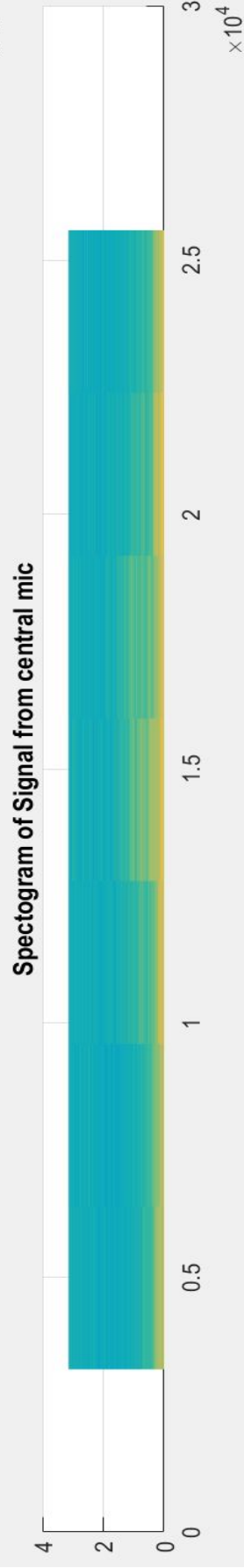
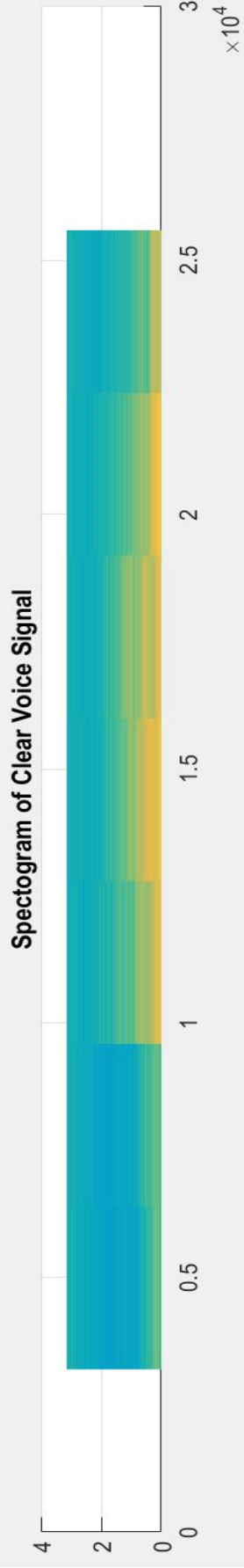
Η έξοδος του Wiener φίλτρου που προκύπτει από την παραπάνω διαδικασία αποθηκεύεται σε αρχείο wav με όνομα real mmse.wav.

2. Οι ζητούμενες κυματομορφές φαίνονται εδώ:



Παρατηρούμε πως με το wiener φίλτρου έχουμε μία περαιτέρω μείωση του θορύβου, το ποίο είναι φανερό και όταν ακούσουμε το σήμα. Ακόμη βέβαια δεν έχει πετύχει η αποθορυβοποίηση που είχαμε στα προσομοιωμένα σήματα.

Τα ζητούμενα σπεκτρογραφήματα έχουν ως εξής:



Ομοίως και εδώ παρατηρούμε πως η έξοδος από το beamformer έχει ελάχιστη βελτίωση από το σήμα στο κεντρικό μικρόφωνο, η οποία βελτιώνεται από το φίλτρο Wiener.

3. Το SSNR στην είσοδο του Wiener φίλτρου είναι εκείνο από την έξοδο του beamformer και είχε υπολογιστεί παραπάνω στα **10.2680**. Στην έξοδο του φίλτρου υπολογίζουμε το SSNR όμοια με το υποερώτημα 2.2.A.2 το οποίο είναι ίσο με **14.8101**. Υπάρχει μια αισθητή αύξηση του SSNR που υποδηλώνει την χρήσιμη πληροφορία στο σήμα έναντι του θορύβου. Επομένως η χρήση του Wiener φίλτρου επέφερε αποτέλεσμα.

4. Ο μέσος όρος των SSNRs των σημάτων εισόδου στο σύστημα delay-and-sum (δηλαδή του σήματος που λαμβάνουμε από το κεντρικό μικρόφωνο) και Wiener post-filter (δηλαδή την έξοδο από τον beamformer) υπολογίζεται στα **9.9769**. Στην τελική έξοδο του συστήματος έχουμε SSNR **14.8101**. Παρατηρείται λοιπόν βελτίωση της τάξης του **48.44%**. Συνεπώς η επιπλέον επεξεργασία με post-Wiener filtering επιφέρει αισθητά καλύτερα αποτελέσματα.