

Ψηφιακή Επεξεργασία Σημάτων – 3^η Εργαστηριακή Άσκηση

Όνομα: Βλαντισλάβ

Επώνυμο: Λέντελ

AM: 03114054

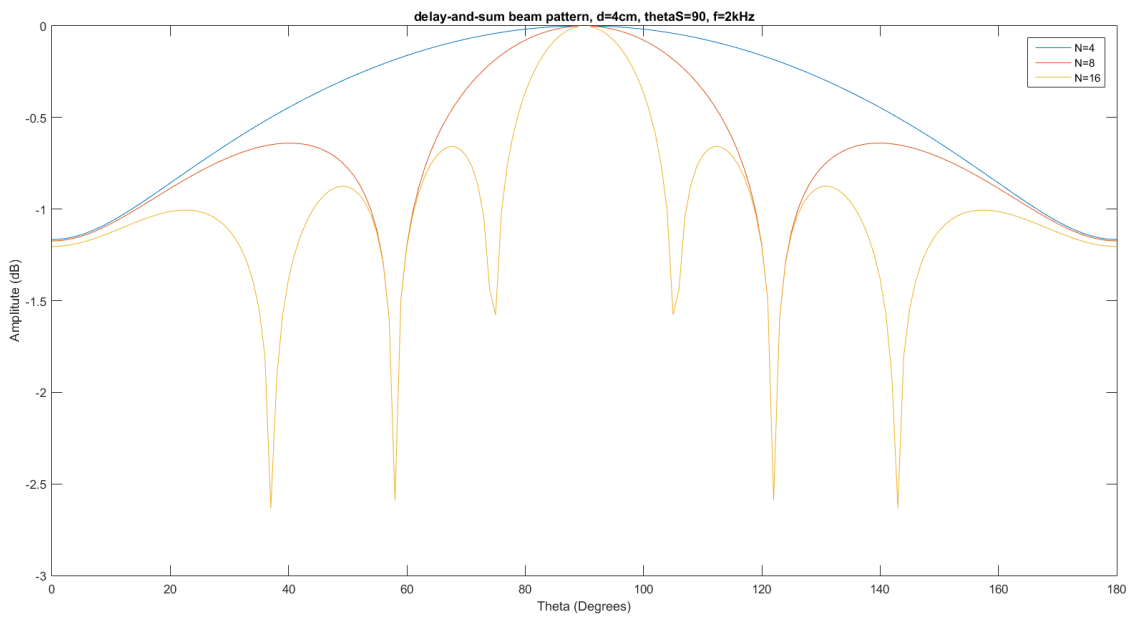
Άσκηση 1.4: ex1_4_1.m, ex1_4_2.m & ex1_4_3.m

Εφόσον έχουμε ομοιόμορφη γραμμική συστοιχία μικροφώνων, ο μαθηματικός τύπος που χρησιμοποιήσαμε είναι :

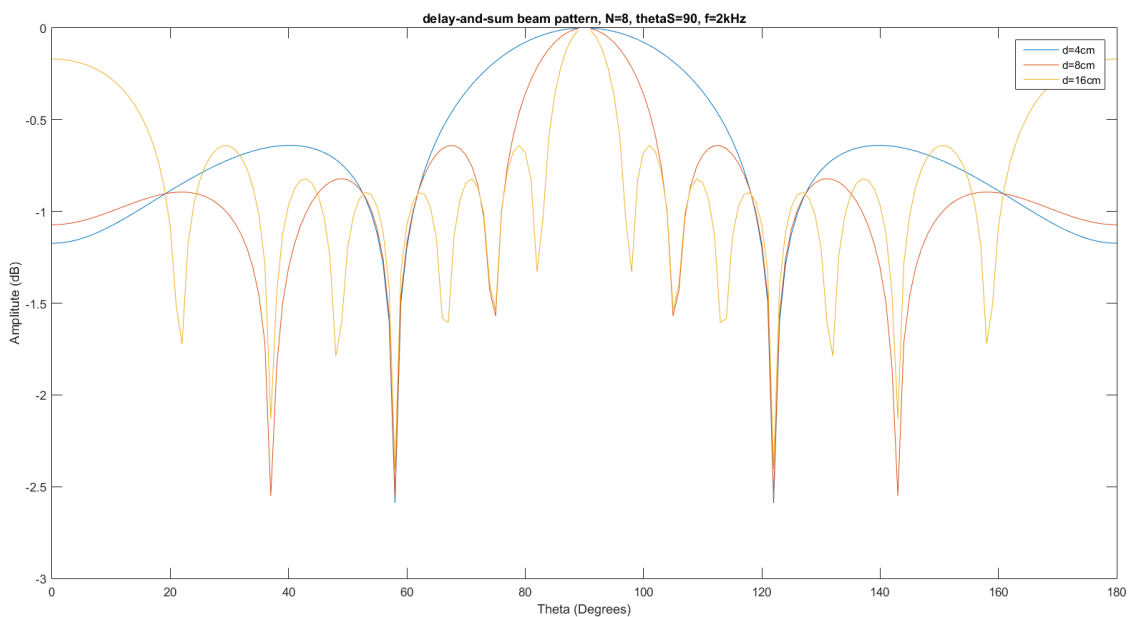
$$B(\omega, \theta) = \frac{1}{N} \frac{\sin\left[\frac{N}{2} \frac{\omega}{c} d (\cos \theta - \cos \theta_s)\right]}{\sin\left[\frac{1}{2} \frac{\omega}{c} d (\cos \theta - \cos \theta_s)\right]} \quad (19)$$

τον οποίο και υλοποιήσαμε σε κώδικα Matlab. Το αποτέλεσμα που προέκυψε για $\theta_s=90$, $f=2\text{kHz}$, $c=340\text{m/s}$ και άρα $\omega=2\pi f$, συναρτήσει της γωνίας θ στο $[0,180]$ είναι:

Για $d=4\text{cm}$:



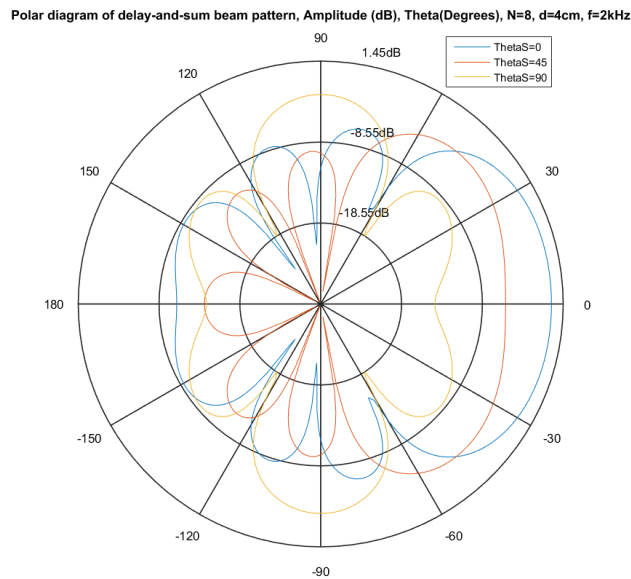
Για $N=8$:



Παρατηρούμε ότι για δεδομένη απόσταση μεταξύ των μικροφώνων, αυξάνοντας το πλήθος τους καταφέρνουμε να έχουμε όλο και πιο στενό λοβό για την γωνία θ_s , γεγονός που είναι επιθυμητό αφού θέλουμε οι τιμές των σημάτων για τις υπόλοιπες γωνίες να εξασθενούν όσο το δυνατόν περισσότερο.

Για δεδομένο πλήθος μικροφώνων, αυξάνοντας την μεταξύ τους απόσταση στενεύει περισσότερο ο κεντρικός λοβός. Ωστόσο, για $d=16\text{cm}$ παρατηρούμε την εμφάνιση ισχυρών πλευρικών λοβών στις γωνίες 0 και 180, γεγονός που θα προκαλέσει, ενδεχομένως, εισαγωγή στο αποτέλεσμα ανεπιθύμητων σημάτων (θόρυβο). Εκτός αν γνωρίζουμε ότι δεν θα αποτελέσουν πρόβλημα οι γωνίες αυτές, προτιμητέο θα ήταν να θέσουμε απόσταση $d=8\text{cm}$ στην περίπτωση των 8 μικροφώνων.

Υπολογίζουμε τώρα τα λογαριθμικά διαγράμματα πλάτους σε πολικές συνταγμένες όταν $N=8$, $d=4\text{cm}$, $f=2\text{kHz}$, συναρτήσει της γωνίας θ στο $[-180,180]$, για $\theta_s=0$, 45 και 90 μοίρες αντίστοιχα:



Παρατηρούμε ότι για $\theta_s=90$ το πλάτος του beam pattern εξασθενεί αρκετά, σε σχέση με τις υπόλοιπες θ_s , σε γωνίες διαφορετικές του 90, και ιδιαίτερα σε αυτές που είναι εκτός του διαστήματος (60,120). Από την άλλη, για $\theta_s=0$ το πλάτος επιτρέπει τη διέλευση, σε σχετικά μεγάλο βαθμό, σημάτων για γωνίες στο (0,60), ενώ για $\theta_s=45$ εισάγει σήματα σε όλο το (0,60) αλλά και σε ένα μέρος του (60,90). Σε αυτές τις περιπτώσεις, είτε η ανίχνευση γωνίας προέλευσης είτε η αποθρομβοποίηση είναι αρκετά πιο δύσκολη. Προτιμητέο λοιπόν είναι να τοποθετούμε τη συστοιχία μικροφώνων κάθετα στην πηγή του ήχου για καλύτερα αποτελέσματα. Παρατηρούμε, επίσης, ότι το διάγραμμα είναι συμμετρικό ως προς τον οριζόντιο άξονα, δηλαδή το beam pattern επιτρέπει ακριβώς με τον ίδιο τρόπο τη διέλευση σημάτων στις αρνητικές γωνίες. Έτσι, ακόμα και για $\theta_s=90$, θα έχουμε εισαγωγή θορύβου από τη γωνία $\theta=-90$.

Άσκηση 2.1.A: ex2_1_A.m

1) Υλοποιούμε το φίλτρο από το οποίο θα περάσουμε το κάθε σήμα χρησιμοποιώντας το διάνυσμα βαρών:

$$\mathbf{H}_{DS}^T(\omega) = \frac{1}{N} \mathbf{d}^H(\mathbf{k}_s), \quad \mathbf{k}_s = \frac{\omega}{c} \mathbf{a}_s, \quad (10)$$

και αφού έχουμε ομοιόμορφη γραμμική συστοιχία, χρησιμοποιούμε τον τύπο:

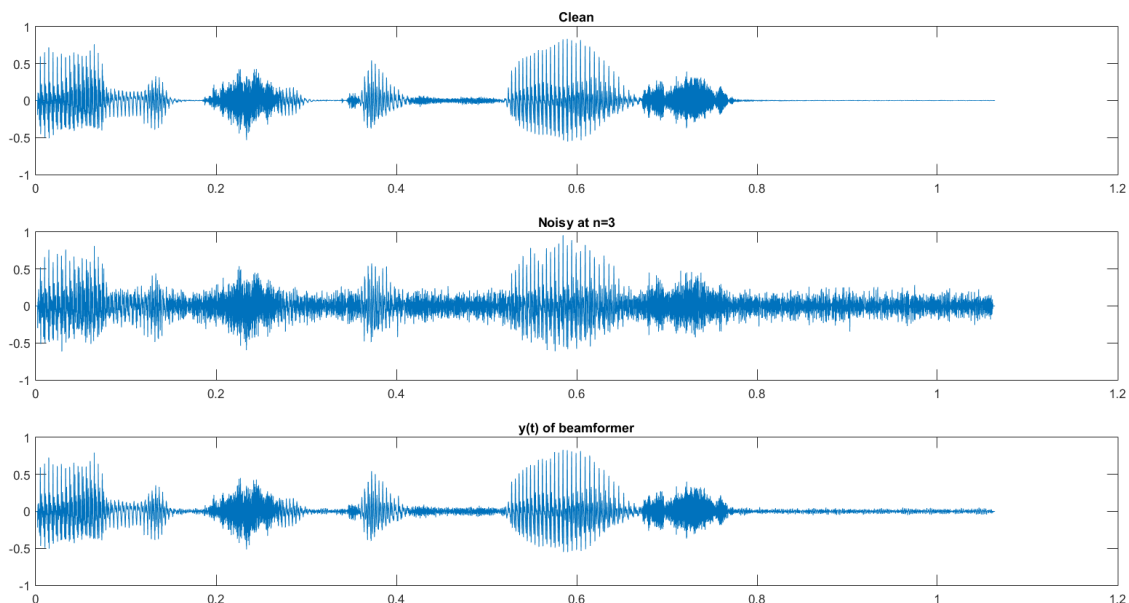
$$\mathbf{d}(\mathbf{k}) = e^{-j \frac{N-1}{2} \frac{\omega}{c} d \cos \theta} [1, e^{j \frac{\omega}{c} d \cos \theta}, \dots, e^{j(N-1) \frac{\omega}{c} d \cos \theta}]^T \quad (17)$$

Αναπαριστούμε το διάνυσμα με την μορφή 2διάστατου πίνακα N γραμμών που αντιστοιχούν στα μικρόφωνα και N_s στηλών (N_s το πλήθος των δειγμάτων) που αντιστοιχούν στις διακριτές συχνότητες. Λόγω της φύσης του DFT, ο οποίος υπολογίζει τις τιμές του μετασχηματισμού στο πρώτο μισό του για θετικές συχνότητες ενώ στο δεύτερο μισό του για αρνητικές συχνότητες, κατασκευάσαμε ανάλογα διάνυσμα συχνοτήτων ω ως:

$$\left[0, \frac{2\pi \cdot 1}{N_s} F_s, \frac{2\pi \cdot 2}{N_s} F_s, \dots, \frac{2\pi \cdot \left(\frac{N_s}{2}-2\right)}{N_s} F_s, \frac{2\pi \cdot \left(\frac{N_s}{2}-1\right)}{N_s} F_s, \frac{2\pi \cdot \left(-\frac{N_s}{2}\right)}{N_s} F_s, \frac{2\pi \cdot \left(-\frac{N_s}{2}+1\right)}{N_s} F_s, \dots, \frac{2\pi \cdot (-2)}{N_s} F_s, \frac{2\pi \cdot (-1)}{N_s} F_s \right].$$

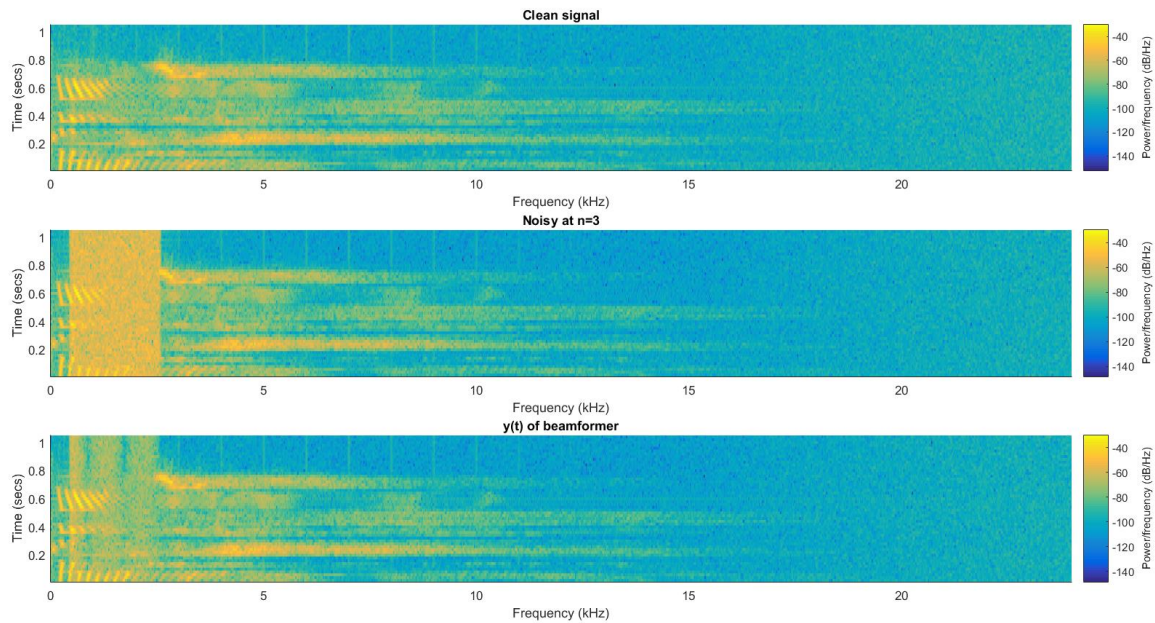
Εφαρμόζουμε το φίλτρο για κάθε ένα σήμα και για κάθε διακριτή συχνότητα σε δομή επανάληψης, και ύστερα υπολογίζουμε τον αντίστροφο DFT για το καθένα, λαμβάνοντας τον συνολικό μέσο όρο ως προς την διάσταση που αντιστοιχεί στην αρίθμηση των μικροφώνων για να λάβουμε την έξοδο του beamformer. Σημειώνουμε ότι λόγω αριθμητικών σφαλμάτων προκύπτει μιγαδικό μέρος κατά την εφαρμογή του αντιστρόφου. Επομένως, παίρνουμε το πραγματικό μέρος, το οποίο αντιστοιχεί ουσιαστικά στο μέσο όρο του άρτιου και περιττού μέρους του πλάτους στο πεδίο συχνότητας, το οποίο είναι κατά προσέγγιση συμμετρικό λόγω της άρτιας συμμετρίας του διανύσματος βαρών, αλλά και λόγω του ότι τα σήματα εισόδου είναι πραγματικά (άρα άρτιος μετασχηματισμός). Στη συνέχεια, ελέγχουμε αν είναι επιτυχής η ευθυγράμμιση ακούγοντας τη διαφορά του καθαρού από το θορυβώδες σήμα.

2) Οι κυματομορφές που προέκυψαν είναι:



Παρατηρούμε ότι η έξοδος του beamformer ακολουθεί σε μεγάλο βαθμό το καθαρό σήμα.

Και τα σπεκτρογραφήματα:



Στο πεδίο συχνοτήτων, λοιπόν, ο θόρυβος που εμφανίζεται στο εύρος [500Hz, 2.5kHz] έχει αρκετά μικρότερο πλάτος στην έξοδο του φίλτρου, που είναι και το επιθυμητό.

3) Υπολογίζουμε το snr σε κώδικα Matlab χρησιμοποιώντας την έτοιμη συνάρτηση $\text{snr}(x,y)$, όπου x το θορυβώδες σήμα και y το σήμα θορύβου. Στο κεντρικό μικρόφωνο η τιμή του είναι: 1.4507 dB, ενώ στην έξοδο του beamformer είναι: 17.1992 dB. Παρατηρούμε βελτίωση κατά 15.7485 dB, που δηλώνει ότι έχουμε επιτύχει αφαίρεση του θορύβου σε μεγάλο βαθμό.

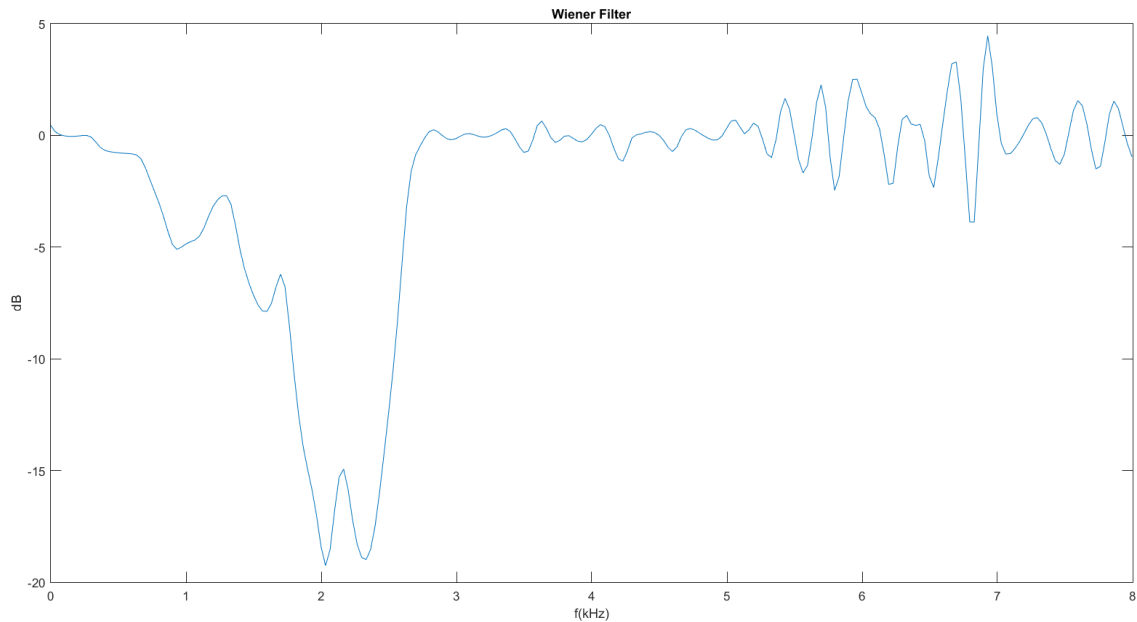
Έχουμε αποθηκεύσει, όπως ζητείται, το αποτέλεσμα με όνομα: “**sim_ds.wav**”.

Άσκηση 2.1.B: ex2_1_B.m

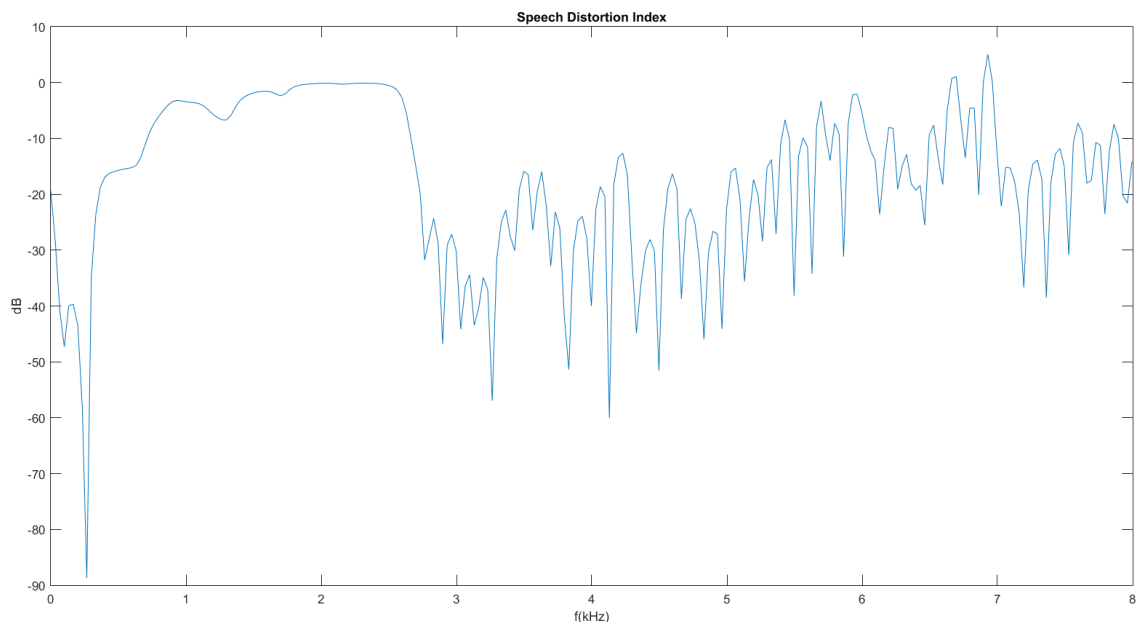
1) Υπολογίζουμε την απόκριση συχνότητας του φίλτρου Wiener (στο πεδίο συχνοτήτων) χρησιμοποιώντας το φάσμα του καθαρού σήματος και του θορυβώδους και τον τύπο:

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

Το φάσμα το εκτιμήσαμε με τη χρήση της συνάρτησης `pwelch`, χωρίς να ορίσουμε εμείς το μήκος του παραθύρου και το βήμα επικάλυψης. Φροντίσαμε όμως να προκύψει ολόκληρη η απεικόνιση του φάσματος (θετικών και αρνητικών συχνοτήτων) ώστε να μην προκύψουν παρακάτω προβλήματα κατά τον υπολογισμό του αντίστροφου μετασχηματισμού. Η απεικόνιση της απόκρισης συχνότητας στο διάστημα [0, 8kHz] είναι:

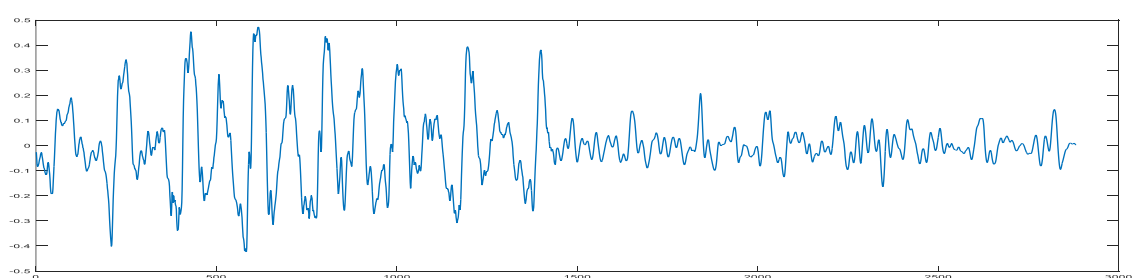


2) Υπολογίζουμε το speech distortion index και έχουμε:

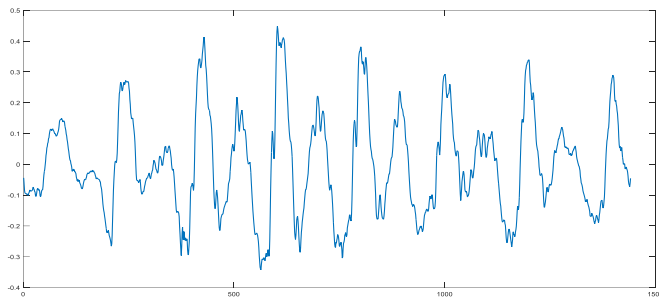


Παρατηρούμε ότι στις συχνότητες [500Hz,2.5kHz] έχουμε ομοιόμορφα την υψηλότερη παραμόρφωση, κάτι που είναι επιθυμητό καθώς σε αυτό το διάστημα βρίσκεται ο θόρυβος.

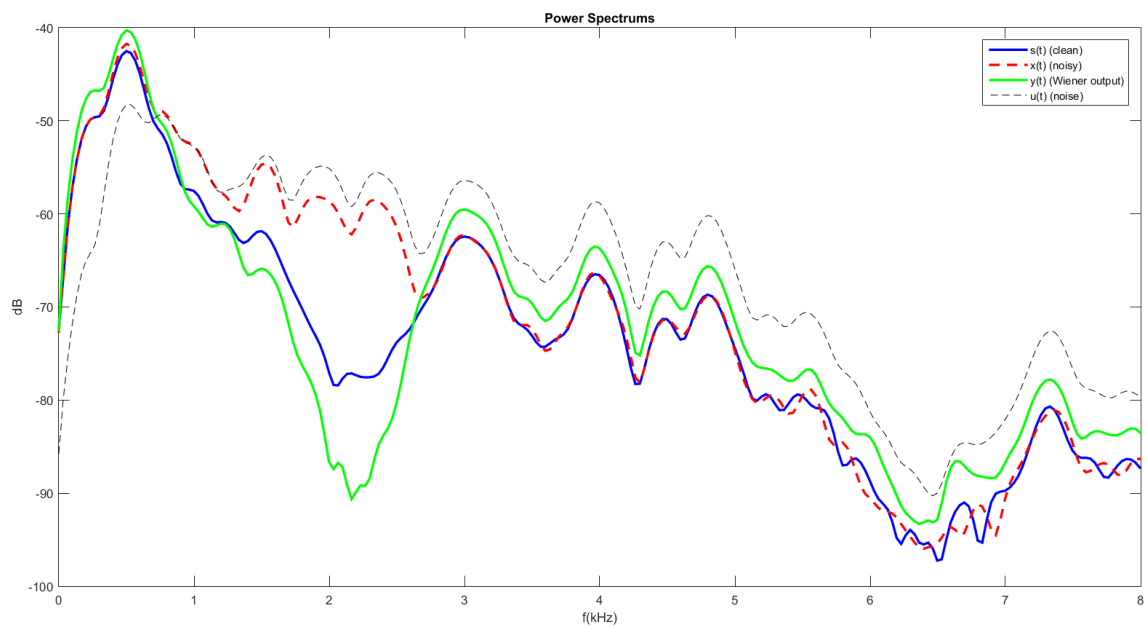
3) Εφαρμόζουμε το φίλτρο στο σήμα του κεντρικού μικροφώνου με τη μορφή γινομένου στο πεδίο της συχνότητας, για να αποφύγουμε τυχόν προβλήματα που θα προκύψουν εάν εφαρμόσουμε συνέλιξη στον χρόνο και αλλάξει το μήκος του προκύπτουν διανύσματος. Βέβαια, το γινόμενο DFT αντιστοιχεί σε κυκλική συνέλιξη που προκαλεί aliasing στον χρόνο. Στην περίπτωση μας δεν είναι σημαντικά εμφανές καθώς εκτός του διαστήματος που μας ενδιαφέρει, το πλάτος της γραμμικής συνέλιξης εξασθενεί:



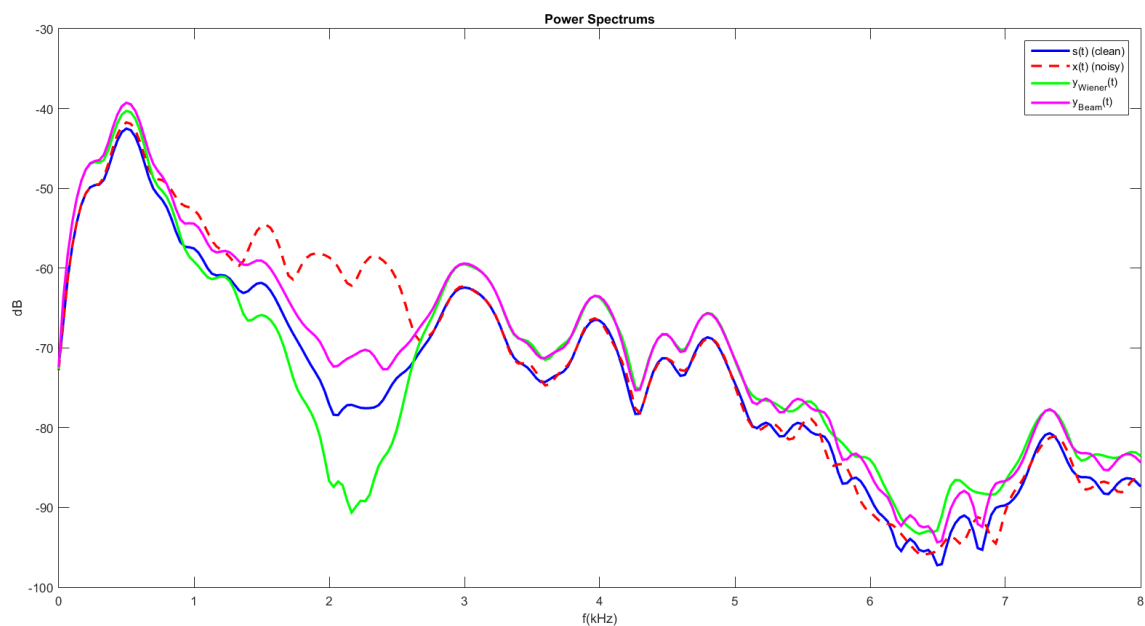
Οπότε προκύπτει με Inverse DFT:



Τα ζητούμενα φάσματα ισχύος είναι:



4) Το snr της εισόδου $x(t)$ είναι: 5.6005 dB, ενώ της εξόδου του Wiener φίλτρου είναι: 6.2345 dB. Παρατηρούμε ότι για το συγκεκριμένο πλαίσιο του σήματος φωνής η αφαίρεση του θορύβου δεν είναι ικανοποιητική. Τα ζητούμενα φάσματα για το συγκεκριμένο χρονικό διάστημα είναι:

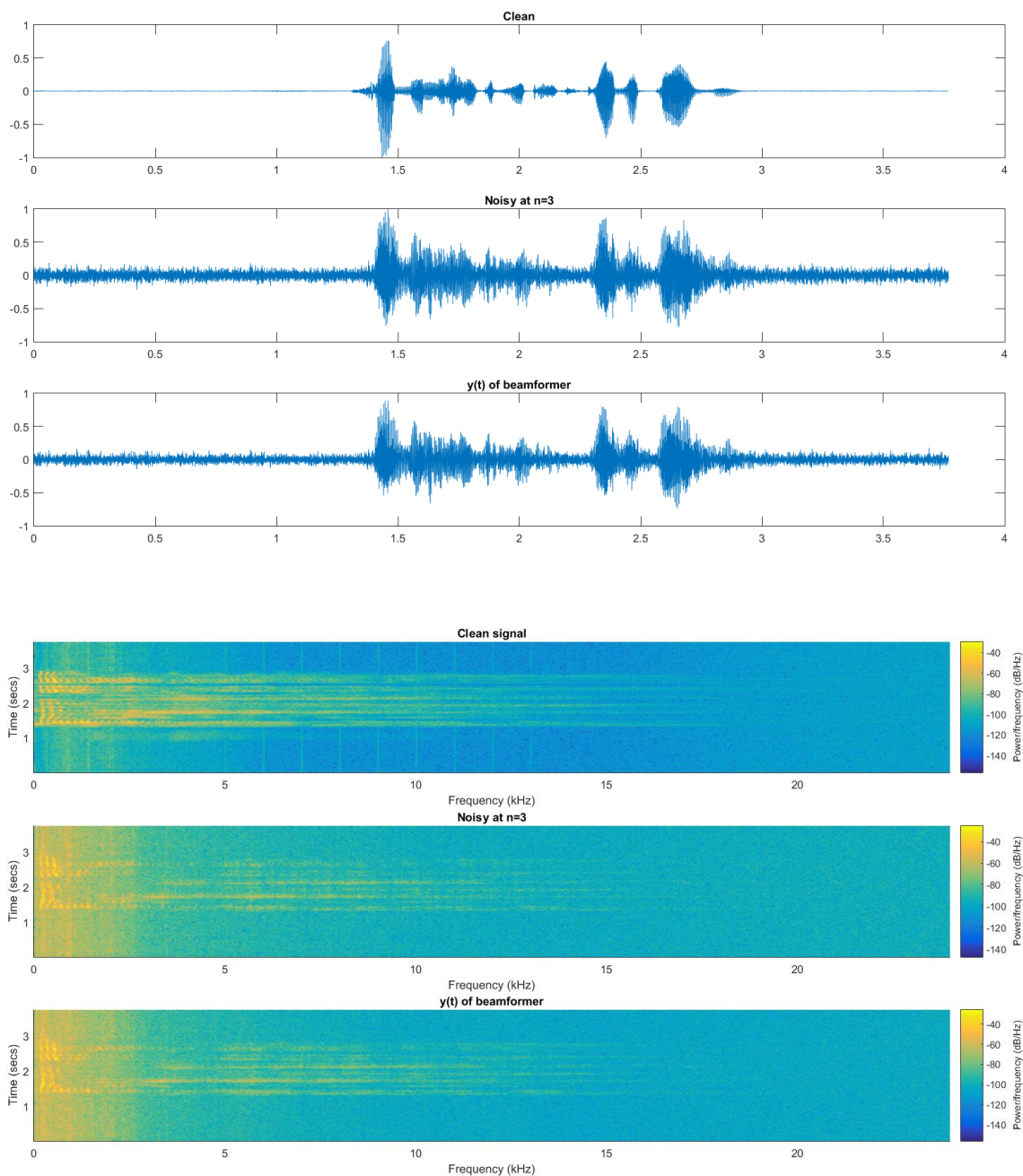


Παρατηρούμε ότι στο διάστημα συχνοτήτων που είναι συγκεντρωμένος ο θόρυβος, η έξοδος του beamformer ακολουθεί πιο πιστά το καθαρό σήμα, συγκριτικά με την έξοδο του φίλτρου Wiener. Θα καταλήγαμε στο

συμπέρασμα ότι προτιμάμε την πολυκαναλική μέθοδο στην περίπτωση που ο θόρυβος προέρχεται από συγκεκριμένη γωνία διαφορετική της γωνίας πηγής, εκμεταλλευόμενοι την γεωμετρία του χώρου.

Άσκηση 2.2.A: ex2_2_A.m

1,2) Με τον ίδιο τρόπο της προηγούμενης άσκησης εφαρμόζουμε το beamforming. Τα διαγράμματα που προκύπτουν είναι:



Παρατηρούμε τόσο από την κυματομορφή όσο και από το σπεκτρογράφημα ότι δεν έχουμε αποτελεσματική μείωση του θορύβου. Στο πεδίο του χρόνου το αποτέλεσμα δεν διαφέρει ιδιαίτερα από το θορυβώδες σήμα, ενώ στο πεδίο συχνότητας ο θόρυβος εξακολουθεί να παρουσιάζει υψηλές τιμές στις ίδιες συχνότητες με το θορυβώδες σήμα. Σημειώνουμε πως ακούγοντας το αποτέλεσμα της διαφοράς καθαρού σήματος και έξοδο του beamformer μπορούσαμε να διακρίνουμε ομιλία, που σημαίνει ότι δεν έχει επιτευχθεί καλή ευθυγράμμιση.

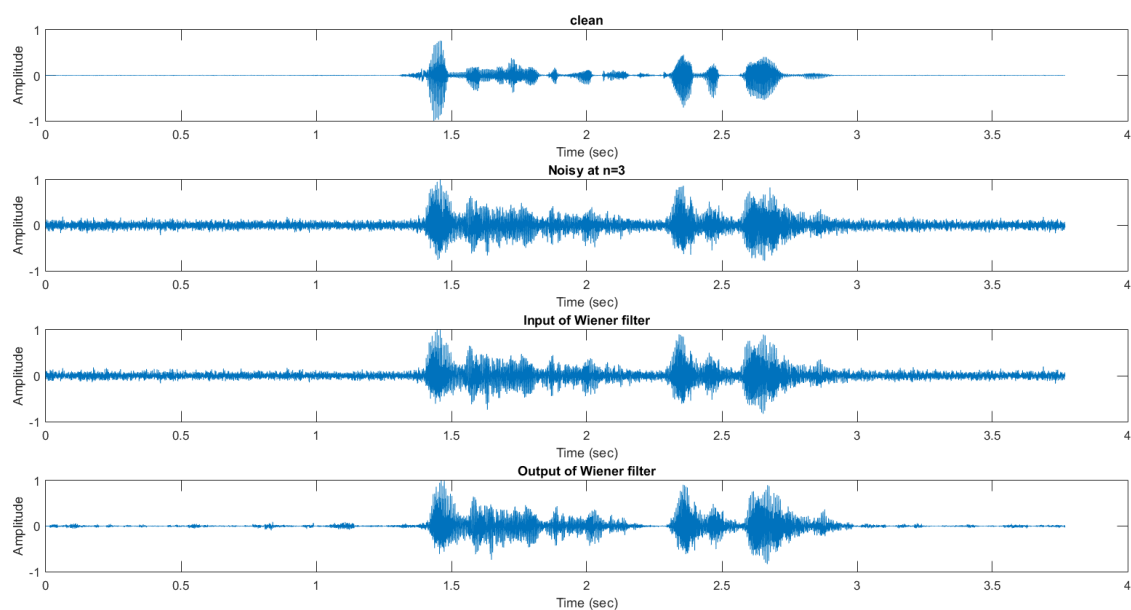
3) Για να υπολογίσουμε το ζητούμενο snr σπάμε το καθαρό σήμα, το σήμα θορύβου στο μικρόφωνο και το σήμα θορύβου στην έξοδο του beamformer σε πλαίσια διάρκειας 40msec και 50% επικάλυψη. Επιλέγουμε τυχαία ένα πλαίσιο του θορύβου τόσο στο μικρόφωνο όσο και στο beamformer για να υπολογίσουμε το στάσιμο φάσμα τους. Για κάθε ένα από τα πλαίσια του καθαρού σήματος, υπολογίζουμε το snr του ως προς τον θόρυβο του μικροφώνου και, κρατώντας μόνο τις τιμές άνω του 0 και μετατρέποντας τιμές άνω των 35 dB σε 35 dB, υπολογίζουμε τον μέσο όρο. Εφαρμόζουμε τον ίδιο αλγόριθμο και για τον θόρυβο του beamformer. Προέκυψε ότι στο μικρόφωνο το snr είναι: 6.2716 dB, και στην έξοδο του φίλτρου: 7.3835 dB. Η βελτίωση δεν είναι ικανοποιητική, οπότε καταλήγουμε με βάση και αυτήν την μετρική ότι στην περίπτωση του diffuse noise field η μέθοδος αποθρομβοποίησης με delay-and-sum beamformer δεν έχει καλή απόδοση.

Αποθηκεύσαμε το αποτέλεσμα στο αρχείο με όνομα: "real_ds.wav".

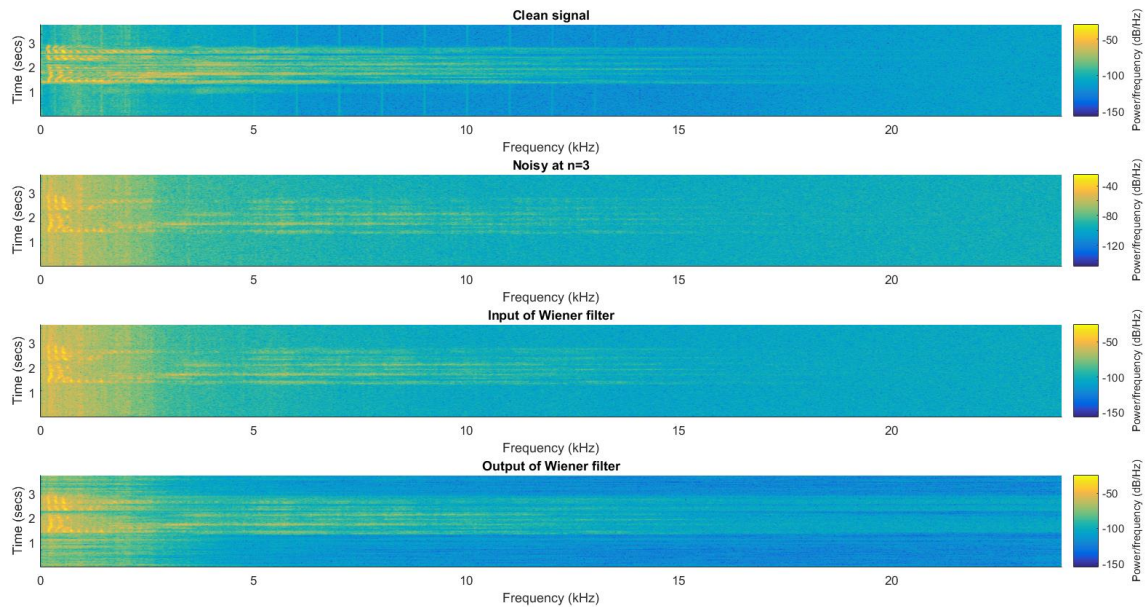
Άσκηση 2.2.B: ex2_2_B.m

1) Χρησιμοποιούμε διάρκεια 0,030sec για τα πλαίσια, άρα $L=0.030*F_s=1920$, και βήμα $L/2$. Αφού έχουμε παράθυρο hamming και το L δεν είναι περιττό (που δημιουργεί πρόβλημα για $R=(L-1)/2$), θα έχουμε αποτελεσματική ανασύνθεση. Υπολογίζουμε για το κάθε πλαίσιο την απόκριση της συχνότητας του Wiener φίλτρου, έχοντας πρώτα υπολογίσει στο χρονικό διάστημα $[0, 0.6\text{sec}]$, στο οποίο ακούγεται μόνο θόρυβος, το φάσμα ισχύος. Σημειώνουμε πως στο πλαίσιο θορύβου αυτό εφαρμόσαμε επίσης παράθυρο hamming ίδιου μήκους πρωτίστως, καθώς παρατηρήσαμε ότι ομαλοποιεί το άκουσμα του θορύβου στο αποτέλεσμα. Επίσης, για την εκτίμηση του φάσματος με την συνάρτηση `pwelch`, δώσαμε ως παράμετρο το μήκος των πλαισίων ίσο με 40 δείγματα και επικάλυψη 20, διότι δοκιμαστικά προέκυψε ομαλότερο άκουσμα. Ύστερα, εκτελούμε overlap - add με κώδικα παρόμοιο με αυτόν της 1^{ης} εργαστηριακής.

2) Τα διαγράμματα που προέκυψαν είναι:



Από τις κυματομορφές φαίνεται να έχει μειωθεί κάπως ο θόρυβος στην διάρκεια της ομιλίας, που είναι και το ζητούμενο.



Παρατηρούμε ότι στην έξοδο του Wiener φίλτρου, με βάση το σπεκτρογράφημα, το πλάτος του θορύβου έχει μειωθεί αρκετά, αφήνοντας να υπερισχύσει το σήμα φωνής.

3) Το $ssnr$ στην είσοδο είναι: 7.3835 dB, και στην έξοδο: 15.0046 dB. Παρατηρούμε αρκετή βελτίωση με το φιλτράρισμα Wiener.

4) Υπολογίζουμε το $ssnr$ των σημάτων εισόδου και βρίσκουμε: 0.0365 dB. Η αύξησή του με το φίλτρο Wiener είναι κατά 14.9682 dB, γεγονός που σημαίνει ότι έχουμε σημαντική βελτίωση της σηματοθορυβικής σχέσης.

Το αποτέλεσμα ήχου είναι ικανοποιητικό, αν και εξακολουθεί να ακούγεται ένας λιγότερο έντονος και αλλοιωμένος θόρυβος, σε σχέση με τον αρχικό. Παρατηρήσαμε, ακόμη, ακούγοντας το αποτέλεσμα, ότι στα σημεία ομιλίας ο θόρυβος «σιωπαίνει» περισσότερο. Σε κάθε περίπτωση, το delay-and-sum beamforming είναι πολύ λιγότερο αποδοτικό και δεν θα ήταν αρκετό.

Το αποτέλεσμα το αποθηκεύσαμε στο αρχείο **“real_mmse.wav”**.