Ψηφιακή Επεξεργασία Σημάτων

Εργαστηριακή Αναφορά 2:

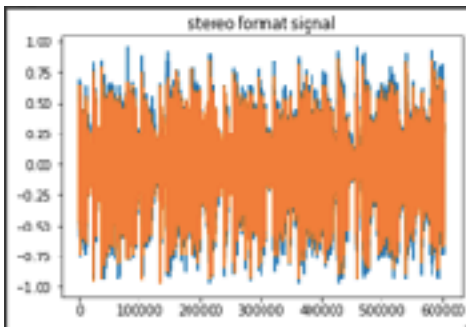
Θέμα: Κωδικοποίηση Σημάτων Μουσικής Βάσει του Ψυχοακουστικού μοντέλου.

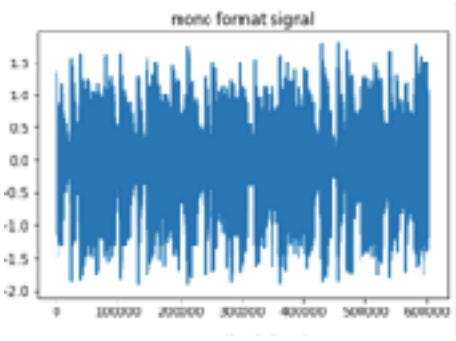
Φοιτητες: Δωροθέα Κουμίδου 03119712
Γιώργος Χαραλάμπους 03119706

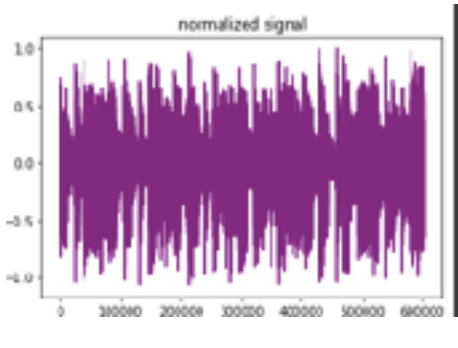
Μέρος 1

Ψυχοακουστικό Μοντέλο 1

Αρχικά εισάγουμε το σήμα μας και αποθηκεύουμε τα πλάτη του σήματος ως sgn ενώ την συχνότητα δειγματοληψίας fs. Στην συνέχεια μετατρέπουμε το σήμα μας από στέρεο σε μονό αθροίζοντας τα δύο chanels και το κανονικοποιούμε διαιρώντας το με το απόλυτο μέγιστο πλάτος του αθροισμένου σήματος.



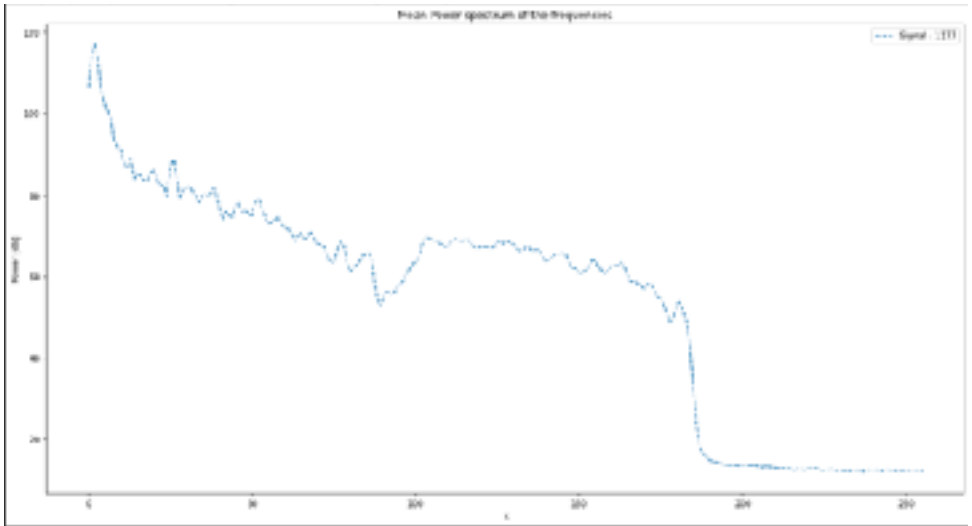




Μετά παραθυροποιούμε το σήμα μας σε πλαίσια των 512 δειγμάτων όμως με μια ματιά παρατηρήσαμε οτι το τελευταίο παράθυρο έχει λιγότερα δείγματα έτσι προσθέσαμε μηδενικά στο τέλος. Η εντολή concatenate είναι αυτή που μας βοήθησε να ενώσουμε τα τελευταία δείγματα με τα μηδενικά και να αναγνωρίζεται από το πρόγραμμα ως ένα παράθυρο των 512.

Έπειτα, δημιουργούμε μια συνάρτηση για την κλίμακα Bark και συμπληρώνουμε τον πίνακα Pk που αποτελεί το φάσμα ισχύος με τρόπο ώστε τελικά να κρατάμε το μονόπλευρο σήμα όπως φαίνεται παρακάτω.

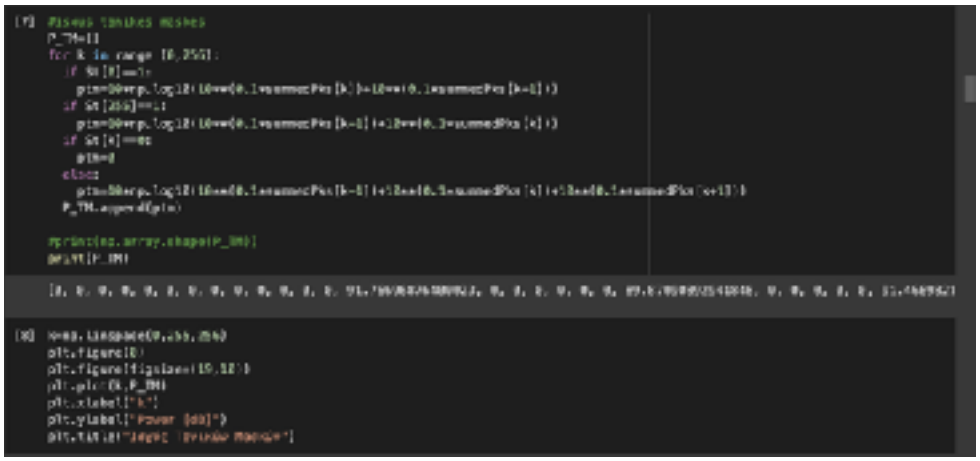
* Για να βοηθηθούμε αποθηκεύσαμε το μέσο όρο όλων των πλαισίων για κάθε συχνότητα και αυτό παραστήσαμε. Ουσιαστηκά το αρχικό power spectrum είχε διαστάσεις [256, 1179], τώρα είναι [256,1].

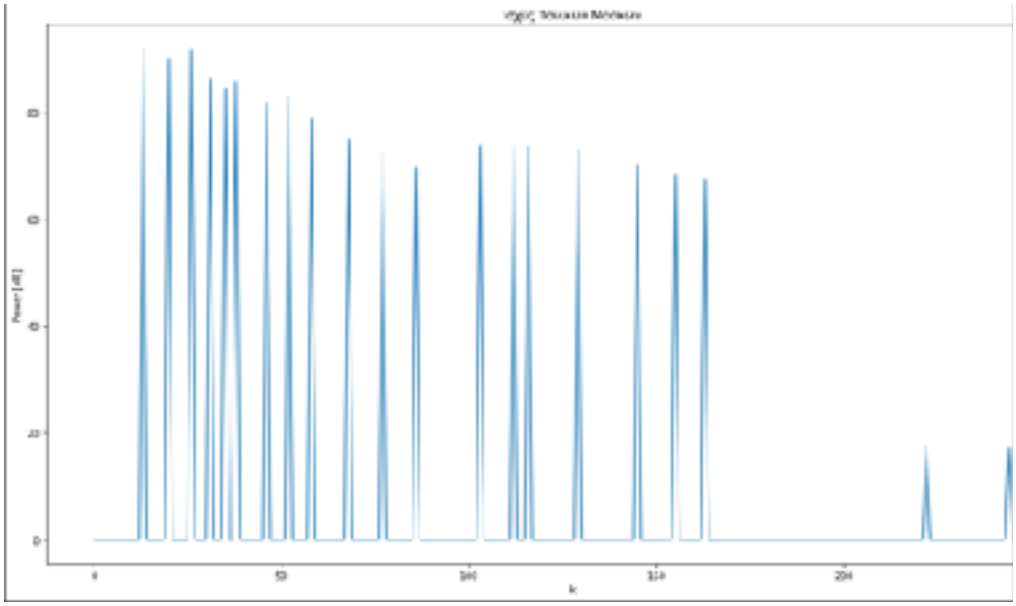


Στο βήμα 1.2 μας ζητήθηκε να βρούμε τις τονικές μάσκες οι οποίες είναι μεγαλύτερα από τις γειτονικές τους συχνότητες τουλάχιστον κατα 7 dB. Για να το καταφέρουμε αυτό, αρχικά φτιάξαμε ένα πίνακα με αποθηκευμένες τις τιμές του Δκ και ορίσαμε τον πίνακα St ως ένα πίνακα γεμάτο μηδενικά και κάθε φορά που βρίσκει τονική μάσκα το μηδέν να αντικαθιστάται από 1, έτσι μας δείχνει την θέση που υπάρχει μάσκα. Άρα περιμένουμε να έχουμε ένα πίνακα και πάλι [256,1]

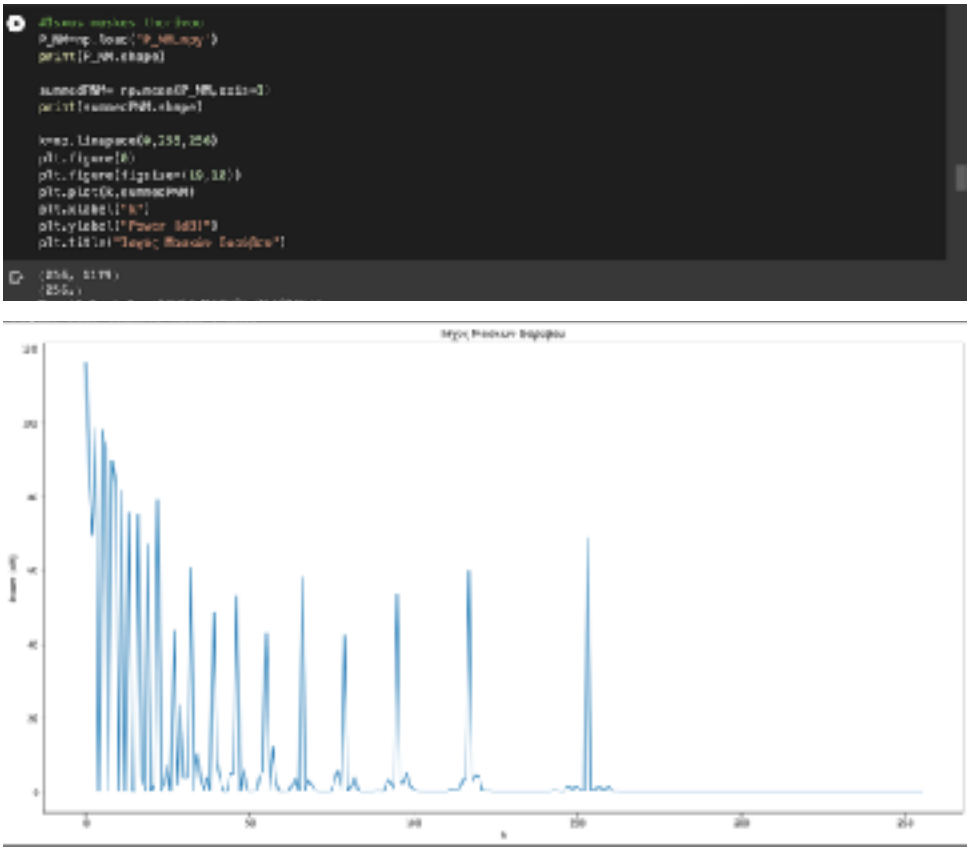
Αφού βρείκαμε τις θέσεις των τονικών μασκών, υπολογίσαμε την ισχύ τους με βάσει τον τύπο:

$$P_{TM}(k) = \begin{cases} 10\log_{10}(10^{0.1P(k-1)} + 10^{0.1P(k)} + 10^{0.1P(k+1)})\,(\text{dB}), & \text{αν } S_T(k) = 1 \\ 0, & \text{αν } S_T(k) = 0 \end{cases}$$

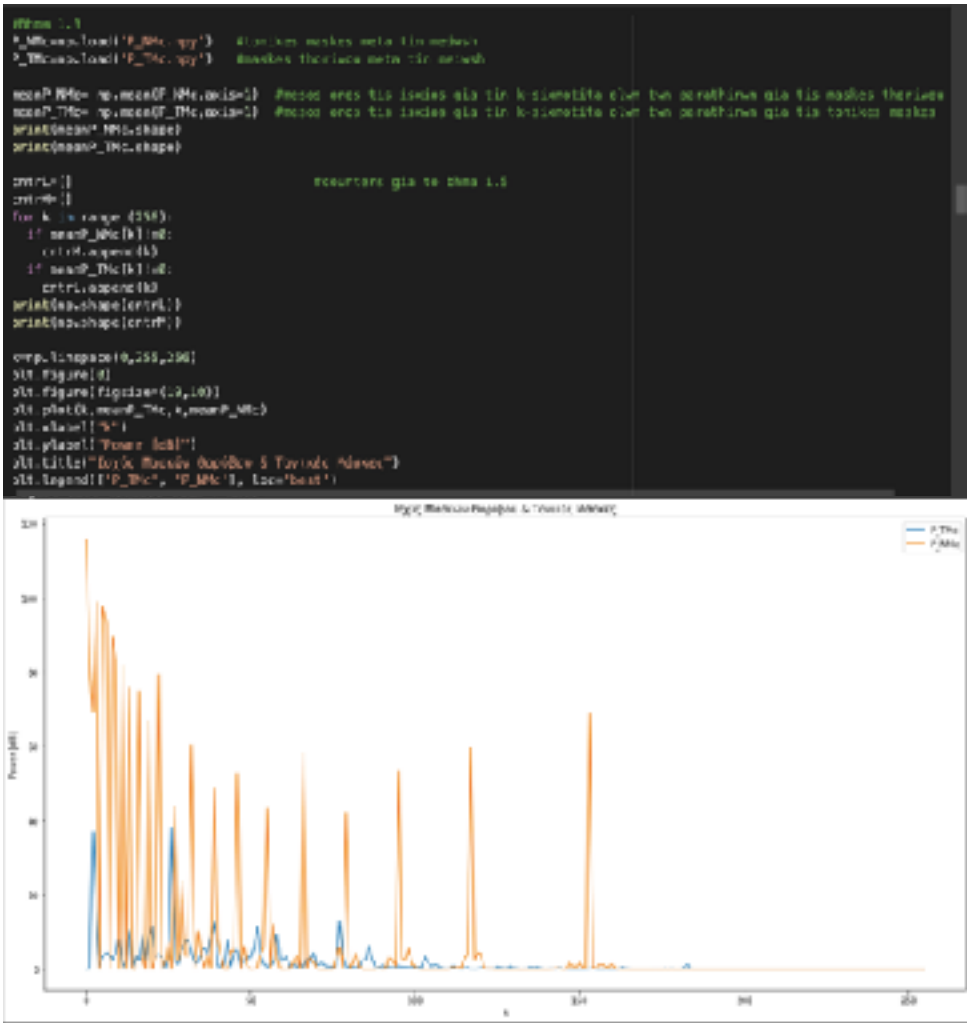




Πάρακάτω φαίνονται και οι μάσκες θορβού.



Στο βήμα 1.3 εισάγαμε τα δύο αρχεία και βρήκαμε το μέσο όρο του καθενός ξεχωριστά. Στον κώδικα φαίνεται επίσης ο counter για το βήμα 1.5

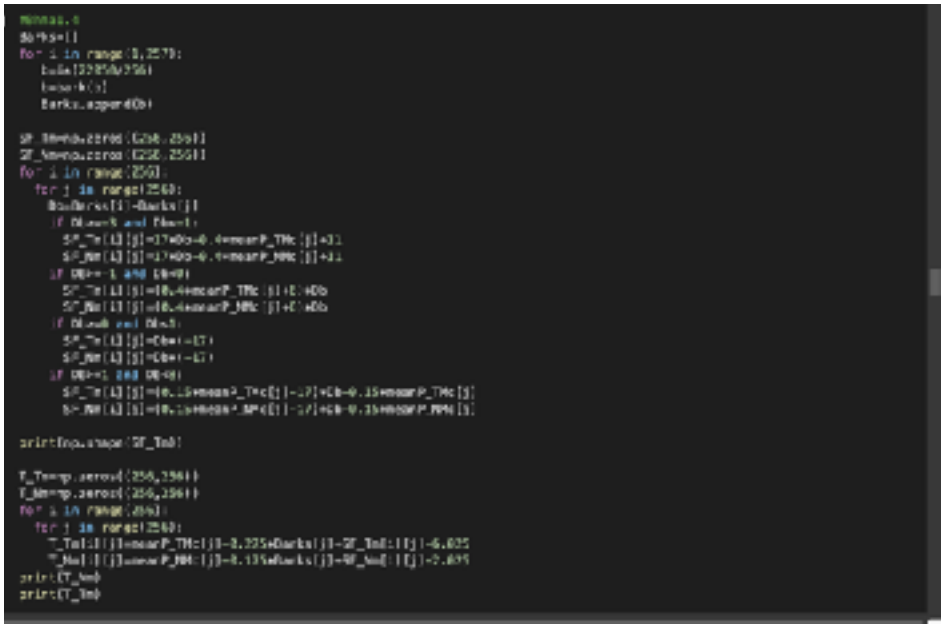


Στο επόμενο βήμα τα δύο διαφορετικά κατώφλια κάλυψης για την μάσκα τόνου και θορύβου με βάσει τους τύπους

$$T_{TM}(i,j) = P_{TM}(j) - 0.275b(j) + SF(i,j) - 6.025(\text{dB SPL})$$

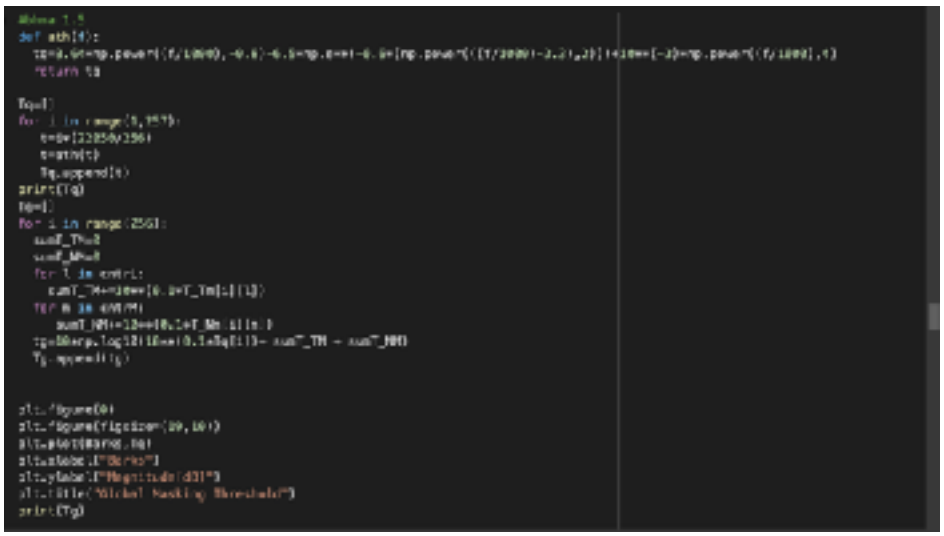
$$T_{NM}(i,j) = P_{NM}(j) - 0.175b(j) + SF(i,j) - 2.025(\text{dB SPL})$$

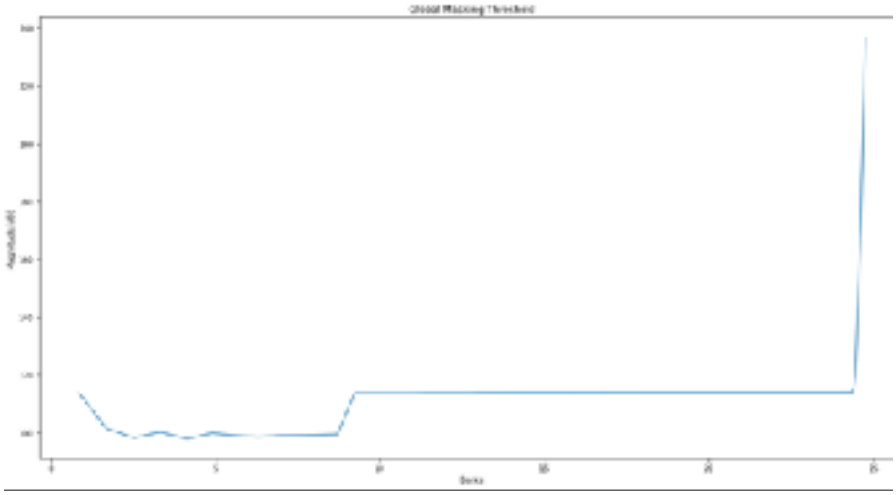
Αρχικά όπως φαίνεται πιο κάτω καλέσαμε την συνάρτηση που φτιάξαμε πιο πάνω για την κλίμακα Bark. Μετά ορίσμα πίνακες SF για τις συνάρτησεις SF των τονικών μασκών και των μασκών θορύβου ξεχωριστά, και δουλέψαμε με τον ίδιο τρόπο για την εύρεση των ST.



Με βάση τα προηγούμενα, στο βήμα 1.5 βρίσκουμε το συνολικό κατώφλι κάλυψης χρησιμοποιώντας τον τύπο:

$$T_g(i) = 10\log_{10}\left(10^{0.1T_q(i)} + \sum_{l=1}^{L} 10^{0.1T_{TM}(l,i)} + \sum_{m=1}^{M} 10^{0.1T_{NM}(m,i)}\right) \text{ dB SPL}$$





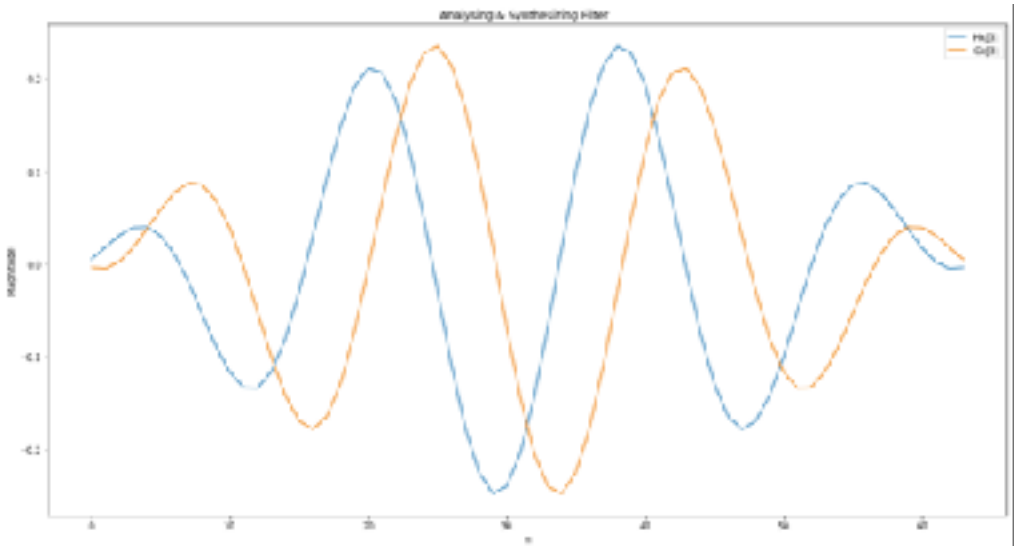
Μέρος 2:

Χρονο-Συχνοτική Ανάλυση με Συστοιχία Ζωνοπερατών Φίλτρων

Στο πρώτο βήμα του μέρους 2, για την ανάλυση του σήματος στις κρίσιμες συνιστώσες του, ανά χρονικό πλαίσιο, χρησιμοποιούμε συστοιχίες ζωνοπερατών φίλτρων. Δημιουργούμε με βάση τους πιο κάτω τύπους για τις κρουστικές αποκρούσεις:

$$h_k(n) = \sin\left[\left(n+\frac{1}{2}\right)\frac{\pi}{2M}\right]\sqrt{\frac{2}{M}}\cos\left[\frac{(2n+M+1)(2k+1)\pi}{4M}\right]$$

$$g_k(n) = h_k(2M-1-n)$$



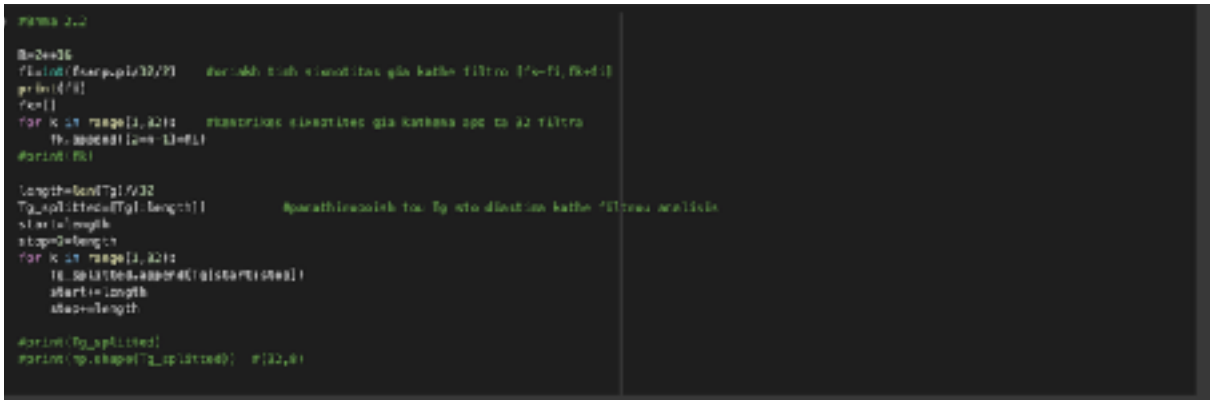
Για το βήμα 2.1, χρησιμοποιήσαμε την μεταβλητή filterNm και κάναμε την υποδειγματοληψία του σήματος κατά παράγοντα M=32.

Σε αυτό το βήμα, μας ζητείται να κβαντοποιήσουμε το σήμα με δύο τρόπους:

α) adaptive(Δ εξαρτημένο απο Bits κωδικοποίησης και το range του εκάστοτε παραθύρου)

β) non-adaptive(Δ =σταυερό και bits κωδικοποίησης=8)

Παρακάτω φαίνεται ο κώδικας για την παραθυροποίηση του Tg(absolute threshold) σε 32 συχνοτικές μπάντες. Αρχικά ορίσαμε την οριακή τιμή συχνότητας για καθένα από τα 32 φίλτρα και βρίκαμε τις κεντρικές συχνότητες(Fk) για κάθε φίλτρο. Στην συνέχεια φαίνεται η συνάρτηση κβαντοποίησης η οποία δέχεται ουσιαστικά 4 παραμέτρους (σήμα προς κβαντοποίηση, ελάχιστο και μέγιστο πλάτος και αριθμό των επιπέδων) και επιστρέφει το κβαντοποιημένο σήμα.

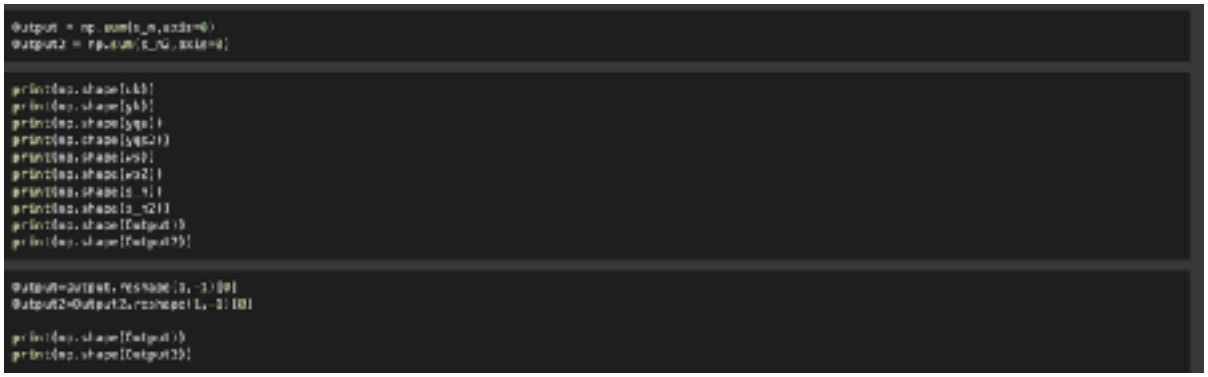


Για τον υπολογισμό του Bk σχηματίσαμε τον δοθέντα τύπο προσθέτοντας την εντολή np.ceil(), η οποία κρατάει το στρογγυλεμένο προς τα πάνω ακέραιο ψηφίο. Στη συνέχεια υπολογίζουμε το βήμα κβάντισης Δ μέσω του τύπου Δ=Xm/(2^B) και τελικά το μέγιστο πλάτος του κβαντιστή μας πολλαπλασιάζοντας το με 2^(B-1) επίπεδα μέσω της συνάρτησης calcl_xmax. Με παρόμοιο

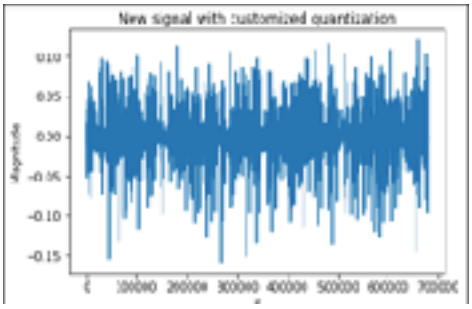
τρόπο υπολογίζονται τα βασικά στοιχεία του non-adaptive κβαντιστή με Xm=2 και Bk=8. Η κβαντοποίηση γίνεται για κάθε πλάισιο του σήματος μας, όπως αυτό παραθυροποιήθηκε στο μέρος 1.

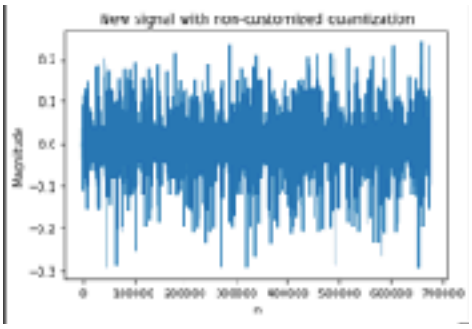
Μετά υπερδειγματοληπτούμε τα δύο κβαντησμένα σήματα και τα αποθηκεύουμε ως ws και ws2. Στην συνέχεια συνελίσουμε το καθένα από αυτα με τα φίλτρα ανάλυσης Gk.

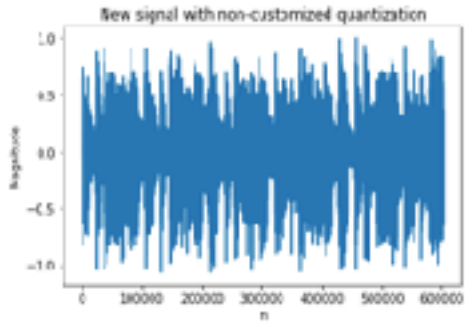
Τα δύο σήματα εξόδου θα είναι το αποτέλεσμα της άθροισης των 32 φιλτραρισμένων σημάτων που υπολογίστηκαν πριν. Προκειμένου να πετύχουμε την απεικόνιση των σημάτων, χρησιμοποιούμε την εντολή reshape για να τα μετατρέψουμε τα σήματα μονοδιάστατα.



Παρακάτω απεικονίζονται τα δύο σήματα εξόδου και το αρχικό για σύγκριση.







Στο άκουσμα των δύο σημάτων είναι αισθητή η διαφορά με το non-adaptive να προσεγγίζει καλύτερα το αρχικό.