

Ψηφιακή Επεξεργασία Σημάτων



1η Εργαστηριακή Άσκηση

Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών ΕΜΠ

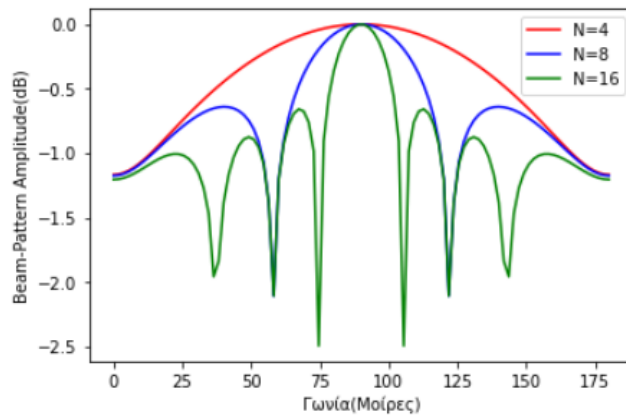
Εξάμηνο 6ο

Σερλής Εμμανουήλ Αναστάσιος, Α.Μ.:03118125, email: manosanastassis@hotmail.com

Δημήτρης Μητρόπουλος, Α.Μ.: 03118608, email: dimitris.7.mitropoulos@gmail.com

Άσκηση 1.5

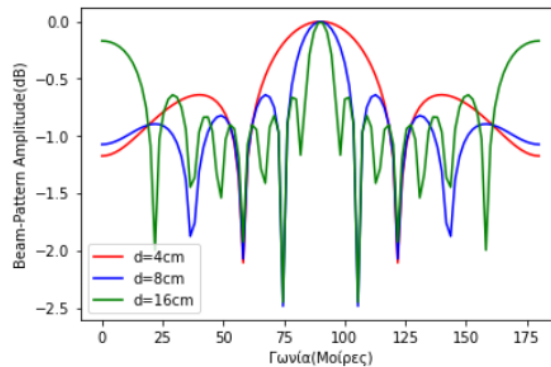
- 1) Για $d=4\text{cm}$ και μεταβαλλόμενο αριθμό μικροφώνων N



Παρατηρούμε ότι και για τις 3 τιμές του N έχουμε μέγιστη τιμή στις 90 μοίρες. Ωστόσο, για μικρές τιμές του N διαθέτουμε “εκτεταμένους” λοβούς, δηλαδή λοβούς οι οποίοι εκτείνονται και σε γωνίες διάφορες των 90 μοιρών. Αντίθετα, για $N=16$, ο κεντρικός λοβός σταματά πολύ πιο απότομα, γεγονός που συνεπάγεται αυξημένη ενίσχυση του σήματος των 90 μοιρών σε σύγκριση με τις πηγές διαφορετικών γωνιών.

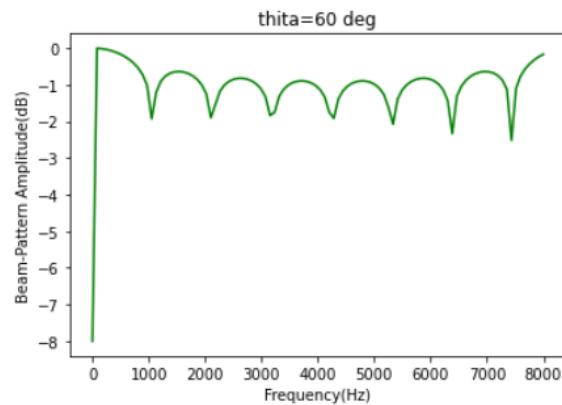
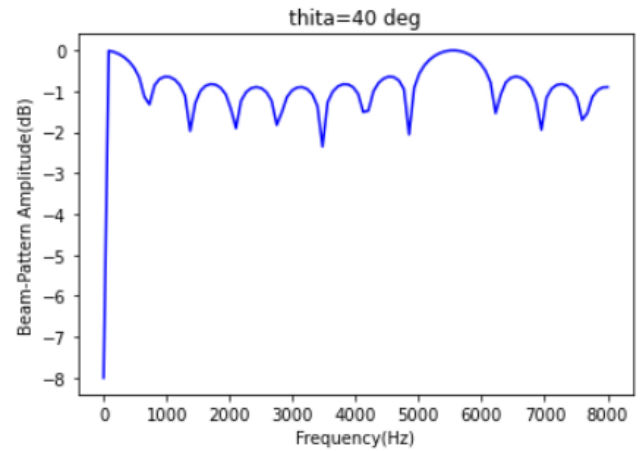
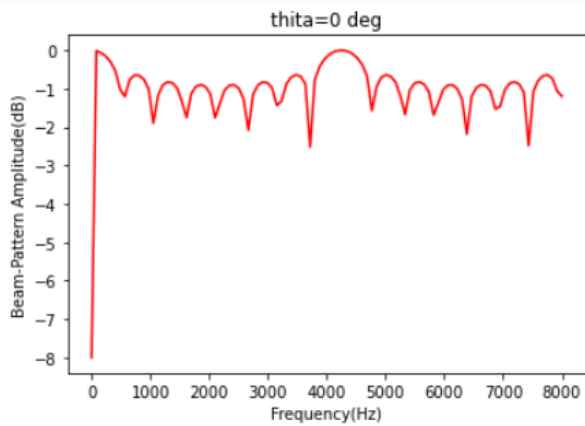
Προφανώς, προτιμούμε αυξημένο N , καθώς έτσι ενισχύεται το σήμα ενδιαφέροντος, που είναι αυτό των 90 μοιρών.

- 2) Κρατώντας $N=8$ και μεταβάλλοντας την απόσταση των μικροφώνων d



Στην λογική του ερωτήματος 1, επιθυμούμε απότομο κόψιμο του κεντρικού λοβού, που αντιστοιχεί σε ενισχυμένο σήμα πηγής. Αυτό προκύπτει από αυξημένο d , άρα ιδανικά θα επιλέγαμε $d=16\text{cm}$.

3) Κρατώντας σταθερά $N=8$ και $d=8\text{cm}$ και μεταβάλλοντας την γωνία θ_s :



Παρατηρούμε ότι αλλαγή της γωνίας θ_s αλλάζει την κύρια συχνότητα ενίσχυσης, μιας και μετατοπίζεται ο μέγιστος λοβός. Μάλιστα, αύξηση του θ_s σημαίνει και προς τα δεξιά μετατόπιση του λοβού, δηλαδή ενίσχυση όλο και μεγαλύτερων-μη θορυβικών-συχνοτήτων. Όσον αφορά την ενίσχυση σημάτων θορύβου από γωνίες διαφορετικές της θ_s , προφανώς αυτή η ενίσχυση θα είναι περιορισμένη μιας ο beamformer είναι προσαρμοσμένος να ενισχύει τα σήματα από μια συγκεκριμένη και προκαθορισμένη κατεύθυνση.

Άσκηση 2.1

2.1.A

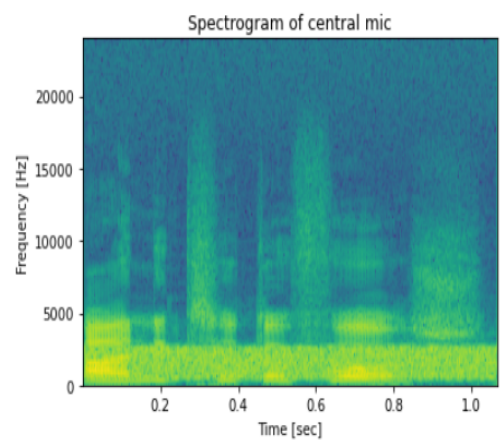
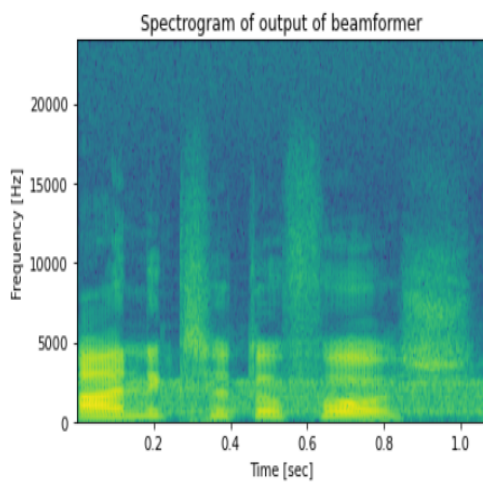
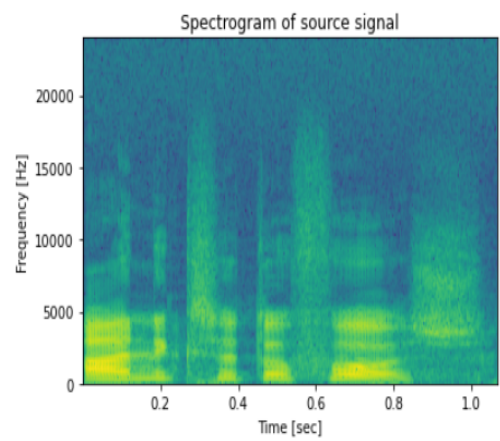
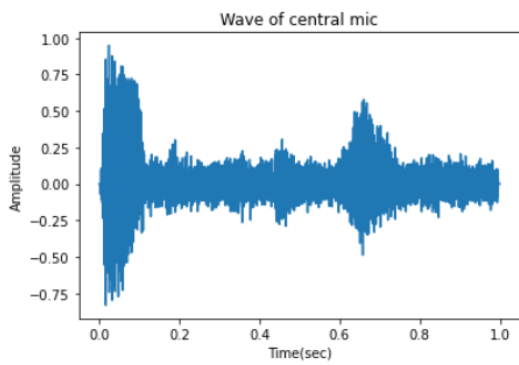
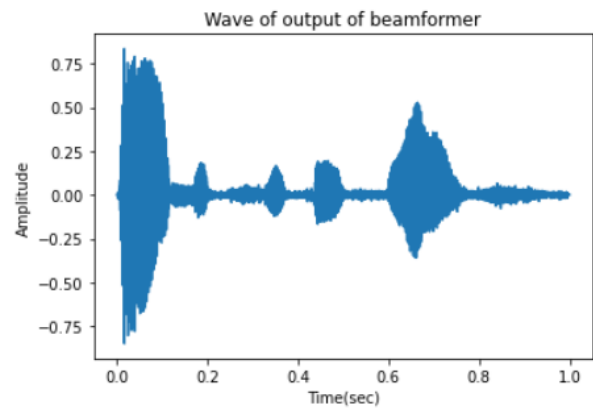
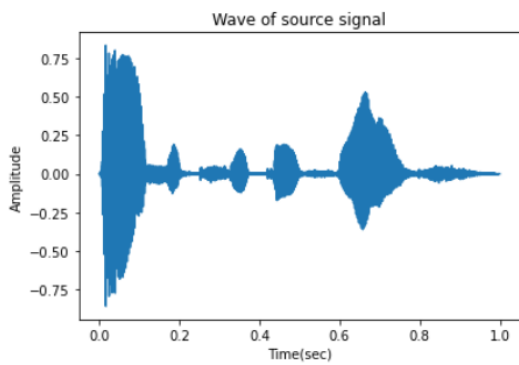
1) Η διαδικασία που ακολουθήθηκε (εν περιλήψει) είναι η εξής:

Αφού διαβάσουμε σε μορφή πινάκων να δοθέντα αρχεία ήχου, στην συνέχεια υπολογίζουμε τα ζητούμενα βάρη από την εξίσωση (10), τα οποία προκύπτουν μέσω της εξίσωσης (17) αν αντιστρέψουμε το πρόσημο των εκθετικών όρων (ώστε να προκύπτει ο συζυγής ανάστροφος του $d(k_s)$). Μετέπειτα, πολλαπλασιάζουμε τον DFT κάθε σήματος με την εξίσωση (17), υλοποιώντας έτσι την ζητούμενη μετατόπιση στο συχνοτικό πεδίο. Τέλος, μεταβαίνουμε στο πεδίου του χρόνου με αντίστροφο DFT και βρίσκουμε τον Μ.Ο. των σημάτων εξόδου.

Όσον αφορά την παράθεση των βαρών, αυτή είναι αδύνατον να γίνει πλήρως, μιας και αναφερόμαστε σε ένα data set μεγέθους $7*fs*1sec=336.000$ στοιχείων.

```
Τα βάρη είναι: [array([0.14285714+0.j          , 0.14285653+0.0004195j ,
0.14285468+0.00083899j, ..., 0.13779553-0.03769025j,
0.13790561-0.03728545j, 0.13801451-0.03688034j]), array([0.14285714+0.j          , 0.14285687+0.00027966j,
0.14285605+0.00055933j, ..., 0.1406001 -0.0252938j ,
0.14064934-0.02501851j, 0.14069805-0.02474312j]), array([0.14285714+0.j          , 0.14285707+0.00013983j,
0.14285687+0.00027966j, ..., 0.14229176-0.01269715j,
0.14230412-0.01255787j, 0.14231635-0.01241857j]), array([0.14285714+0.j, 0.14285714+0.j, 0.14285714+0.j, ...,
0.14285714+0.j, 0.14285714+0.j, 0.14285714+0.j]), array([0.14285714+0.j          , 0.14285707-0.00013983j,
0.14285687-0.00027966j, ..., 0.14229176+0.01269715j,
0.14230412+0.01255787j, 0.14231635+0.01241857j]), array([0.14285714+0.j          , 0.14285687-0.00027966j,
0.14285605-0.00055933j, ..., 0.1406001 +0.0252938j ,
0.14064934+0.02501851j, 0.14069805+0.02474312j]), array([0.14285714+0.j          , 0.14285653-0.0004195j ,
0.14285468-0.00083899j, ..., 0.13779553+0.03769025j,
0.13790561+0.03728545j, 0.13801451+0.03688034j]))]
[array([0.14285714+0.j          , 0.14285653+0.0004195j ,
0.14285468+0.00083899j, ..., 0.13779553-0.03769025j,
0.13790561-0.03728545j, 0.13801451-0.03688034j]), array([0.14285714+0.j          , 0.14285687+0.00027966j,
0.14285605+0.00055933j, ..., 0.1406001 -0.0252938j ,
0.14064934-0.02501851j, 0.14069805-0.02474312j]), array([0.14285714+0.j          , 0.14285707+0.00013983j,
0.14285687+0.00027966j, ..., 0.14229176-0.01269715j,
0.14230412-0.01255787j, 0.14231635-0.01241857j]), array([0.14285714+0.j, 0.14285714+0.j, 0.14285714+0.j, ...,
0.14285714+0.j, 0.14285714+0.j, 0.14285714+0.j]), array([0.14285714+0.j          , 0.14285707-0.00013983j,
0.14285687-0.00027966j, ..., 0.14229176+0.01269715j,
0.14230412+0.01255787j, 0.14231635+0.01241857j]), array([0.14285714+0.j          , 0.14285687-0.00027966j,
0.14285605-0.00055933j, ..., 0.1406001 +0.0252938j ,
0.14064934+0.02501851j, 0.14069805+0.02474312j]), array([0.14285714+0.j          , 0.14285653-0.0004195j ,
0.14285468-0.00083899j, ..., 0.13779553+0.03769025j,
0.13790561+0.03728545j, 0.13801451+0.03688034j]))]
```

2) Οι ζητούμενες γραφικές και τα σπεκτρογραφήματα:



3) Οι ζητούμενοι σηματοθορυβικοί λόγοι:

SNR of central mic(in) is: 11.102068519257195 dB

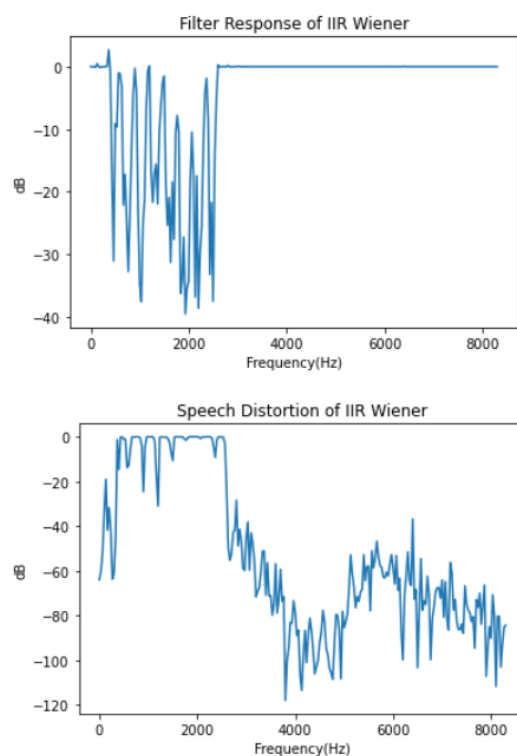
SNR of beamformer(out) is: 46.84653667083072 dB

Προφανώς, παρατηρούμε ότι η έξοδος του beamformer διαθέτει καλύτερο σηματοθορυβικό λόγο σε σχέση με το σήμα του κεντρικού μικροφώνου($n=3$), γεγονός που αποδεικνύει την αποθορυβοποίηση που πραγματοποιεί η εν λόγω μέθοδος.

Μάλιστα, αυτό επαληθεύεται και από τα αποτελέσματα του ερωτήματος 2, μιας τα σπεκτρογραφήματα της εξόδου και της πηγής δεν διαθέτουν την συνεχή κίτρινη “λωρίδα” θορύβου που διαθέτει το κεντρικό μικρόφωνο μεταξύ των συχνοτήτων [0,2kHz].

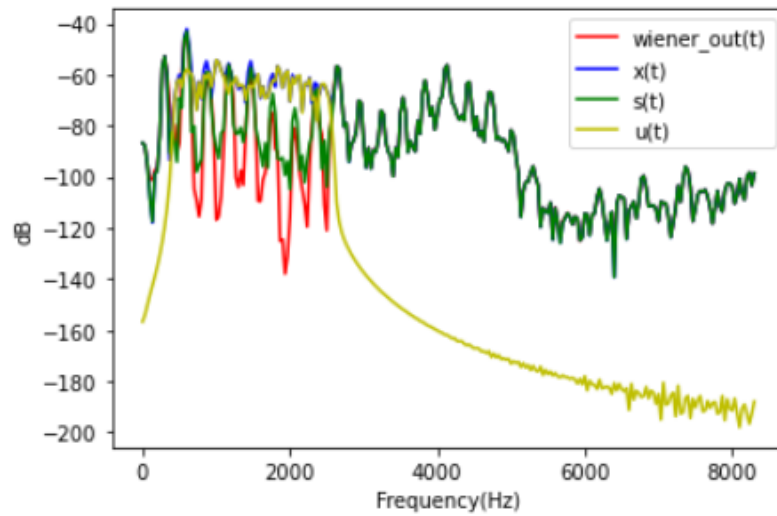
2.1.B

1) Οι ζητούμενες γραφικές παραστάσεις:



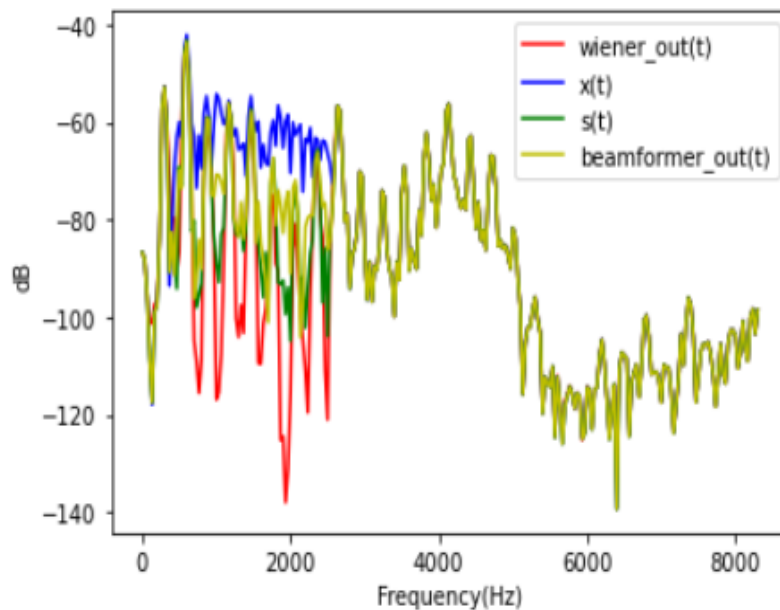
Παρατηρούμε ότι η τιμή του $n_{sd}(\omega)$ είναι αντιστρόφως ανάλογη της συχνότητας, γεγονός που ισοδυναμεί με κακή απόδοση του φίλτρου στις χαμηλές συχνότητες, όπου θα παρατηρείται αυξημένη παραμόρφωση. Αυτό είναι εύλογο αν αναλογισθούμε ότι ο θόρυβος αποτελεί ένα σήμα χαμηλών συχνοτήτων, οπότε προφανώς είναι δυσκολότερη η αποθορυβοποίηση του εκάστοτε σήματος εισόδου στις χαμηλές συχνότητες από ότι στις υψηλές.

2) Η-από κοινού- γραφική παράσταση των ζητούμενων σημάτων



Παρατηρούμε ότι η έξοδος του wiener φίλτρου (κόκκινο) είναι προβληματική στις χαμηλές συχνότητες, όπου αδυνατεί να προσεγγίσει ικανοποιητικά το σήμα της πηγής (με πράσινο). Αυτό οφείλεται στην ύπαρξη θορύβου στις εν λόγω συχνότητες, τον οποίον το φίλτρο αδυνατεί να αποκόψει πλήρως.

3) Η-από κοινού- γραφική παράσταση των ζητούμενων σημάτων



4) Τα ζητούμενα SNR:

SNR of $x(t)$ (in) is: 12.13638873281335 dB

SNR of wiener filter(out) is: 16.70027809806258 dB

SNR of sample of beamformer(out) is: 47.24172611082214 dB

Τα αποτελέσματα του ερωτήματος επαληθεύουν την ανωτερότητα του beamformer έναντι του wiener φίλτρου, μιας και η έξοδος του πρώτου όχι μόνο έχει μεγαλύτερο SNR από την έξοδο του δεύτερου, αλλά και η γραφική της ταυτίζεται σε πολύ μεγαλύτερο βαθμό με την γραφική του σήματος πηγής σε σχέση με την γραφική του wiener_out(t). Όπως έχει ήδη αναφερθεί, η εν λόγω διαφορά στην επίδοση των δύο φίλτρων είναι ανάλογη της ικανότητας αποκοπής της θορυβικής συνιστώσας που επικρατεί στις χαμηλές συχνότητες.

Άσκηση 2.2

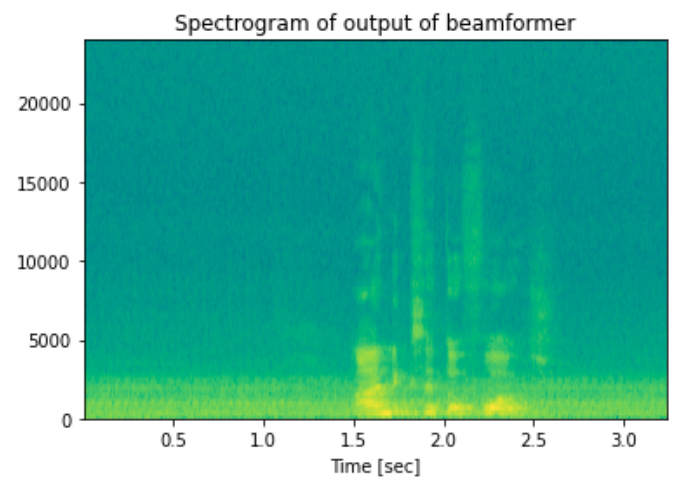
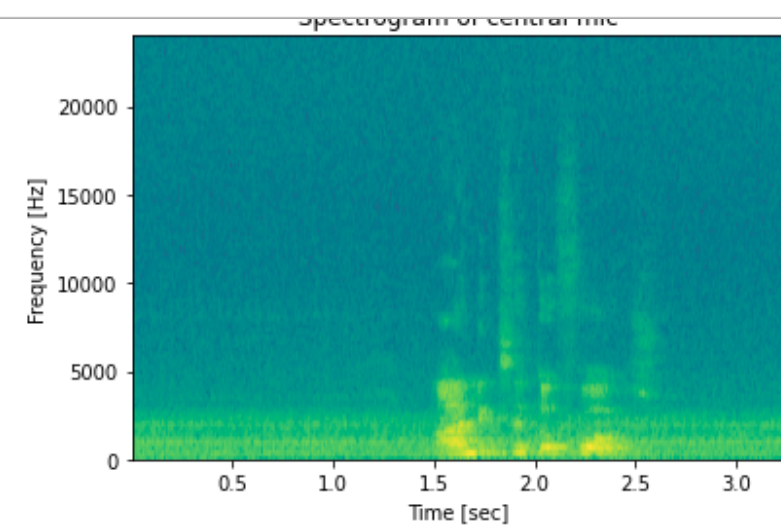
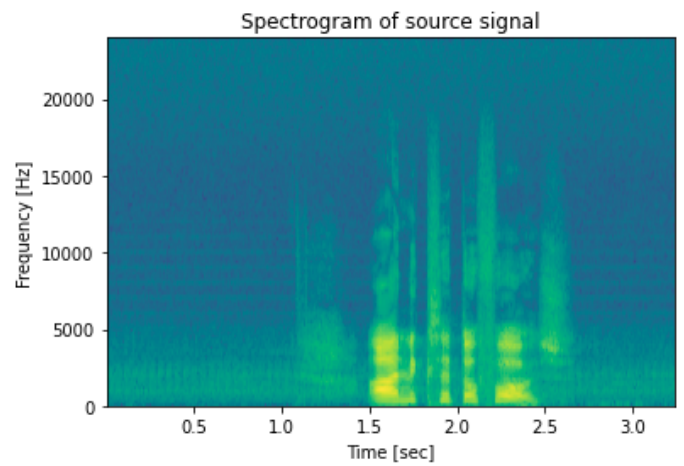
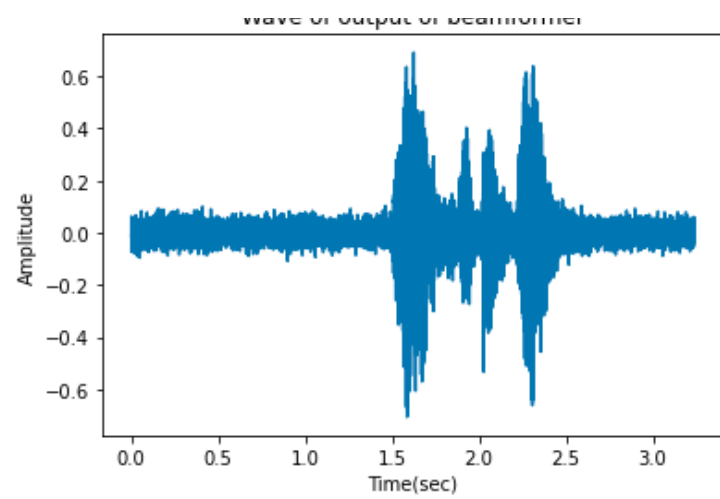
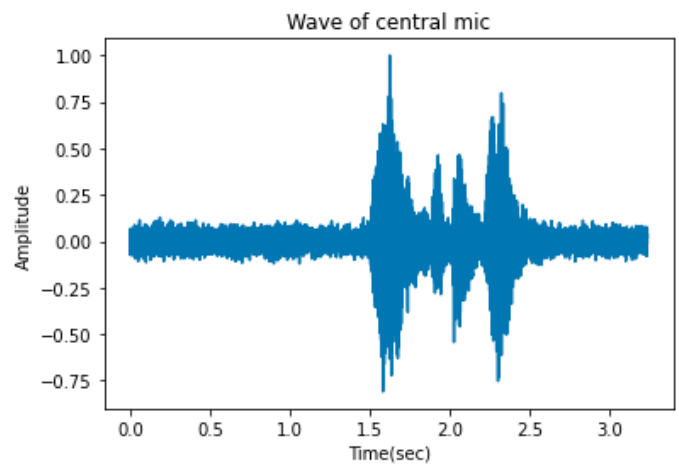
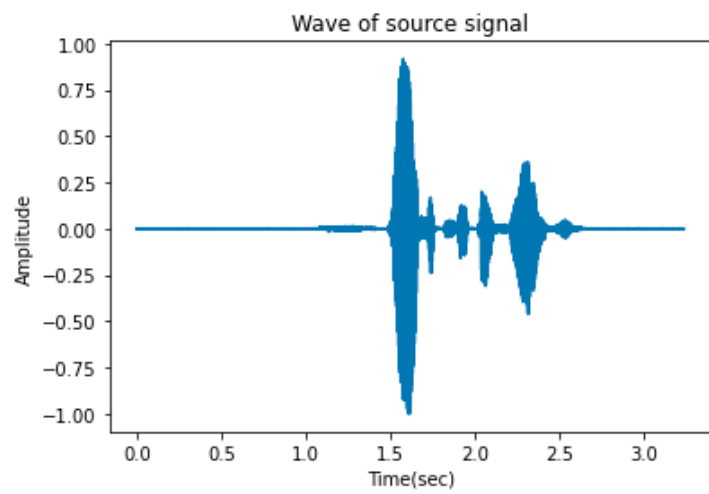
2.2.A

1) Αντίστοιχα με το 2.1.A, υπολογίζουμε τα βάρη από την εξίσωση 10 τα οποία προκύπτουν μέσω της εξίσωσης (17) αν αντιστρέψουμε το πρόσημο των εκθετικών όρων (ώστε να προκύπτει ο συζυγής ανάστροφος του $d(ks)$). Μετέπειτα, πολλαπλασιάζουμε τον DFT κάθε σήματος με την εξίσωση (17), υλοποιώντας έτσι την ζητούμενη μετατόπιση στο συχνοτικό πεδίο. Τέλος, μεταβαίνουμε στο πεδίου του χρόνου με αντίστροφο DFT και βρίσκουμε τον Μ.Ο. των σημάτων εξόδου.

Τα ζητούμενα βάρη

```
Τα βάρη είναι:
[array([0.14285714+0.00000000e+00j, 0.14285713+6.91458345e-05j,
        0.14285708+1.38291653e-04j, ..., 0.14162324-1.87355619e-02j,
        0.14163229-1.86670111e-02j, 0.14164131-1.85984559e-02j]), array([0.14285714+0.00000000e+00j, 0.14285714+4.6097
2240e-05j,
        0.14285711+9.21944432e-05j, ..., 0.1423083 -1.25104146e-02j,
        0.14231233-1.24644938e-02j, 0.14231635-1.24185717e-02j]), array([0.14285714+0.00000000e+00j, 0.14285714+2.3048
6123e-05j,
        0.14285714+4.60972240e-05j, ..., 0.14271987-6.26122390e-03j,
        0.14272087-6.23819735e-03j, 0.14272188-6.21517064e-03j]), array([0.14285714+0.j, 0.14285714+0.j, 0.14285714+0.
j, ...,
        0.14285714+0.j, 0.14285714+0.j, 0.14285714+0.j]), array([0.14285714+0.00000000e+00j, 0.14285714-2.30486123e-05
j,
        0.14285714-4.60972240e-05j, ..., 0.14271987+6.26122390e-03j,
        0.14272087+6.23819735e-03j, 0.14272188+6.21517064e-03j]), array([0.14285714+0.00000000e+00j, 0.14285714-4.6097
2240e-05j,
        0.14285711-9.21944432e-05j, ..., 0.1423083 +1.25104146e-02j,
        0.14231233+1.24644938e-02j, 0.14231635+1.24185717e-02j]), array([0.14285714+0.00000000e+00j, 0.14285713-6.9145
8345e-05j,
        0.14285708-1.38291653e-04j, ..., 0.14162324+1.87355619e-02j,
        0.14163229+1.86670111e-02j, 0.14164131+1.85984559e-02j])]
```

2) Οι ζητούμενες γραφικές και τα σπекτρογραφήματα



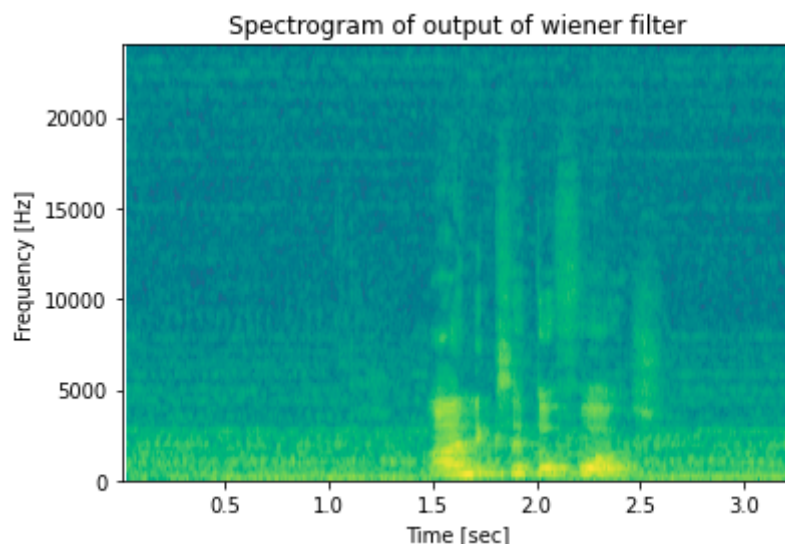
3) Για το SSNR χρησιμοποιήθηκε ο τύπος (22), βάση του οποίου υπολογίστηκε το SSNR κάθε πλαισίου των 25ms ξεχωριστά και μετά υπολογίστηκε ο μέσος όρος. Επίσης βάση του τύπου ορίζουμε μέγιστη τιμή τα 35db και κατώτατη τα -5db (ως κατώφλι). Οι τιμές των SSNR είναι:

```
SNR of central mic(in) is: 4.713797551672598  
SNR of beamformer(out) is: 3.200677968816983
```

Παρατηρούμε ότι το SSNR του beamformer είναι χειρότερο από αυτό του κεντρικού μικροφώνου. Αυτό σημαίνει πως το beamformer στην συγκεκριμένη περίπτωση δεν δουλεύει όπως θα θέλαμε και άρα χρειάζεται καλύτερο φιλτράρισμα.

2.2.B

2) Τα σπεκτρογραφήματα από το (α)-(γ) βρίσκονται στο ερώτημα (2.2.A.2). Το σπεκτρογράφημα της εξόδου του φίλτρου είναι το εξής:



Από το σπεκτρογράφημα μπορούμε να παρατηρήσουμε πως η έξοδος του φίλτρου είναι αρκετά βελτιωμένη από αυτή του beamformer (αν και ο θόρυβος δεν έχει εξαλειφθεί τελείως).

3) Παρατηρούμε από την εικόνα παρακάτω πως υπάρχει μια μεγάλη βελτίωση στον δείκτη SSNR (σχεδόν διπλασιασμός). Σε σύγκριση και με το SSNR του 4ου μικροφώνου παρατηρείται σημαντική βελτίωση στην εξάλειψη του θορύβου από το diffuse noise field και άρα το μοντέλο μας είναι σαφώς πιο επιδραστικό.

```
SSNR of wiener(in) is: 3.200677968816983  
SSNR of wiener(out) is: 5.679030057115995
```