Ψηφιακή Επεξεργασία Σήματος

2η Εργαστηριακή άσκηση

Σιφναίος Σάββας, ΑΜ: 03116080

Στούμπου Βασιλική, ΑΜ: 03116026

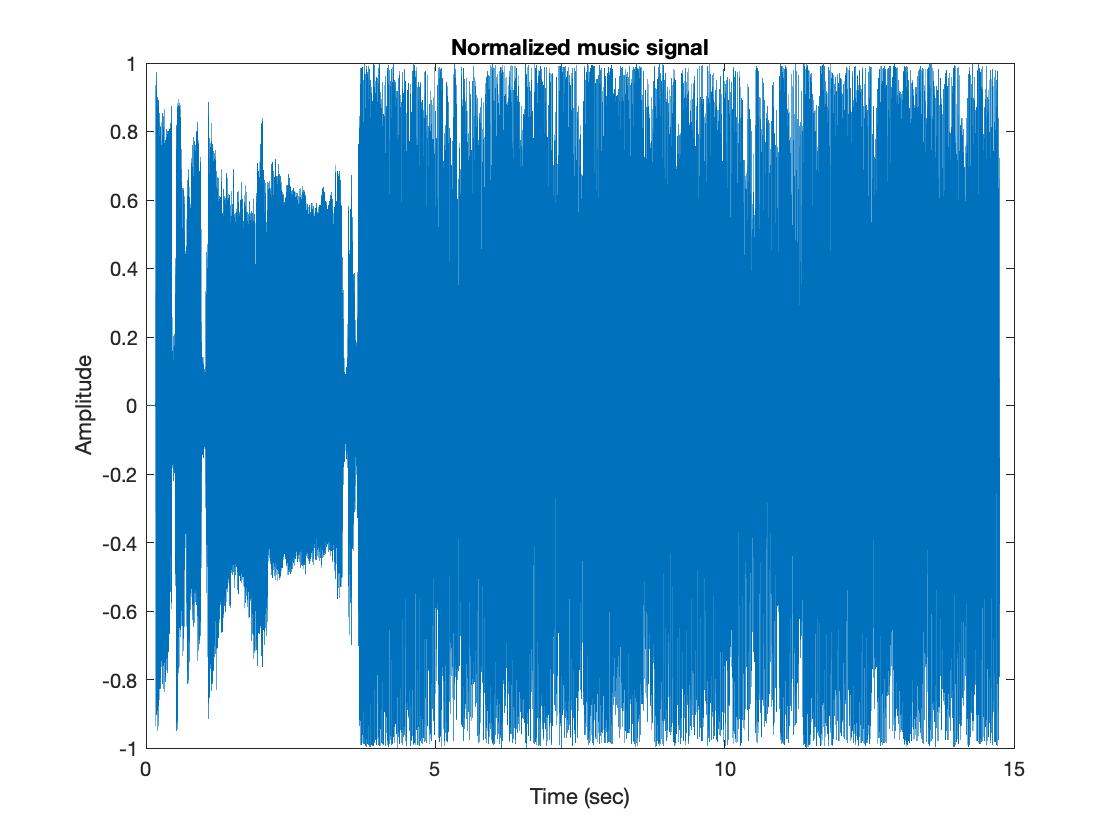
6ο εξάμηνο, ΗΜΜΥ

1ο Μέρος

Στόχος του μέρους αυτού είναι η υλοποίηση του ψυχοακουστικού μοντέλου 1.

**Βήμα 1.0**

Διαβάζουμε το σήμα μουσικής χρησιμοποιώντας τη συνάρτηση audioread. Στη συνέχεια, κανονικοποιούμε το σήμα μουσικής, ώστε όλες οι τιμές του να βρίσκονται στο διάστημα [-1,1]. Παρακάτω, φαίνεται η γραφική παράσταση του σήματος:



Στη συνέχεια, χωρίζουμε το σήμα μας σε χρονικά πλαίσια των 512 δειγμάτων. Γεμίζουμε το τελευταίο πλαίσιο, που έχει λιγότερα δείγματα, με μηδενικά. Προκύπτουν συνολικά 1271 πλαίσια. Η ανάλυση στα επόμενα βήματα γίνεται ανά πλαίσιο, γι’ αυτό και χρησιμοποιούμε επαναληπτικό βρόχο.

**Βήμα 1.1**

Ορίζουμε το κατώφλι ακοής, καθώς επίσης και την κλίμακα Bark με βάση τις σχέσεις που δίνονται στην εκφώνηση. Χρησιμοποιώντας παράθυρο Ηanning, παραθυροποιούμε το σήμα και υπολογίζουμε το φάσμα ισχύος σε dB:

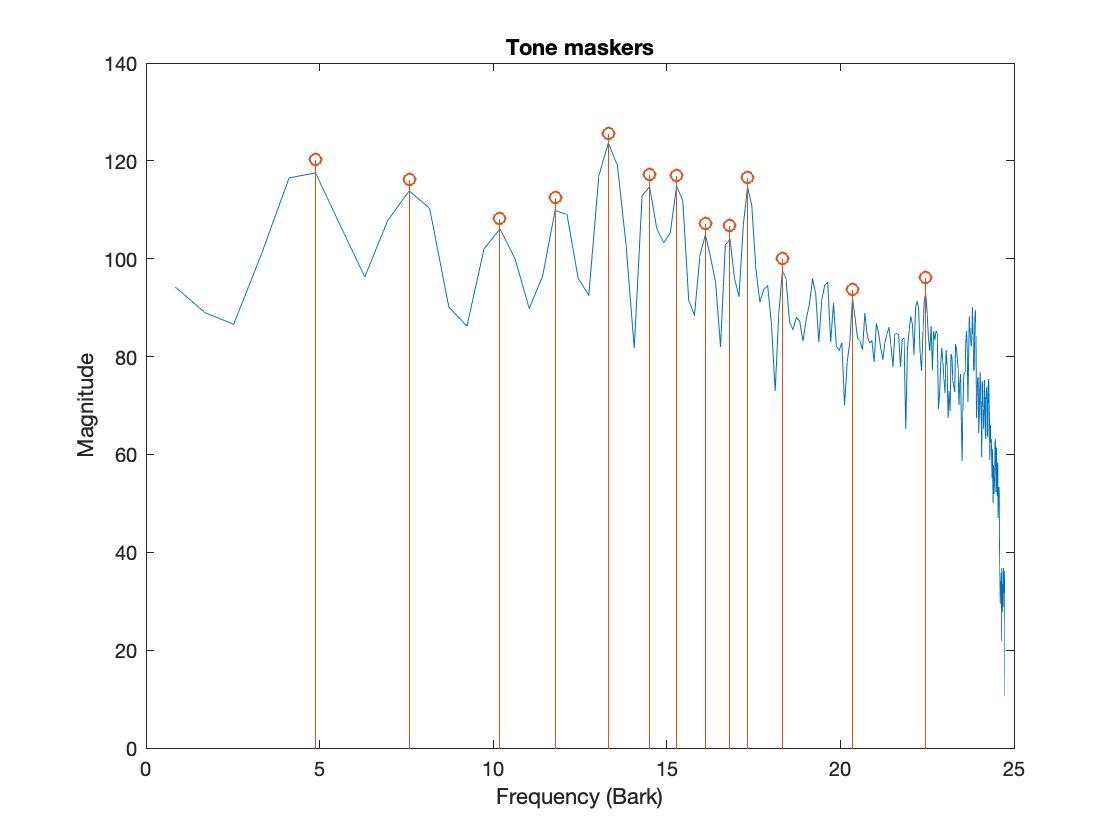
όπου και .

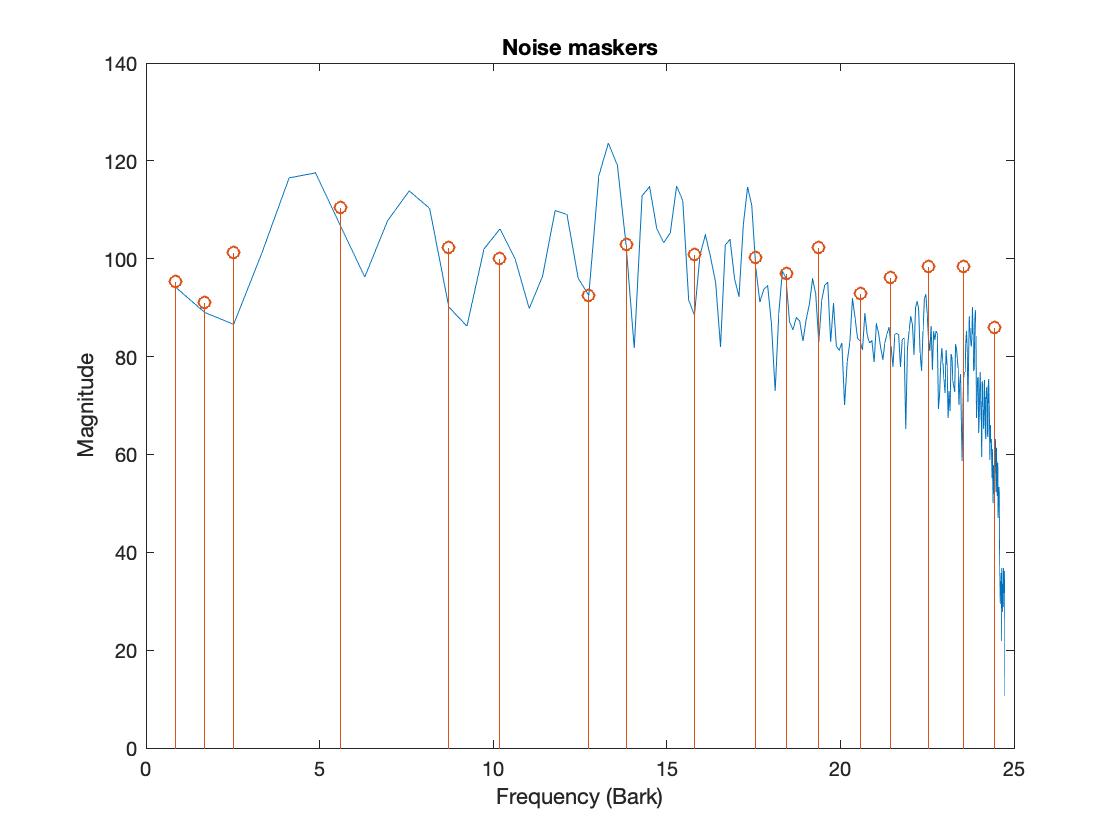
Λόγω συμμετρίας του φάσματος, κρατάμε μόνο τα πρώτα 256 δείγματα, δηλαδή .

**Βήμα 1.2**

Στόχος αυτού του βήματος είναι ο εντοπισμός των μασκών τόνων και θορύβου. Για την εύρεση των τονικών μασκών, χρησιμοποιούμε τη συνάρτηση , η οποία λαμβάνει την τιμή 1 μόνο στις θέσεις όπου τα τοπικά μέγιστα του φάσματος ισχύος είναι μεγαλύτερα κατά τουλάχιστον 7 dB από τις γειτονικές συχνότητες. Σε αυτές τις θέσεις, έχουμε μία μάσκα που επικαλύπτει τις γειτονικές συχνότητες. Το εύρος της γειτονιάς που εξετάζουμε εξαρτάται από την περιοχή συχνοτήτων που βρισκόμαστε. Σε μεγαλύτερες συχνότητες, οι μάσκες καλύπτουν ευρύτερες γειτονιές. Στη συνέχεια, με βάση την εξής σχέση (σε dB):

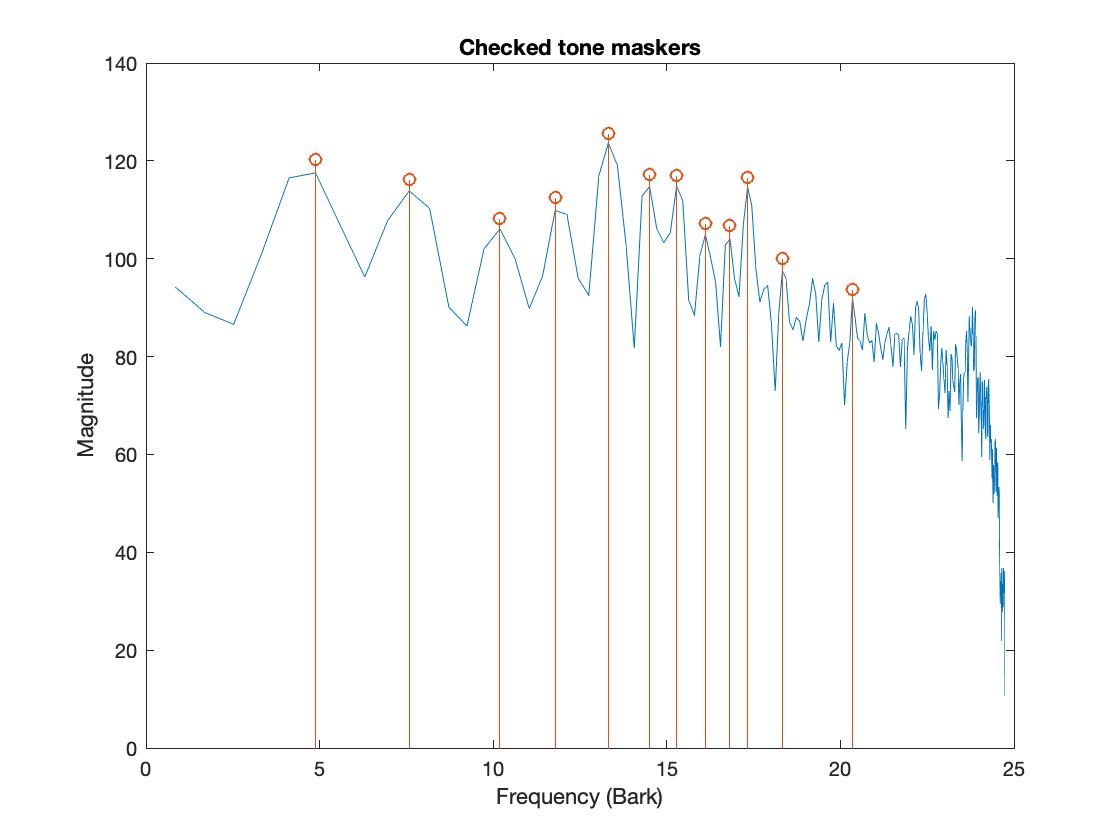
υπολογίζουμε την ισχύ των τονικών μασκών. Για την εύρεση των μασκών θορύβου και της ισχύος αυτών, χρησιμοποιούμε την έτοιμη συνάρτηση που μας δίνεται. Παρουσιάζονται σε κοινό διάγραμμα το φάσμα ισχύος και η ισχύς των τονικών μασκών, καθώς επίσης και το φάσμα ισχύος με την ισχύ των μασκών θορύβου για το 100ο πλαίσιο ανάλυσης.

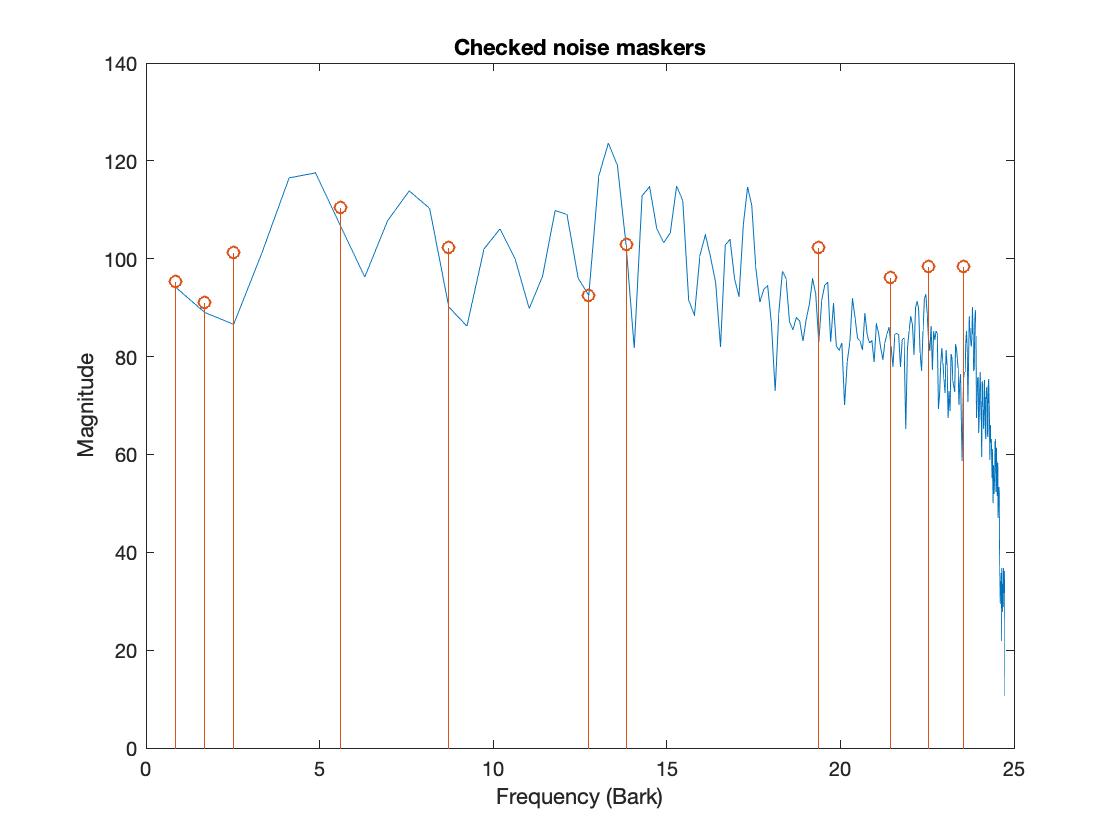




**Βήμα 1.3**

Στη συνέχεια, περιορίζουμε τον αριθμό των μασκών, χρησιμοποιώντας την έτοιμη συνάρτηση που μας δίνεται. Και εδώ παρουσιάζονται σε κοινό διάγραμμα το φάσμα ισχύος και η ισχύς των περιορισμένων τονικών μασκών, καθώς επίσης και το φάσμα ισχύος με την ισχύ των περιορισμένων μασκών θορύβου για το 100ο πλαίσιο ανάλυσης.

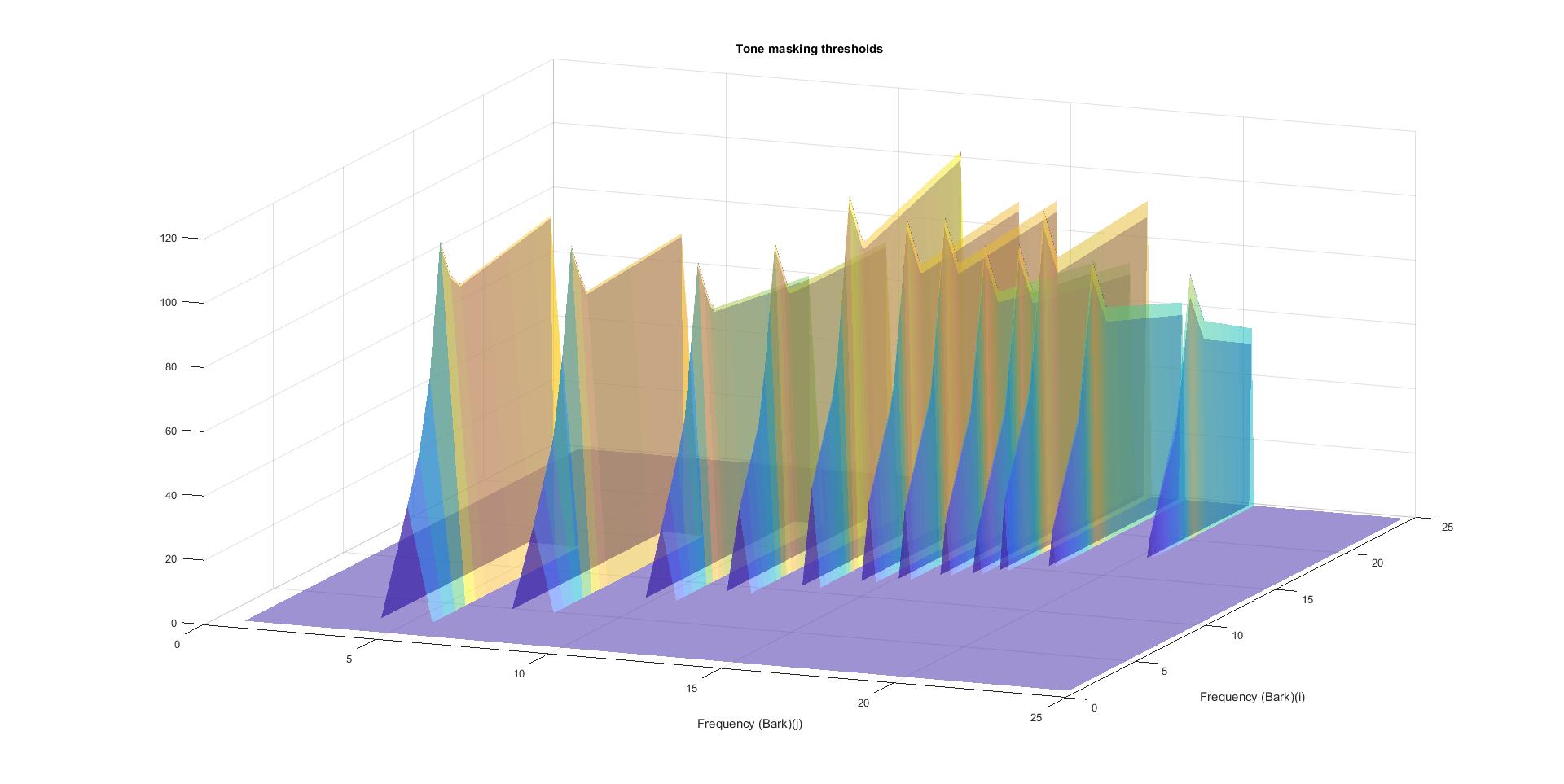


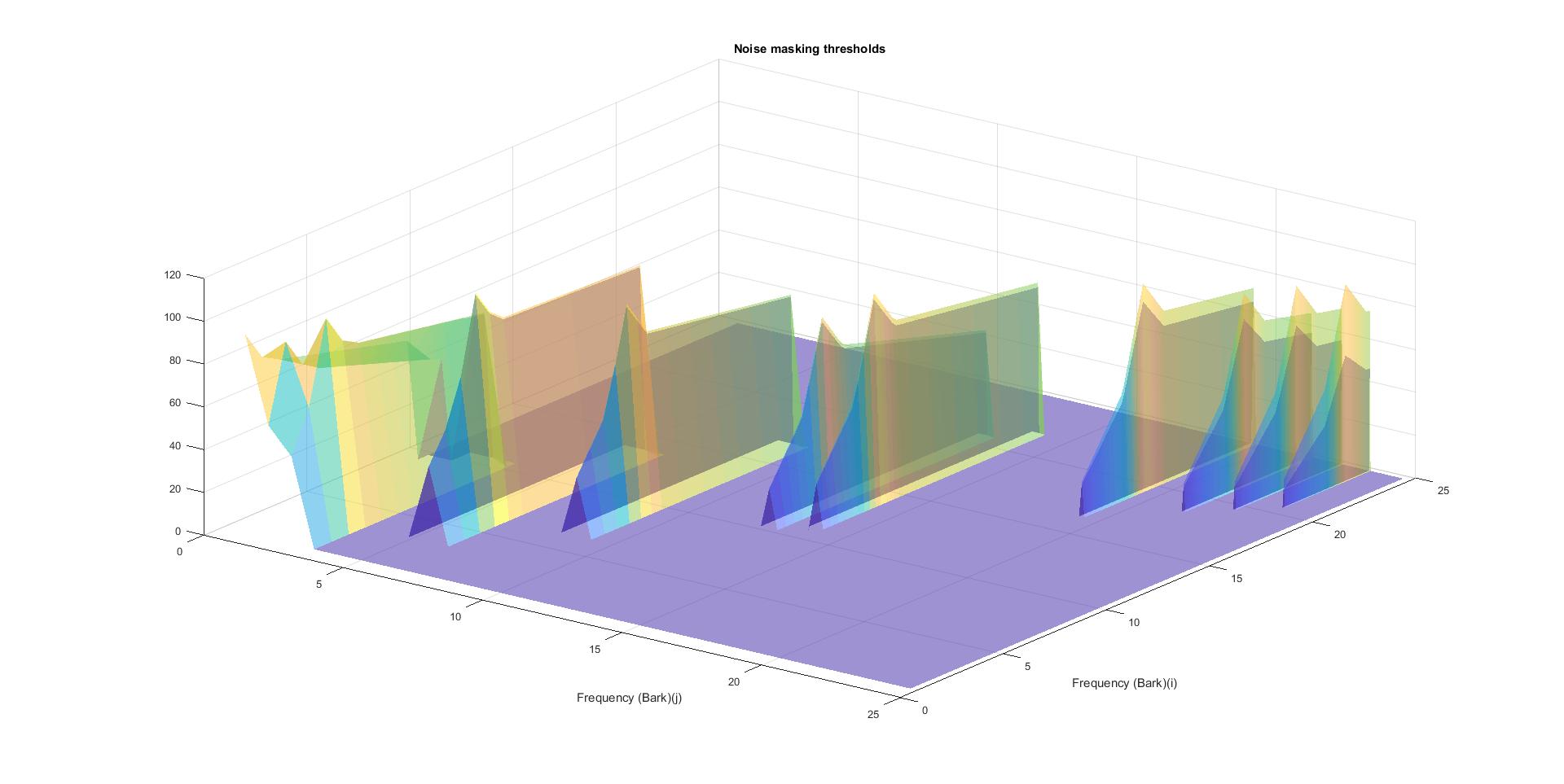


**Βήμα 1.4**

Σε αυτό το βήμα υπολογίζουμε τα κατώφλια κάλυψης για την περίπτωση των μασκών τόνων και θορύβου. Το κάθε κατώφλι αντιπροσωπεύει το ποσοστό κάλυψης σε ένα συγκεκριμένο σημείο i, λόγω της μάσκας που υπάρχει σε ένα σημείο j. Το εύρος των συχνοτήτων στο οποίο εξετάζουμε την επικάλυψη που προκαλεί μία μάσκα είναι της τάξης των 12 Bark (3 τιμές της κλίμακας Bark κάτω έως και 8 τιμές πάνω από την τιμή b(j)). Για τον υπολογισμό των κατωφλίων και σε κάποιο σημείο i λόγω της μάσκας στο σημείο j, χρησιμοποιούμε την ισχύ της μάσκας και αντίστοιχα, την τιμή b(j), καθώς επίσης και τη συνάρτηση . Η συνάρτηση αυτή εκφράζει την έκταση της κάλυψης και η τιμή της εξαρτάται από την ισχύ της μάσκας και τη διαφορά . Τα κατώφλια και δίνονται από τους τύπους (σε dB):

Παρακάτω, ακολουθεί η 3-διάστατη γραφική απεικόνιση των κατωφλίων και για το 100ο πλαίσιο ανάλυσης.

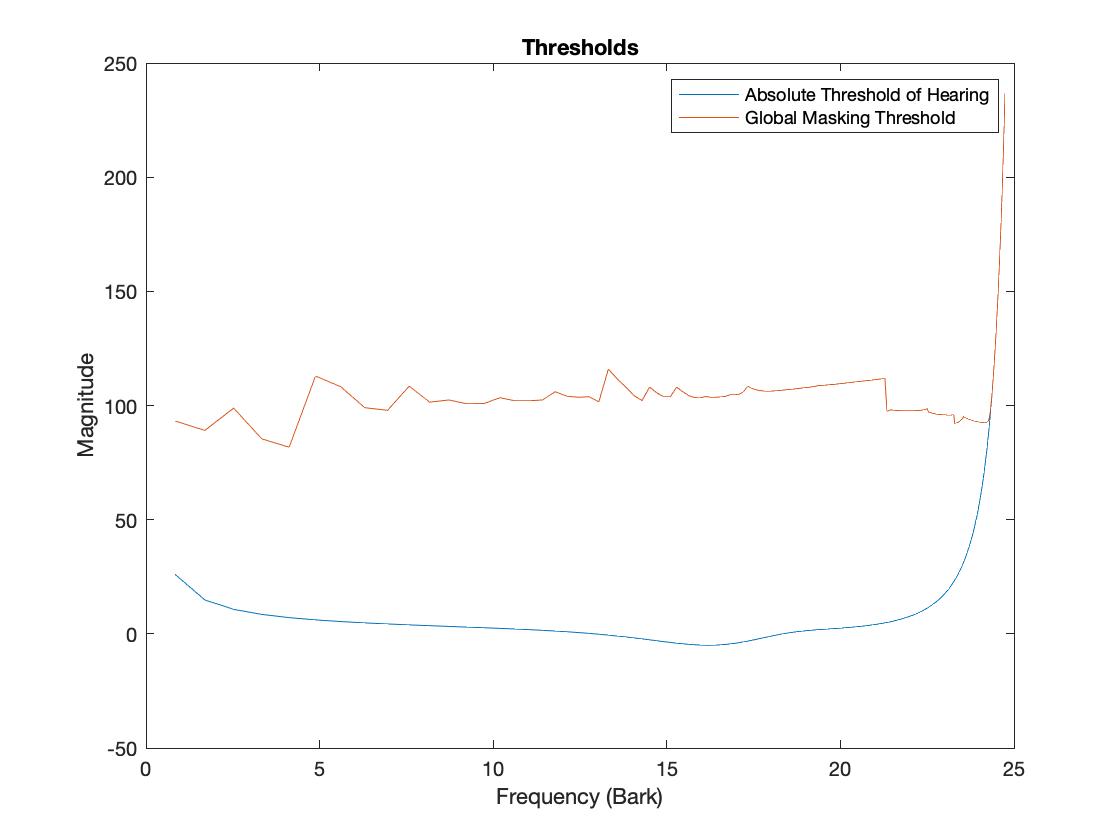
**



**Βήμα 1.5**

Τέλος, υπολογίζουμε το συνολικό κατώφλι κάλυψης σε κάθε διακριτή συχνότητα, με βάση την εξής σχέση:

Ακολουθεί, σε κοινό διάγραμμα, η γραφική παράσταση του συνολικού κατωφλίου κάλυψης και του κατωφλίου ακοής ως προς τη συχνότητα (σε Bark).



Παρατηρούμε ότι, λόγω των μασκών, η τιμή του συνολικού κατωφλίου κάλυψης είναι σημαντικά μεγαλύτερη σε σχέση με αυτή του κατωφλίου ακοής για τις περισσότερες τιμές συχνοτήτων.

2ο Μέρος

Σε αυτό το μέρος της άσκησης καλούμαστε να συμπιέσουμε το σήμα μας με τη βοήθεια του ψυχοακουστικού μοντέλου που υλοποιήσαμε στο 1ο μέρος. Συγκεκριμένα, αναλύουμε το αρχικό σήμα σε συγκεκριμένες συχνοτικές ζώνες, προκειμένου να μπορέσουμε να το κβαντίσουμε με τρόπο κατάλληλο ώστε το λάθος στο τελικό σήμα να μην είναι αντιληπτό . Αυτό επιτυγχάνεται με την χρήση συστοιχίας 32 ζωνοπερατών φίλτρων, των οποίων οι ζώνες διέλευσης καλύπτουν ολόκληρο το φάσμα συχνοτήτων του σήματος. Το αποτέλεσμα της παραπάνω ανάλυσης είναι η διαίρεση/διάσπαση του προς επεξεργασία σήματος σε συχνοτικές υποζώνες διευκολύνοντας έτσι τον διαχωρισμό των αντιληπτικά κρίσιμων σημείων του σήματος από τα περιττά.

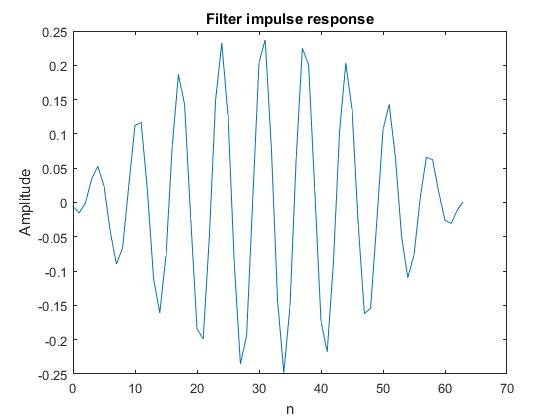
Παρακάτω παρουσιάζονται τα βήματα μέσω τον οποίων παράγεται το τελικό σήμα.

**Βήμα 2.0**

Το βήμα αυτό θα μπορούσε να θεωρηθεί εισαγωγικό, καθώς εδώ ορίζουμε τις κρουστικές αποκρίσεις των συστοιχιών ζωνοπερατών φίλτρων. Ειδικότερα, για την ανάλυση χρησιμοποιείται ο Modified Discrete Cosine Transform, ο οποίος είναι πλήρως αντιστρέψιμος και δεν εισάγει λάθη στην κωδικοποίηση του σήματος και ορίζεται από την σχέση :

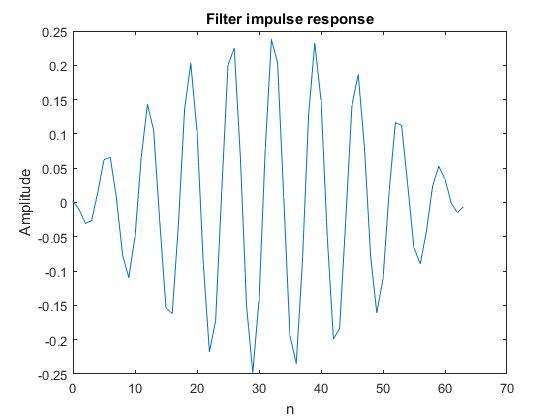
, όπου Μ = 32 και k ανεξάρτητη ακέραια μεταβλητή που λειτουργεί ως δείκτης σε ένα από τα 32 ζωνοπερατά φίλτρα της συστοιχίας (k = 1, 2, ..., M).

Παρακάτω παρατίθεται το γράφημα της κρουστικής απόκρισης του 10ου φίλτρου ανάλυσης ( ).



Επιπλέον, η κρουστική απόκριση των φίλτρων σύνθεσης του τελικού σήματος ορίζεται ως , η οποία είναι μια κλιμακωμένη και χρονικά μετατοπισμένη εκδοχή της .

Παρακάτω παρατίθεται το γράφημα της κρουστικής απόκρισης του 10ου φίλτρου σύνθεσης ().



**Βήμα 2.1**

Στο σημείο αυτό ξεκινάμε την ανάλυση του σήματος μουσικής. Αρχικά, κάθε παράθυρο του σήματος φιλτράρεται από τα 32 ζωνοπερατά φίλτρα ανάλυσης της συστοιχίας. Η παραπάνω διαδικασία αντιστοιχεί στον υπολογισμό της συνέλιξης του κάθε παραθύρου με τις κρουστικές αποκρίσεις των 32 αυτών φίλτρων. Στο σημείο αυτό, αξίζει να σημειωθεί πως το αποτέλεσμα του φιλτραρίσματος των παραθύρων (μήκους 512 δειγμάτων), δηλαδή της συνέλιξής τους με την κρουστική απόκριση των φίλτρων (μήκους 64 δειγμάτων), είναι σήματα εξόδου μήκους 512+64-1 = 575 δειγμάτων.

Στην συνέχεια, πραγματοποιούμε τον αποδεκατισμό των φιλτραρισμένων παραθύρων κατά παράγοντα υποδειγματοληψίας 32 , ο οποίος είναι και ο μέγιστος δυνατός ώστε να μην υπάρχει επικάλυψη του σήματος (aliasing). Η κύρια ιδέα του αποδεκατισμού είναι ότι από όλα τα δείγματα του αρχικού σήματος εμείς επιλέγουμε να κρατήσουμε εκείνα τα οποία είναι ακέραιο πολλαπλάσιο του 32, δηλαδή εκείνα που ικανοποιούν την σχέση sample\_number mod 32 == 0. Τονίζουμε πως στον κώδικα Matlab η παραπάνω διαδικασία εκτελείται με τη χρήση της συνάρτησης "downsample". Τέλος, σημειώνεται ότι το αποτέλεσμα του αποδεκατισμού του σήματος κατά παράγοντα 32 , δίνει ένα νέο σήμα μήκους 575/32 = 18 δειγμάτων.

**Βήμα 2.2**

Σε αυτό το βήμα καλούμαστε να κβαντίσουμε τα σήματα που προκύπτουν από το προηγούμενο στάδιο, δηλαδή την έξοδο των φίλτρων ανά πλαίσιο ανάλυσης. Σκοπός της κβάντισης είναι η αντιστοίχιση των τιμών του αρχικού σήματος σε συγκεκριμένες, προκαθορισμένες, ισαπέχουσες κβαντισμένες τιμές, οι οποίες υπολογίζονται βάσει της χαμηλότερης στάθμης του κβαντιστή και ενός βήματος κβάντισης. Στη συνέχεια, υλοποιούμε δύο τύπους γραμμικών κβαντιστών έναν **προσαρμοζόμενο** και έναν **μη προσαρμοζόμενο**. Η αρχή λειτουργίας μη προσαρμοζόμενου κβαντιστή είναι η εξής:

* Επιλογή της τάξης του κβαντιστή, του αριθμού δηλαδή των bits που χρησιμοποιούνται για την αναπαράσταση των επιπέδων κβάντισης.
* Ορισμός των 2^τάξη επιπέδων.
* Ορισμός του χαμηλότερου επιπέδου κβάντισης ως το -1.
* Ορισμός του βήματος κβάντισης και υπολογισμός του ως Βήμα = (1-(-1))/2^τάξη.
* Υπολογισμός των επόμενων διακριτών επιπέδων κβάντισης μέσω της σχέσης **επόμενο επίπεδο = προηγούμενο επίπεδο + Βήμα**.
* Αντιστοίχιση κάθε τιμής του σήματος προς κβάντιση στο πλησιέστερο επίπεδο του κβαντιστή.

Ο προσαρμοζόμενος κβαντιστής διαφοροποιείται από τον παραπάνω λόγω της δυνατότητας του να επιλέγει την τάξη του ανάλογα με το πλαίσιο το οποίο επεξεργάζεται κάθε φορά. Επιπλέον, εκτός της δυνατότητας προσαρμογής της τάξης του, ανάλογα με τις απαιτήσεις του σήματος, στον συγκεκριμένο κβαντιστή ως χαμηλότερο επίπεδο ορίζεται η ελάχιστη τιμή του προς κβάντιση σήματος. Με τις δύο αυτές τροποποιήσεις επιτυγχάνεται η μείωση των απωλειών πληροφορίας.

Παρακάτω γίνεται επεξήγηση της διαδικασίας "προσαρμογής" της τάξης του κβαντιστή για κάθε πλαίσιο ανάλυσης.

Αρχικά υπολογίζονται οι στάθμες έντασης του σήματος R (δηλαδή το πλήθος των διαφορετικών του τιμών).

Για R=2^16 υπολογίζουμε την τάξη του κβαντιστή, βάσει του παρακάτω τύπου:

όπου είναι η ελάχιστη τιμή του συνολικού κατωφλίου κάλυψης για το οποίο οι δείκτες i ικανοποιούν την εξής σχέση:

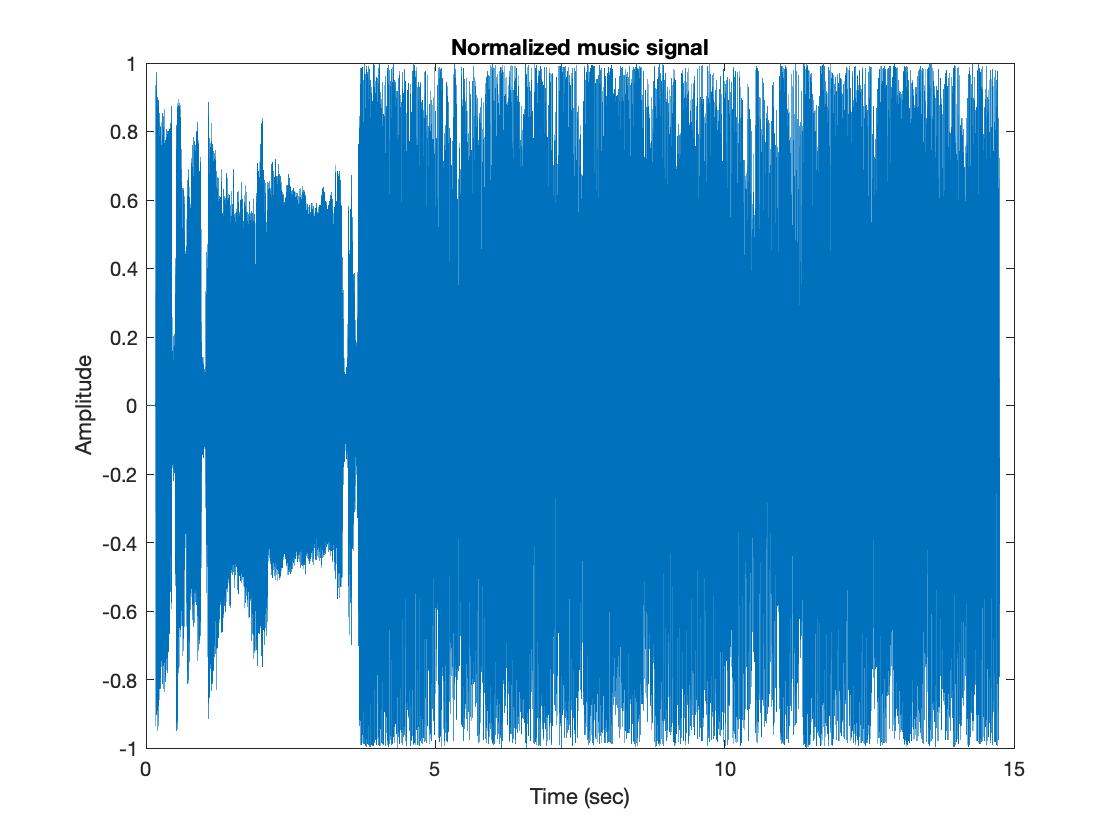
όπου οι κεντρικές συχνότητες τω φίλτρων ανάλυσης-σύνθεσης και υπολογίζονται ως εξής:

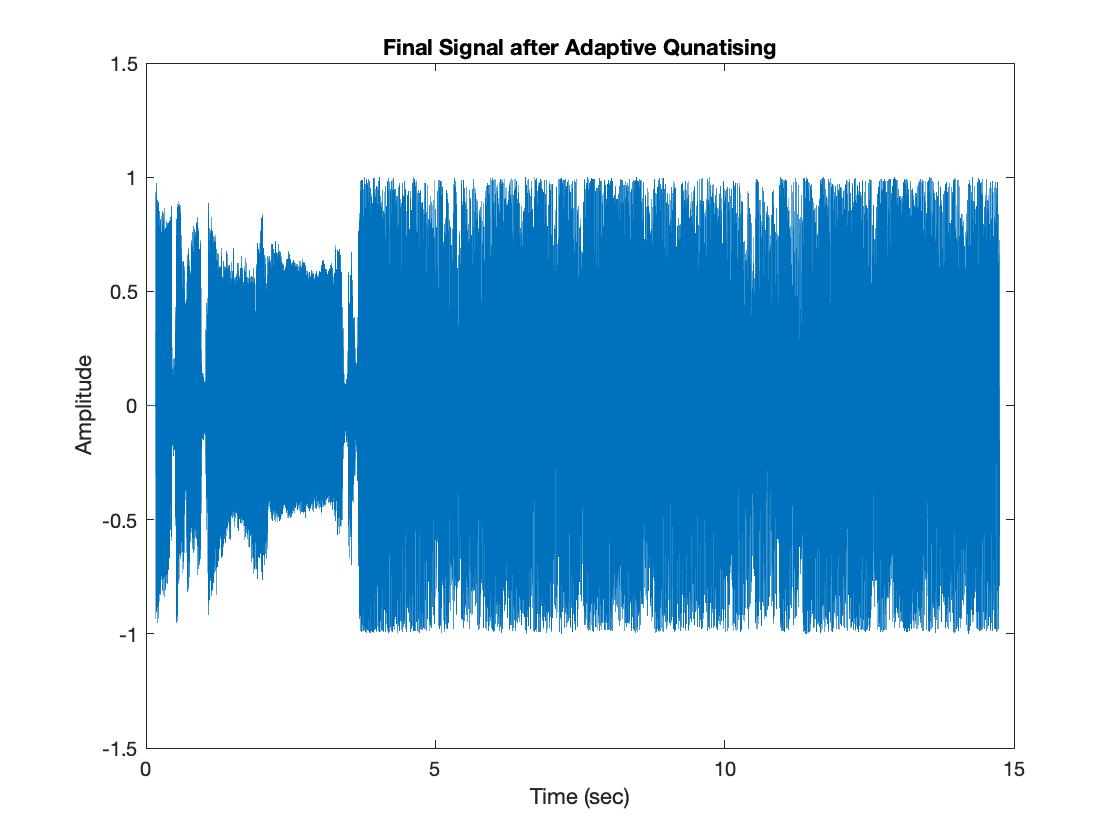
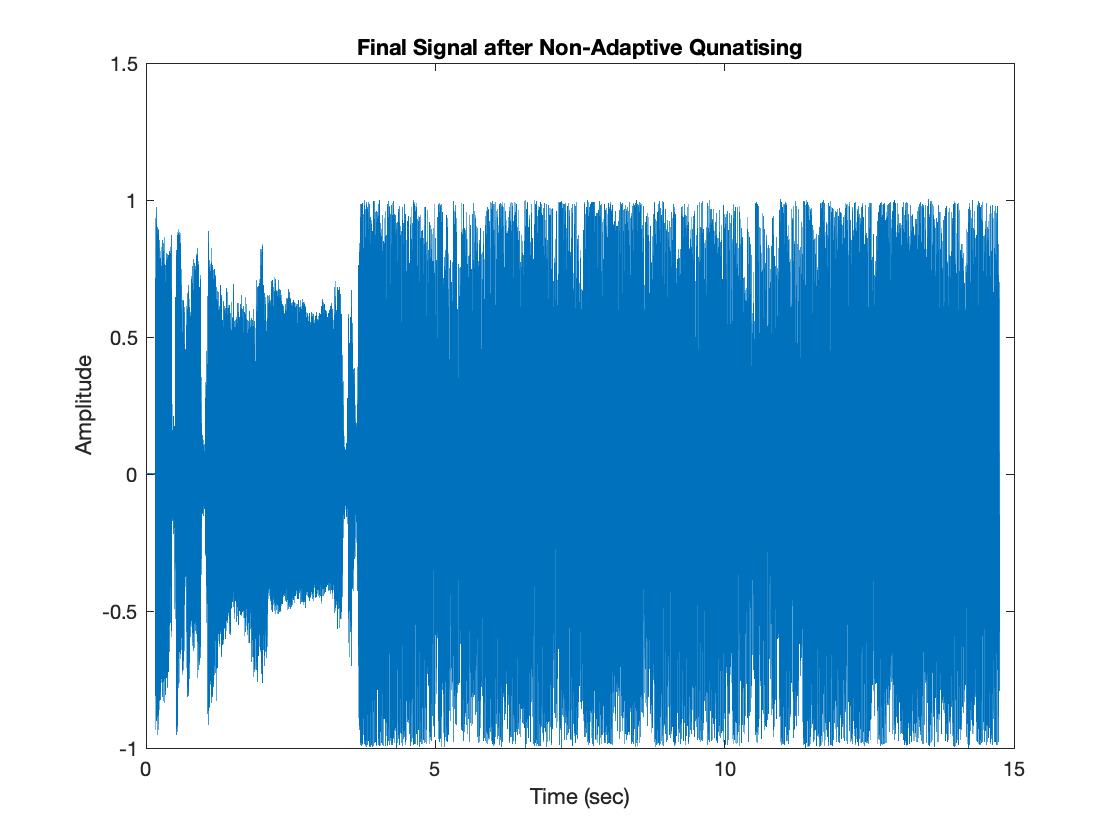
**Βήμα 2.3**

Στο τελευταίο αυτό βήμα χρησιμοποιούμε τα κβαντισμένα σήματα που έχουν παραχθεί στο παραπάνω ερώτημα. Στα κβαντισμένα αυτά σήματα κάνουμε υπερδειγματοληψία κατά παράγοντα 32, με αποτέλεσμα τα νέα σήματα, ανά πλαίσιο ανάλυσης, να αποτελούνται από τις κβαντισμένες τιμές του σήματος στις θέσεις που είναι ακέραια πολλαπλάσια του Μ (Μ=32), ενώ σε όλες τις υπόλοιπες θέσεις να έχουν μηδενική τιμή. Η παρεμβολή αυτή μηδενικών μεταξύ των 18 διαδοχικών τιμών ονομάζεται Zero-Padding. Επιπλέον, αξίζει να σημειωθεί πως στον κώδικα Matlab η διαδικασία της υπερδειγματοληψίας γίνεται με χρήση της συνάρτησης "upsample" . Στην συνέχεια τα σήματα που παράγονται από την υπερδειγματοληψία φιλτράρονται από τα 32 φίλτρα σύνθεσης με τις εξόδους κάθε φίλτρου να αθροίζονται ανά πλαίσιο ανάλυσης, ώστε τελικά να δημιουργηθεί το συνολικό κβαντισμένο σήμα εξόδου. Η παραπάνω διαδικασία, αντιστοιχεί σε συνέλιξη των σημάτων που προκύπτουν από την υπερδειγματοληψία με τις κρουστικές αποκρίσεις των φίλτρων σύνθεσης, . Για το κάθε παράθυρο, το τελικό σήμα προκύπτει ως άθροισμα των παραπάνω εξόδων για όλες τις τιμές του κ, δηλαδή για όλα τα τμήματα της συστοιχίας. Συνεπώς, γίνεται κατανοητό πως τα σήματα εξόδου, ως αποτελέσματα συνέλιξης, θα έχουν μήκος 576+64-1 = 639 δείγματα. Τέλος, αξίζει να σημειωθεί πως αν αθροίζονταν, απλά, οι έξοδοι των φίλτρων σύνθεσης το τελικό σήμα θα είχε μήκος 1271\*639 = 812.169 δείγματα. Ωστόσο, αύτη η παράθεση των εξόδων των φίλτρων ανά πλαίσιο έχει σαν αποτέλεσμα την εισαγωγή μεγάλου πλήθους μηδενικών στο σήμα εξόδου, γεγονός το οποίο ακουστικά μεταφράζεται σε έντονη παραμόρφωση και τελικά σε μεγάλη απόκλιση του τελικού σήματος από το δοθέν. Για τον λόγο αυτό, οι έξοδοι των φίλτρων σύνθεσης διατάσσονται με τέτοιο τρόπο ώστε να υπάρχει επικάλυψη των τελευταίων δειγμάτων ενός πλαισίου από τα, αντίστοιχα σε πλήθος, πρώτα δείγματα του αμέσως επόμενου πλαισίου ανάλυσης. Η μέθοδος αύτη σύνθεσης του τελικού σήματος ονομάζεται Overlap-Add και έχει ως αποτέλεσμα τα τελικό σήμα εξόδου να έχει μήκος ίσο με αυτό του σήματος εισόδου.

Σύγκριση Μεθόδων και Αποτελεσμάτων

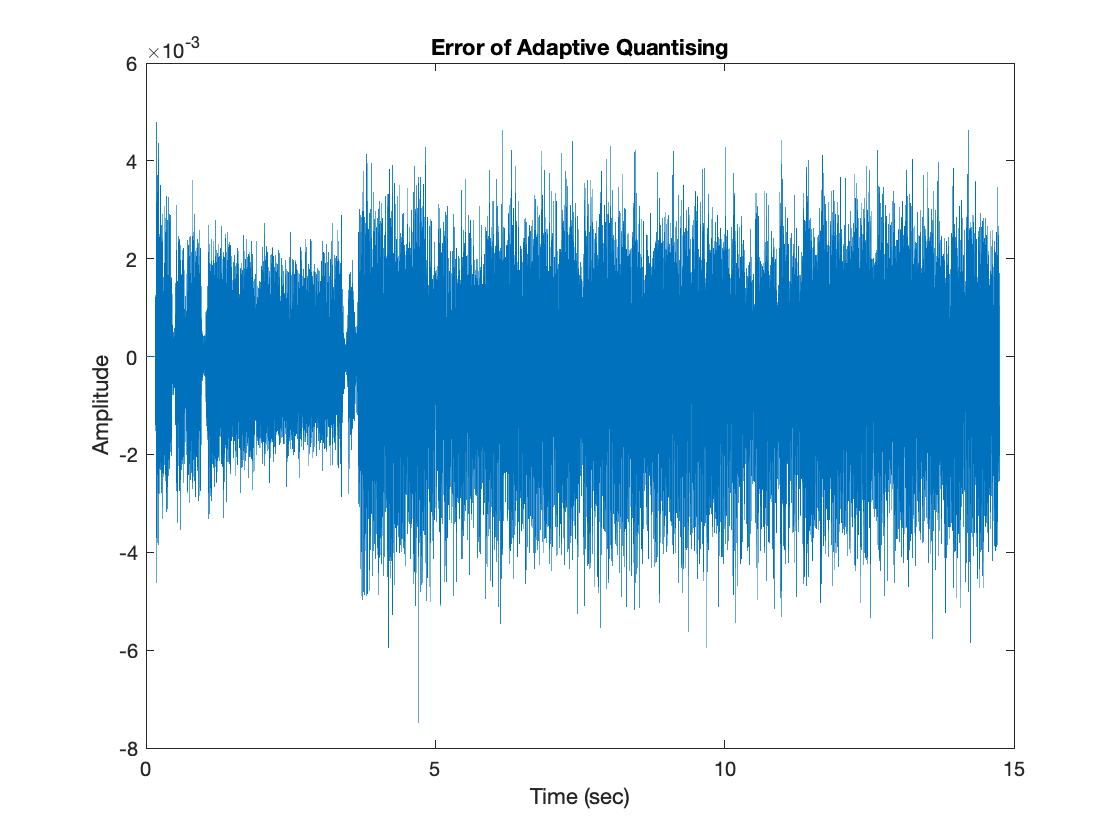
Παρακάτω παρατίθενται τα γραφήματα του αρχικού σήματος μετά την κανονικοποίησή του και των σημάτων που παράγονται με την χρήση προσαρμοζόμενου και μη προσαρμοζόμενου κβαντιστή.

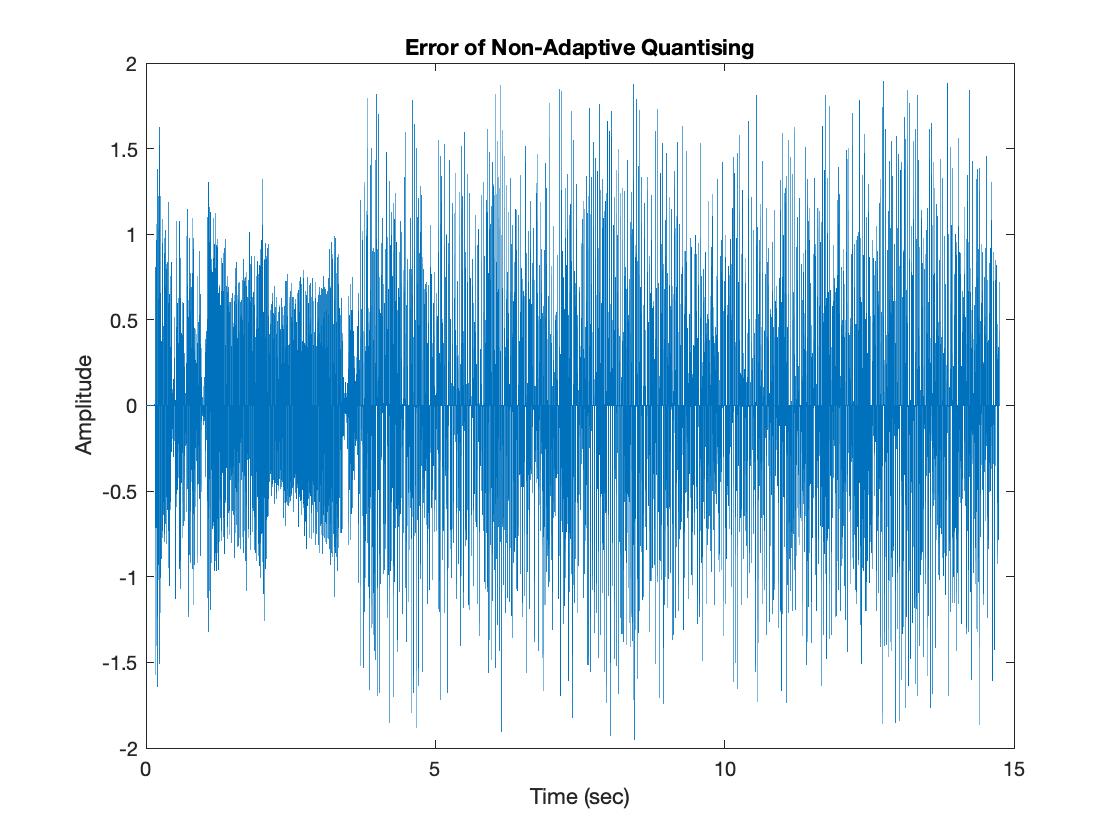




Στο σημείο αυτό, αξίζει να σημειωθεί πως το προκύπτον, από τον προσαρμοζόμενο κβαντιστή, σήμα είναι, ακουστικά, σχεδόν πανομοιότυπο με το δοθέν σήμα, σε αντίθεση με το προκύπτον από τον μη προσαρμοζόμενο κβαντιστή σήμα το οποίο έχει αισθητές διαφορές (θόρυβο) από το αρχικό.

Επιπλέον, παρατίθενται τα γραφήματα των σφαλμάτων που προκύπτουν από τις δύο μεθόδους κβάντισης. Παρατηρούμε από τη διαφορετική κλίμακα στην οποία παρουσιάζονται, πως το σφάλμα του προσαρμοζόμενου κβαντιστή είναι περίπου 3 τάξεις μικρότερο σε σχέση με το σφάλμα του μη προσαρμοζόμενου κβαντιστή.

.



Τέλος, μέσω του κώδικα Matlab, υπολογίσαμε το μέσο τετραγωνικό σφάλμα και το ποσοστό συμπίεσης για τις δύο μεθόδους κβάντισης και τα αποτελέσματα που προκύπτουν είναι τα εξής

**Για προσαρμοζόμενο κβαντιστή**

* Μέσο τετραγωνικό σφάλμα =
* Ποσοστό συμπίεσης = 55.4688%

**Για μη προσαρμοζόμενο κβαντιστή**

* Μέσο τετραγωνικό σφάλμα = 0.0416
* Ποσοστό συμπίεσης = 50%

Παρατηρούμε πως ο προσαρμοζόμενος κβαντιστής έχει μεγαλύτερο ποσοστό συμπίεσης από τον μη προσαρμοζόμενο και κατά συνέπεια ότι το μέσο τετραγωνικό σφάλμα του πρώτου είναι σημαντικά μικρότερο από το αντίστοιχο του δεύτερου.