



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

**ΜΕΛΕΤΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΗΜΑΤΩΝ ΠΑΡΑΘΥΡΟΥ
ΣΕ ΠΕΡΙΒΑΛΛΟΝ
ΣΧΕΔΟΝ ΑΡΜΟΝΙΚΩΝ ΜΟΝΤΕΛΩΝ**

ΤΗΛΕΜΑΧΟΣ Δ. ΕΛΕΥΘΕΡΙΟΥ

Διπλωματική Εργασία υποβληθείσα στο
Τμήμα Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης
ως μέρος των απαιτήσεων για την απόκτηση Πτυχίου.

ΥΠΕΥΘΥΝΟΣ ΚΑΘΗΓΗΤΗΣ : ΙΩΑΝΝΗΣ ΣΤΥΛΙΑΝΟΥ
ΕΠΙΒΛΕΠΩΝ : ΓΕΩΡΓΙΟΣ ΚΑΦΕΤΖΗΣ

ΗΡΑΚΛΕΙΟ
ΜΑΡΤΙΟΣ 2016

Στην οικογένεια μου ...



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

**ΜΕΛΕΤΗ ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΗΜΑΤΩΝ ΠΑΡΑΘΥΡΟΥ
ΣΕ ΠΕΡΙΒΑΛΛΟΝ
ΣΧΕΔΟΝ ΑΡΜΟΝΙΚΩΝ ΜΟΝΤΕΛΩΝ**



ΤΗΛΕΜΑΧΟΣ Δ. ΕΛΕΥΘΕΡΙΟΥ

ΥΠΕΥΘΥΝΟΣ ΚΑΘΗΓΗΤΗΣ : ΙΩΑΝΝΗΣ ΣΤΥΛΙΑΝΟΥ
ΕΠΙΒΛΕΠΩΝ : ΓΕΩΡΓΙΟΣ ΚΑΦΕΝΤΖΗΣ

ΗΡΑΚΛΕΙΟ
ΜΑΡΤΙΟΣ 2016

Ευχαριστίες

Στην διαδρομή μου σε όλα τα στάδια της εκπαίδευσης μέχρι και την ανάδειξή μου ως πτυχιούχο Επιστήμης Υπολογιστών, υπήρξαν αρκετοί άνθρωποι που με βοήθησαν και θέλω να ευχαριστήσω με αυτό το κείμενο. Θέλω αρχικά να ευχαριστήσω τους καθηγητές μου που υποστήριξαν και μου έδωσαν αυτή την εργασία, Στυλιανού Ιωάννη και Γιώργο Καφεντζή, για την βοήθεια αλλά και τις πολλές εξειδικευμένες γνώσεις που έλαβα κατά την εκπόνηση της παρούσας εργασίας. Ο Γιώργος, που ήταν και ο επόπτης μου, στάθηκε δίπλα μου σε καθημερινή βάση για σχεδόν ένα εξάμηνο και είμαι πραγματικά ευγνώμων για την καθοδήγηση, την ενθάρυνση και κυρίως την εμπιστοσύνη και την υπομονή που απλόχερα μου έδειχνε κάθε στιγμή που χρειαζόμουν την βοήθειά του.

Επιπλέον, θέλω να ευχαριστήσω τους καθηγητές μου που με βοήθησαν να εισέλθω σε ένα από τα καλύτερα Πανεπιστήμια της χώρας. Ευχαριστώ μέσα από την καρδιά μου για τις γνώσεις και την διαπαιδαγώγηση που μου προσέφεραν, τον μαθηματικό μου, Δημήτρη Δελληγιάννη, τον φυσικό μου, Κωστή Δελληγιάννη και την φιλόλογό μου, «Ρωξάνη». Δεν θα ήμουν όμως σε αυτή τη σχολή χωρίς την καθοδήγηση, την υποστήριξη και τις γνώσεις που μου προσέφερε ο Αιμίλιος Μπακατσέλος. Είναι ο άνθρωπος που μου «μεταλαμπάδευσε το μικρόβιο» του πραγματισμού και με ώθησε να επιλέξω το συγκεκριμένο τμήμα για τις σπουδές μου.

Στη συνέχεια θέλω να ευχαριστήσω για την υποστήριξη και την φιλία τους όλα αυτά τα χρόνια τον Στράτο Σουλιωτάκη, τον Παναγιώτη Τσακιρίδη και τον Νίκο Τσιώτα, οι οποίοι στάθηκαν σαν αδέρφια δίπλα μου και με βοήθησαν με κάθε τρόπο να αφοσιωθώ στις σπουδές μου. Επίσης, δεν θέλω να παραλείψω τους Γιώργο Καραδήμο, Νίκο Ρηγάκη, Ελευθερία Τσιούχα Ροδούλα Κυριακάκη και Χρυσή Παρίση με τους οποίους περάσαμε τέλεια φοιτητική ζωή και οι αναμνήσεις από αυτά τα χρόνια θα μείνουν βαθιά χαραγμένες μέσα μας. Τέλος, είμαι ιδιαίτερα ευγνώμων στην κοπέλα μου, Ζέτα Ακτουδιανάκη, για την αγάπη, την φροντίδα, την ανεκτικότητα και την στήριξή της, γιατί ήταν θεμελιώδης παράγοντας στην στροφή μου προς τη μελέτη και την ολοκλήρωση των σπουδών μου.

Τελευταίους αλλά όχι εσχάτους, θα ήθελα να ευχαριστήσω θερμά την οικογένειά μου: τους γονείς μου Δημήτρη Ελευθερίου και Ελένη Ψυχογιού, για τα πάντα. Τους χρωστάω αυτό που είμαι σήμερα. Τον αδερφό μου, Γιώργο, γιατί απλά είναι το σημαντικότερο στήριγμα στη ζωή μου. Χρωστάω πάρα πολλά και στους παππού Γιώργιο και την γιαγιά Αλεξάνδρα που η αγάπη και η στήριξη τους όλα αυτά τα δύσκολα χρόνια, μου έδινε δύναμη να συνεχίσω, καθώς και στην γιαγιά Κατίνα για την έγνοια και την αγάπη της. Δεν θα μπορούσα να παραλείψω τον παππού Τηλέμαχο, που έκανε το «μεγάλο ταξίδι» στην αρχή της ακαδημαϊκής μου εκπαίδευσης, τον ευχαριστώ για όσα μου δίδαξε καθώς και για όλα όσα προσέφερε στην οικογένειά μας και εύχομαι να καμαρώνει για μένα από εκεί ψηλά. Τέλος, ευχαριστώ από τα βάθη της καρδιάς μου τους καθηγητές της ζωής: Δημήτρη και Θανάση Ψυχογιό καθώς και τον Σπύρο Σουλιωτάκη που με διαπαιδαγώγισαν όπως μόνο αυτοί ξέρουν και μου έδωσαν αξίες που κανένα Πανεπιστήμιο και Κολλέγιο του κόσμου δεν θα μπορούσε ποτέ να μου προσφέρει.

Περίληψη

Μία από τις πιο ευρέως διαδεδομένες και χρησιμοποιούμενες παραμετρικές μεθόδους για την επεξεργασία σήματος φωνής και ήχου, είναι η ημιτονοειδής μοντελοποίηση. Η ακριβής αναπαράσταση των σημάτων που αναλύονται, είναι άμεσο αποτέλεσμα του ακριβούς υπολογισμού των βασικών ημιτονοειδών παραμέτρων (πλάτη, συχνότητες, φάσεις).

Πρόσφατα προταθέντα μοντέλα, παράγουν εξαιρετικά ορθούς υπολογισμούς των ημιτονοειδών αυτών παραμέτρων. Συνεπώς, αναπαριστούν σήματα φωνής με ορθό και συμπαγή τρόπο. Τα εν λόγω μοντέλα είναι, το Σχεδόν-Αρμονικό Μοντέλο (QHΜ), το Προσαρμόσιμο Σχεδόν-Αρμονικό Μοντέλο (aQHΜ) και το εκτεταμένο Προσαρμόσιμο Σχεδόν-Αρμονικό Μοντέλο (eaQHΜ).

Όπως όλες οι τεχνικές ανάλυσης σημάτων, έτσι και τα παραπάνω μοντέλα κόβουν το σήμα σε μικρά πλαίσια και εφαρμόζουν σε αυτά κάποιο παράθυρο ανάλυσης για να εκτιμήσουν τις παραμέτρους του μοντέλου, παράγοντας πολύ υψηλή ποιότητα στην ανασύνθεση του σήματος φωνής. Υπάρχει ένα μεγάλο εύρος τέτοιων παραθύρων ανάλυσης, το καθένα με τα δικά του χαρακτηριστικά, τα πεδία του χρόνου και της συχνότητας.

Η παρούσα εργασία, βασίζεται στη μελέτη των σχεδόν-αρμονικών μοντέλων, την σύγκριση απόδοσης εναλλακτικών παραθύρων στο QHΜ, καθώς και την εφαρμογή αυτών με τα βέλτιστα παραχθέντα αποτελέσματα, στην τελευταία χρονικά έκδοση του μοντέλου (eaQHΜ), με σκοπό την βελτιστοποίηση του υπολογισμού των ημιτονοειδών παραμέτρων.

Περιεχόμενα

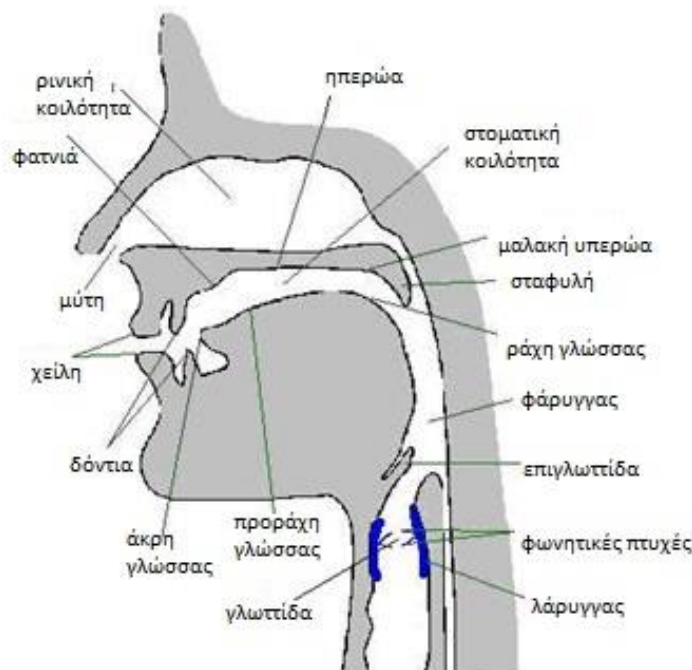
Τίτλος	7
Ευχαριστίες	9
Περίληψη	11
1 Εισαγωγή	
1.1 Μηχανισμός παραγωγής φωνής	15
1.2 Γραμμικό σύστημα	16
1.3 Ημιτονοειδή μοντέλα	17
1.4 Η εργασία	17
2 Θεωρητικό Μέρος	
2.1 Προσαρμοστικά Ημιτονοειδή Μοντέλα	19
2.2 Quasi Harmonic Model (QHM)	20
2.3 Adaptive Quasi Harmonic Model (aQHM)	24
2.4 Extended-adaptive Quasi Harmonic Model (eaQHM)	26
2.4.1 Αλγόριθμος eaQHM	28
2.5 Πλήρους εύρους (full-band) eaQHM	28
2.5.1 Ανάλυση	29
2.5.2 Προσαρμοστικότητα	30
2.5.3 Σύνθεση	31
2.5.4 Παράδειγμα ανακατασκευής σήματος	33
3 Πειραματικό Μέρος	
3.1 Συνθετικά Πειράματα	35
3.1.1 Εισαγωγή	35
3.1.2 Παράθυρα Ανάλυσης	35
3.1.3 Bias	37
3.1.4 MSE	39
3.1.5 Συμπεράσματα	43
3.2 Πειράματα σε πραγματικά σήματα	45
3.2.1 Εισαγωγή	45
3.2.2 Βάση Δεδομένων	46
3.2.3 Δομή Πειράματος	46
3.2.4 Αποτέλεσμα Πειράματος	47
3.2.5 Συμπεράσματα	49
4 Συμπέρασμα και μελλοντική εργασία	51
5 Παράρτημα	53
6 Βιβλιογραφία	73

1 Εισαγωγή

1.1 Μηχανισμός παραγωγής φωνής

Αναμφίβολα, η ομιλία αποτελεί την σημαντικότερη μορφή επικοινωνίας μεταξύ των ανθρώπων ανα τους αιώνες. Η ομιλία βασίζεται στο φωνητικό σύστημα, το οποίο και συμβάλλει στην παραγωγή της ανθρώπινης φωνής.

Ο τρόπος λειτουργίας του φωνητικού συστήματος είναι αρκετά πολύπλοκος. Πράγματι, αν τμηματοποιήσουμε τη διαδικασία παραγωγής του λόγου θα διαπιστώσουμε ότι η φωνή παράγεται με τις κινήσεις των φωνητικών χορδών που βρίσκονται στο λάρυγγα. Ο αέρας από τους πνεύμονες περνάει μεταξύ των φωνητικών χορδών, τις κάνει να δονούνται και παράγουν ήχο. Η ομιλία γενικά και ειδικότερα η άρθρωση του λόγου, εκτελούνται με τις βασικές λειτουργίες της στοματογναθικής περιοχής. Στη διαδικασία αυτή συμμετέχει ολόκληρη η στοματική κοιλότητα, με τη συνεργασία της γλώσσας, της υπερώας, των δοντιών και των χειλιών. (Σχήμα 1.1)



Σχήμα 1.1

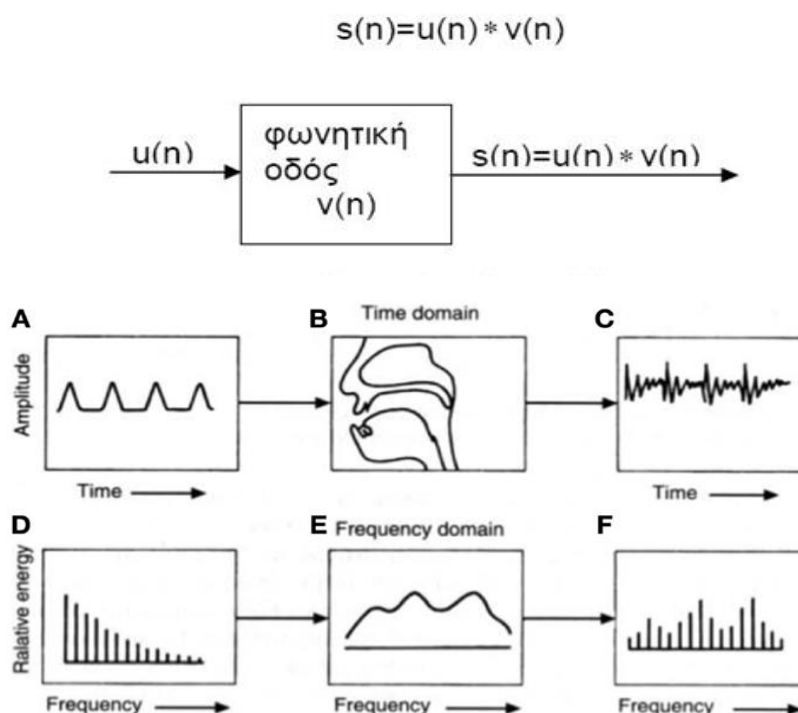
Ανατομία φωνητικής διόδου

Ακριβέστερα, η φωνή του ανθρώπου και οι αλλαγές της σε ένταση, ύψος, τόνο και χροιά, διαμορφώνουν την ατομική εκείνη εικόνα της ομιλίας, η οποία την καθιστά ευχάριστη ή δυσάρεστη στην ακοή. Η βάση και ο σημαντικότερος παράγοντας για την ομαλή εξέλιξη του λόγου είναι η αναπνοή.

Με αφετηρία την παραγωγή της φωνής καθώς και τη διαφορετικότητά της από άνθρωπο σε άνθρωπο, η επιστημονική κοινότητα άρχισε να αναζητά τρόπους για την επεξεργασία και ανάλυση της ανθρώπινης φωνής με αποτέλεσμα να καταλήξει σε αρκετά μαθηματικά μοντέλα, τα οποία μπορούν να χωριστούν χονδρικά σε δύο “στρατόπεδα” , τα οποία παρατίθενται συνοπτικά.

1.2 Γραμμικό σύστημα

Το σήμα φωνής είναι ιδιαίτερα πολύπλοκο, εξαιτίας των διαφορετικών φύσεων διεργασιών που συντελούν στη σχεδόν περιοδικότητα του έμφωνου λόγου και στη μη στατικότητα του. Κατά συνέπεια, η περιγραφή της παραγωγής ομιλίας με ένα ακριβές μαθηματικό μοντέλο είναι εξαιρετικά δύσκολη. Ωστόσο, σε ένα σχετικά απλό μοντέλο μπορεί να θεωρηθεί το φωνητικό σύστημα σαν ένα γραμμικό χρονικά αναλλοίωτο σύστημα (linear – time invariant, LTI). Σε αυτή την περίπτωση το σήμα εξόδου $s(n)$ μπορεί να θεωρηθεί ότι προέρχεται από τη συνέλιξη του σήματος στην έξοδο των φωνητικών χορδών $u(n)$ με τη συνάρτηση $v(n)$, η οποία περιγράφει τα χαρακτηριστικά της φωνητικής οδού. (Σχήμα 1.2).



Σχήμα 1.2

Διαμόρφωση σήματος ήχου στο χρόνο και στη συχνότητα

Αν η $u(n)$ είναι περιοδική τότε και η $s(n)$ είναι περιοδική . Στην πραγματικότητα ωστόσο, το σύστημα μπορεί να είναι μη γραμμικό με αποτέλεσμα οι φωνητικοί ήχοι $s(n)$ να μην περιγράφονται σωστά από LTI συστήματα.

Γενικά, ένα σήμα ομιλίας είναι μια ακολουθία ήχων. Το σήμα αλλάζει σχετικά αργά με το χρόνο, μεταξύ 5 – 30msec τα χαρακτηριστικά του είναι σχεδόν σταθερά, οπότε και το σήμα σε ένα τέτοιο διάστημα θεωρείται στάσιμο.

1.3 Ημιτονειδή μοντέλα

Η ανάγκη για ένα πιο αξιόπιστο μαθηματικό μοντέλο αναπαράστασης της ανθρώπινης φωνής οδήγησε τους επιστήμονες στην εφαρμογή ημιτονοειδών μοντέλων. Σε αυτά τα μοντέλα, το σήμα φωνής αναπαρίσταται σαν ένα αθροισμα απο ημίτονα με παραμέτρους πλάτη, συχνότητα και φάσεις και η χρήση τους δίνει ένα πανίσχυρο πλαίσιο ανάλυσης και «τυποποίησης» των μεθόδων που χρησιμοποιούνται. Το βασικότερο και πιο γνωστό εργαλείο ανάλυσης σημάτων είναι η ανάλυση σε σειρες Fourier, η οποία μας δίνει πληροφορίες για το συχνοτικό τους περιεχόμενο. Αυτό πρακτικά σημαίνει οτι μας πληροφορεί για το ποιά και πόσα ημίτονα πρέπει να “προσθέσουμε” για να αναπαραστήσουμε μαθηματικά το σήμα που αναλύουμε.

Ωστόσο, οι προσεγγίσεις των αρχικά σχεδιασμένων ημιτονοειδών μοντέλων, είχαν ως δεδομένο την τοπική στατικότητα στα διαστήματα του αναλυθέντος σήματος, δηλαδή θεωρούσαν ότι το σήμα έμενε ίδιο για κάποιο χρονικό διάστημα . Η τοπική στατικότητα είναι μία υπόθεση που δεν ισχύει στην πράξη. Το γεγονός αυτό δημιούργησε την ανάγκη για τρόπους ημιτονοειδούς μοντελοποίησης με ένα σύνολο από μη-παραμετρικές και μη-στατικές συναρτήσεις βάσης.

1.4 Η εργασία

Στην παρούσα εργασία θα παρουσιαστούν κάποια απο τα πιο πρόσφατα προταθέντα μη-στατικά ημιτονοειδή μοντέλα για την ανάλυση και ανασύνθεση φωνής. Τα μοντέλα αυτά , βασίζονται στην ιδέα της σχεδόν-αρμονικότητας . Επίσης , θεμελιώδης είναι η έννοια της προσαρμοστικότητας. Η γενική ιδέα του μοντέλου αυτού είναι ότι παρέχει ένα μηχανισμό διόρθωσης της αρχικά υπολογισμένης συχνότητας με σκοπό την επανεκτίμησή της , όσο το δυνατόν κοντινότερη στην πραγματική τιμή της.

Στο πειραματικό στάδιο , και αφού θα έχει επεξηγηθεί πλήρως ο τρόπος λειτουργίας των μοντέλων και αποσαφηνιστεί η έννοια του παραθύρου που αυτά χρησιμοποιούν για την σύνθεση του σήματος, θα μελετηθεί πώς η εναλλαγή του τύπου παραθύρου θα επηρεάσει την ποιότητα του ανασυνθέντος τελικού σήματος.

2 Θεωρητικό Μέρος

2.1 Προσαρμοστικά Ημιτονοειδή Μοντέλα

Ο όρος «προσαρμοστικά ημιτονοειδή μοντέλα» (adaptive Sinusoidal Models – aSM) αναφέρεται σε μία οικογένεια ημιτονοειδών μοντέλων, τα οποία έχουν τη δυνατότητα να προσαρμόζουν τις βασικές ημιτονοειδείς παραμέτρους (στιγμιαία φάση και πλάτος), του σήματος φωνής το οποίο αναλύεται. Η μοντελοποίηση αυτή, έχει ως δεδομένο ότι τα σήματα φωνής είναι μη-στατικά και το γεγονός αυτό μας βοηθά στην ορθότερη προσέγγιση των σημάτων σε σχέση με τα στατικά ημιτονοειδή μοντέλα που βασίζονταν στην ψευδή έννοια της τοπικής στατικότητας. Στο πλαίσιο αυτό είναι καλό να αποσαφηνιστεί η έννοια της προσαρμοστικότητας.

Δεδομένου ενός συνόλου από πολύπλοκες βασικές ημιτονοειδείς συναρτήσεις, το οποίο αντικατοπτρίζει ένα σήμα φωνής με μαθηματική μορφή, **προσαρμοστικότητα** είναι μία επαναλαμβανόμενη ανακατασκευή αυτού του συνόλου συναρτήσεων του σήματος. Η ανακατασκευή αφορά το πλάτος ή/και την συχνότητα του σήματος και επιτυγχάνει να βελτιώσει μετά από κάποιο πεπερασμένο αριθμό επαναλήψεων τις στιγμιαίες παραμέτρους του σήματος (στιγμιαία συχνότητα, πλάτος και φάση).

Αρχικά, τα aSM μοντέλα θεμελιώθηκαν σαν θεωρία για την αναπαράσταση τμημάτων ενός σήματος, σε ένα σύνολο από μη-παραμετρικές και μη-στατικές ημιτονοειδείς συναρτήσεις. Στα μικρά αυτά τμήματα, όπου εφαρμόζεται κάποιο παράθυρο ανάλυσης, τα συμβατικά μοντέλα θεωρούν ότι τα χαρακτηριστικά του λόγου μένουν αναλλοίωτα σε τοπικό επίπεδο. Με δεδομένο την θέση αυτή, οι βασικές συναρτήσεις παραμένουν σταθερές σε πλάτος και συχνότητα διευκολύνοντας τον τοπικό υπολογισμό τους.

Γενικά, ένα aSM μπορεί να αναπαρασταθεί ως το άθροισμα :

$$x(t) = \left(\sum_{k=-K}^K C_k(t) \psi_k(t) \right) w(t)$$

όπου το $\psi_k(t)$ αναπαριστά το σύνολο των μη-στατικών συναρτήσεων βάσης, το $C_k(t)$ το πλάτος, $2K+1$ είναι ο αριθμός των εκθετικών του αθροίσματος και τέλος $w(t)$ είναι το παράθυρο ανάλυσης που είναι μη μηδενικό στο $[-T, T]$. Στα συμβατικά ημιτονοειδή μοντέλα, όπως το Sinusoidal Model [MQ86], το Harmonic Model [LSM93] κ.α., το σύνολο των συναρτήσεων βάσης κατά τη διάρκεια της ανάλυσης παρουσιάζει στατικότητα σε συχνότητα και πλάτος. Για παράδειγμα, η συνάρτηση βάσης στο Sinusoidal Model έχει τη μορφή:

$$\psi_k^{SM}(t) = 1 \cdot e^{j2\pi f_k t}, \quad C_k^{SM}(t) = a_k$$

Όπου οι συχνότητες και τα πλάτη είναι σταθερά μέσα στο παράθυρο ανάλυσης. Ωστόσο τα προσαρμοστικά μοντέλα είναι μη-παραμετρικά και βασίζονται στα πραγματικά χαρακτηριστικά του αναλυόμενου σήματος. Έτσι, προκύπτει η βασική συνάρτηση από τον παρακάτω τύπο:

$$\psi_k^{aSM}(t) = \alpha_k(t) \cdot e^{j\phi_k(t)},$$

όπου $\alpha_k(t)$ είναι το στιγμιαίο πλάτος της k -οστής συνάρτησης βάσης, $\phi_k(t)$ η στιγμιαία συχνότητα της k -οστής συνάρτησης βάσης.

Στη συνέχεια, θα παρουσιάσουμε τρία μοντέλα που βασίζονται στην προσαρμοστικότητα και την σχεδόν-αρμονικότητα για την ανάλυση του σήματος φωνής.

Το πρώτο βήμα στα aSM έγινε με το Σχεδόν-Αρμονικό Μοντέλο (Quasi Harmonic Model – QHM), το οποίο μελετήθηκε και παρουσιάστηκε αρχικά από τον Στυλιανού Ιωάννη [Sty96] και στη συνέχεια αναλύθηκε και επεκτάθηκε από τον Ιωάννη Πανταζή [PRS08] [PRS09a] [PRS09b] [PRS10] σε προσαρμοστικό Σχεδόν-Αρμονικό Μοντέλο (adaptive Quasi Harmonic Model – aQHM). Στην τελευταία έκδοση του, αυτή που χρησιμοποιήθηκε και για τα πραγματικά πειράματα της παρούσας εργασίας, το μοντέλο παρουσιάστηκε από τον Καφεντζή Γεώργιο [KOY11] σαν εκτεταμένο προσαρμοστικό Σχεδόν Αρμονικό Μοντέλο (extended-adaptive Quasi Harmonic Model – eaQHM). Όλα όμως τα παραπάνω μοντέλα βασίζονται στο μηχανισμό διόρθωσης συχνότητας του Σχεδόν Αρμονικού Μοντέλου (QHM), από το οποίο και θα ξεκινήσουμε τη σχετική συζήτηση.

2.2 Quasi Harmonic Model

Το QHM δεν είναι προσαρμοστικό από τη φύση του, παρ' όλα αυτά παρέχει τον μηχανισμό προσαρμογής τον οποίο χρησιμοποιούν τα προσαρμοστικά σχεδόν-αρμονικά μοντέλα. Ο μηχανισμός αυτός βασίζεται στη διόρθωση της συχνότητας μέσω ενός μηχανισμού εκτίμησης της διαφοράς, ανάμεσα στην πραγματική και την αρχικά-υπολογισμένη συχνότητα. Στη συνέχεια, αυτή η διαφορά προστίθεται στην ήδη υπολογισμένη συχνότητα με σκοπό την ανακατασκευή ενός σήματος με συχνότητα που είναι πολύ κοντά σε αυτή του αυθεντικού σήματος.

Γενικότερα, όπως αναφέραμε στην ημιτονοειδή μοντελοποίηση, ένα σήμα φωνής μπορεί να αναπαρασταθεί με το παρακάτω άθροισμα.

$$x(t) = \left(\sum_{k=-K}^K a_k e^{j2\pi f_k t} \right) w(t),$$

όπου $2K+1$ ο αριθμός των εκθετικών με πλάτη a_k , συχνότητες f_k και παράθυρο ανάλυσης $w(t)$. Ωστόσο, η συχνότητα f_k αναφέρεται στις διορθωμένες συχνότητες των αρμονικών του σήματος. Στα ημιτονοειδή μοντέλα, έχουμε αρχικά έναν κατα προσέγγιση υπολογισμό με μεθόδους όπως το peak-picking ή λαμβάνοντας υπόψιν αρμονικές της θεμελιώδους συχνότητας. Οι κατα προσέγγιση συχνότητες θα συμβολίζονται με \hat{f}_k . Έτσι, μπορούμε να γράψουμε:

$$f_k = \hat{f}_k + \eta_k, \quad k = -K, \dots, K$$

Αν το σφάλμα η_k , είναι ψηλό τότε και η εκτίμηση της σύνθεσης των πλατών a_k , έχει μεγάλο bias. Κατά την αναζήτηση λύσης σε αυτό το πρόβλημα, στις δημοσιεύσεις [PRS09a] και [PRS08], παρουσιάζεται η χρήση του QHM για την ανασύνθεση σημάτων φωνής και προτάθηκε η σχέση:

$$x(t) = \left(\sum_{k=-K}^K (a_k + tb_k) e^{j2\pi \hat{f}_k t} \right) w(t),$$

όπου το b_k αναπαριστά την κλίση της k -οστής συνιστώσας. Στο πεδίο της συχνότητας, η k -οστή συνιστώσα αναπαριστάται ως :

$$X_k(f) = a_k W(f - \hat{f}_k) + j \frac{b_k}{2\pi} W'(f - \hat{f}_k)$$

όπου $W(f)$, ο μετασχηματισμός Fourier του παραθύρου ανάλυσης και $W'(f)$ η παράγωγος του $W(f)$. Στο [PRS08], αποδείχθηκε ότι το QHM έχει τη δυνατότητα να διορθώσει τη διαφορά της υπολογισμένης συχνότητας από την πραγματική, χρησιμοποιώντας την προβολή του b_k πάνω στο a_k :

$$b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k$$

όπου $j a_k$ αναπαριστά το κάθετο διάνυσμα στο a_k . Οι παράμετροι ρ_1 και ρ_2 υπολογίζονται :

$$\rho_1 = \frac{\Re\{a_k\}\Re\{b_k\} + \Im\{a_k\}\Im\{b_k\}}{|a_k|^2},$$

$$\rho_2 = \frac{\Re\{a_k\}\Im\{b_k\} - \Im\{a_k\}\Re\{b_k\}}{|a_k|^2},$$

Τα $\Re\{a_k\}$, $\Re\{b_k\}$ και $\Im\{a_k\}$, $\Im\{b_k\}$ αναπαριστούν το πραγματικό και φανταστικό μέρος των a_k και b_k αντίστοιχα. Με πρόσθεση στο $X_k(f)$ και σύμφωνα με το ανάπτυγμα σε σειρές Taylor του $W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi})$, και της τιμής του $W''(f)$ με \hat{f}_k να παίρνει την ελάχιστη δυνατή τιμή της, τότε για μικρές τιμές του $\rho_{2,k}$, μπορεί να αποδειχθεί [PRS08] ότι η k -οστή συνιστώσα μπορεί να αναπαρασταθεί στο πεδίο της συχνότητας ως:

$$X_k(f) = a_k \left[W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j \frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

Εφαρμόζοντας το ανάπτυγμα σε σειρά Taylor στο $W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi})$, παράγεται ότι :

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

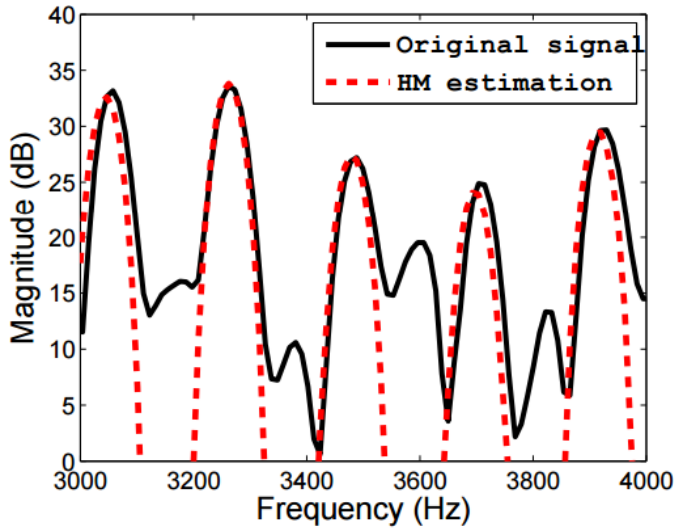
Και η τιμή του $W''(f)$ στη συχνότητα \hat{f}_k όσο το δυνατόν μικρότερη, τότε για μικρές τιμές του $\rho_{2,k}$, μπορεί να αποδειχθεί [PRS08] ότι η k -οστή συνιστώσα μπορεί να αναπαρασταθεί στο πεδίο του χρόνου ως:

$$x_k(t) \approx a_k \left[e^{j(2\pi \hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi \hat{f}_k t} \right] w(t)$$

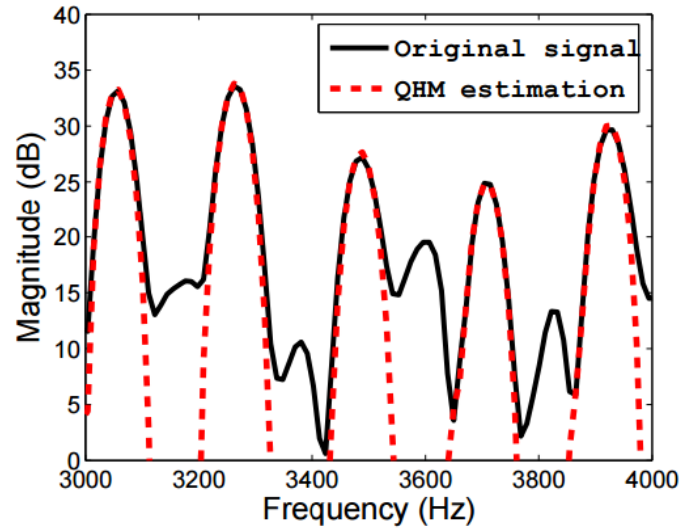
Έτσι, η παράμετρος $\rho_{2,k}/2\pi$ εκτιμά το σφάλμα η_k :

$$\hat{\eta}_k = \frac{\rho_{2,k}}{2\pi} = \frac{1}{2\pi} \frac{\Re\{a_k\}\Im\{b_k\} - \Im\{a_k\}\Re\{b_k\}}{|a_k|^2},$$

όπου το $\rho_{1,k}$ υπολογίζει την κανονικοποιημένη κλίση του πλάτους της k -οστής συνιστώσας. Στη συνέχεια, αποδεικνύεται επίσης ότι η διόρθωση εξαρτάται από το μέγεθος του $\rho_{2,k}$ και την τιμή του $W''(f)$ στη συχνότητα f_k . Ο μηχανισμός διόρθωσης του σφάλματος συχνότητας, απεικονίζεται στα σχήματα 2.2.1 1 (α) και (β) όπου στο πρώτο οι αρμονικές υπολογίζονται ως πολλαπλάσιες της θεμελιώδους συχνότητας f_0 , ενώ στο δεύτερο έχει εφαρμοστεί ο μηχανισμός διόρθωσης συχνότητας.



(Α) Χωρίς διόρθωση συχνότητας



(Β) Με διόρθωση συχνότητας

Σχήμα 2.2.1

Διόρθωση συχνότητας με τη χρήση του QHM.

Στη συνέχεια, ας δούμε πώς υπολογίζονται οι παράμετροι a_k, b_k του μοντέλου :

$$\mathbf{x} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$$

Το σφάλμα ορίζεται σε διακριτό χρόνο ως το άθροισμα:

$$\begin{aligned} \epsilon(\mathbf{a}, \mathbf{b}) &= \sum_{n=-N}^N |s[n] - s_q[n]|^2 \\ &= \sum_{n=-N}^N (s[n] - s_q[n])^* (s[n] - s_q[n]) \end{aligned}$$

όπου $s[n]$, το αυθεντικό σήμα αφού έχει εφαρμοστεί το παράθυρο ανάλυσης, $s_q[n]$ η σχεδόν-αρμονική αναπαράσταση του αρχικού σήματος και $2N+1$ το μέγεθος του παραθύρου. Στη σημειογραφία πινάκων, αν διαχωριστούν οι τιμές του παραθύρου από τα δείγματα, ο υπολογισμός του σφάλματος γίνεται:

$$\begin{aligned} \epsilon(\mathbf{a}, \mathbf{b}) &= (\mathbf{W}\mathbf{s} - \mathbf{W}\mathbf{s}_q)^H (\mathbf{W}\mathbf{s} - \mathbf{W}\mathbf{s}_q) \\ &= (\mathbf{W}(\mathbf{s} - \mathbf{s}_q))^H \mathbf{W}(\mathbf{s} - \mathbf{s}_q) \\ &= (\mathbf{s} - \mathbf{s}_q)^H \mathbf{W}^H \mathbf{W}(\mathbf{s} - \mathbf{s}_q) \end{aligned}$$

Με το \mathbf{W} να αναπαριστά ένα τετραγωνικό πίνακα ο οποίος έχει τις τιμές του παραθύρου ανάλυσης στη διαγώνιό του, \mathbf{s} είναι τα δείγματα του αυθεντικού σήματος σε ένα διάλυσμα και το \mathbf{H} δηλώνει τον τελεστή Hermitian στον τετραγωνικό πίνακα.

Με τα παραπάνω δεδομένα, το QHM μπορεί να γραφεί σε μορφή πινάκων ως:

$$\begin{aligned} s_q[n] &= \sum_{n=-N}^N (a_k + nb_k) e^{j2\pi f_k n / f_s} \\ &= \sum_{n=-N}^N a_k e^{j2\pi f_k n / f_s} + \sum_{n=-N}^N nb_k e^{j2\pi f_k n / f_s} \\ \mathbf{s}_q &= \mathbf{E}_0 \mathbf{a} + \mathbf{E}_1 \mathbf{b} = [\mathbf{E}_0 | \mathbf{E}_1] \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{E} \mathbf{x} \end{aligned}$$

Όπου $\mathbf{E}_0 = (E_0)_{n,k} = e^{j2\pi \frac{f_k n}{f_s}}$, $\mathbf{E}_1 = (E_1)_{n,k} = n(E_0)_{n,k} = ne^{j2\pi \frac{f_k n}{f_s}}$

Και $\mathbf{E} = [\mathbf{E}_0 | \mathbf{E}_1]$

Επομένως, η ελαχιστοποίηση γίνεται θέτοντας :

$$\begin{aligned} \frac{\partial \epsilon(\mathbf{x})}{\partial \mathbf{x}} &= 0 \\ \frac{\partial}{\partial \mathbf{x}} (\mathbf{s} - \mathbf{E} \mathbf{x})^H \mathbf{W}^H \mathbf{W} (\mathbf{s} - \mathbf{E} \mathbf{x}) &= 0 \end{aligned}$$

Η λύση δίνεται απο την εξίσωση:

$$\mathbf{x} = \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{bmatrix} = (\mathbf{E}^H \mathbf{W}^H \mathbf{W} \mathbf{E})^{-1} \mathbf{E}^H \mathbf{W}^H \mathbf{W} \mathbf{s}$$

Τελικά το σήμα μπορεί τοπικά να υπολογιστεί ως :

$$x(t) = \sum_{k=-K}^K |\hat{a}_k| e^{j(2\pi(\hat{f}_k + \hat{\eta}_k)t + \hat{\phi}_k)} w(t)$$

όπου $\hat{\phi}_k = \angle \hat{a}_k$

2.3 The adaptive Quasi-Harmonic Model, aQHM

Ωστόσο, η υπόθεση της τοπικής στατικότητας στην οποία στηρίζονται όλα τα μοντέλα που προαναφέρθηκαν, ελαφρύνθηκε αργότερα από ένα νέο μοντέλο [PRS11], το προσαρμοστικό σχεδόν Αρμονικό Μοντέλο (aQHM), το οποίο αναπαρίσταται από τη σχέση:

$$x(t) = \left(\sum_{k=-K}^K (a_k + tb_k) e^{j(\hat{\phi}_k(t+t_i) - \hat{\phi}_k(t_i))} \right) w(t), \quad t \in [-T, T]$$

όπου $\phi_k(t)$ δηλώνει την στιγμιαία φάση στην k -οστή συνιστώσα και t_i , το κέντρο του παραθύρου ανάλυσης. Ο όρος b_k παρέχει τον τρόπο της ανανέωσης της συχνότητας του ανακατασκευασμένου σήματος στο κέντρο του παραθύρου t_i , όπως ακριβώς και στο QHM. Δεδομένων των δειγμάτων του αναλυόμενου σήματος στο διάνυσμα s , οι παράμετροι εισόδου του μοντέλου παράγονται από τη μέθοδο Least Squares, όπως στο QHM, οπότε παίρνουμε:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (\mathbf{E}^H \mathbf{W}^H \mathbf{W} \mathbf{E})^{-1} \mathbf{E}^H \mathbf{W}^H \mathbf{W} s$$

όπου W είναι ο πίνακας που περιέχει τις τιμές του παραθύρου στη διαγώνιό του, s είναι το διάνυσμα του εισαγόμενου σήματος, ο πίνακας E μπορεί να αναπαρασταθεί από την σχέση $E = [E_0 | E_1]$, οι υποπίνακες E_i , $i=0,1$, έχουν στοιχεία που δίνονται από τις σχέσεις:

$$(E_0)_{n,k} = e^{j(\phi_k(t_n+t_i) - \phi_k(t_i))}$$

και
$$(E_1)_{n,k} = t_n e^{j(\phi_k(t_n+t_i) - \phi_k(t_i))} = t_n (E_0)_{n,k},$$

και τέλος, η στιγμιαία φάση της k -οστής συνιστώσας μπορεί να υπολογιστεί ως:

$$\hat{\phi}_k(t) = \hat{\phi}_k(t_i) + \int_{t_i}^{t_i+t} 2\pi f_k(u) du, \quad t \in [-T, T],$$

με $f_k(t)$ να είναι η καμπύλη της συχνότητας της k -οστής συνιστώσας.

Χρησιμοποιώντας τον ορισμό της συχνότητας, η στιγμιαία φάση μιας συνιστώσας, $\phi(t)$, υπολογίζεται ως το ολοκλήρωμα της στιγμιαίας συχνότητας, $f(t)$, η οποία παρέχεται από την αρχική εκτίμηση των παραμέτρων, όπως ακριβώς και στο QHM. Έτσι παρεμβάλλοντας τιμές φάσης ανάμεσα σε δύο χρονικά στιγμιότυπα t_i , t_{i+1} , καταλήγουμε στην παρακάτω ισότητα:

$$\phi(t) = \hat{\phi}(t_i) + \int_{t_i}^{t_i+t} 2\pi \hat{f}(u) du$$

Όπου $\hat{\phi}(t_i)$, η εκτίμηση της στιγμιαίας φάσης στη χρονική στιγμή t_i . Όμως, η λύση αυτή δεν λαμβάνει υπόψη την κατάσταση στο όριο της χρονικής στιγμής t_{i+1} . Η τιμή της φάσης στο χρονικό σημείο t_{i+1} δίδεται από τη σχέση:

$$\phi(t) \Big|_{t=t_{i+1}} = \hat{\phi}(t_{i+1}) + 2\pi M$$

Με M έναν ακέραιο κατάλληλα επιλεγμένο ώστε να είναι όσο το δυνατόν κοντύτερα στην τιμή:

$$M = \text{round}\left(\frac{\phi(t_{i+1}) - \hat{\phi}(t_i)}{2\pi}\right)$$

με round να είναι η διαδικασία που επιστρέφει τον κοντινότερο ακέραιο στην τιμή που προκύπτει από την τιμή του ορίσμά της.

Στη συνέχεια, αποσκοπώντας στην εξασφάλιση της συνέχειας της φάσης στα όρια της, προτάθηκε η αλλαγή του υπολογισμού της φάσης ως εξής:

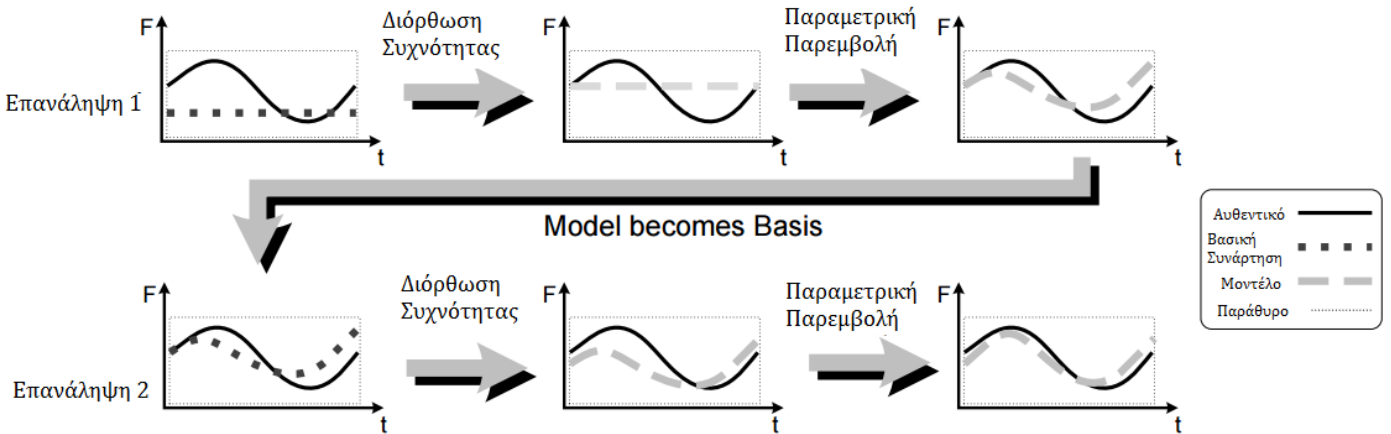
$$\phi(t) = \hat{\phi}(t_i) + \int_{t_i}^{t+t_i} (2\pi f(u) + c(u)) du$$

με το $c(u)$ να δίδεται από την σχέση

$$c(u) = r(t_{i+1}) \sin\left(\frac{\pi(u - t_i)}{t_{i+1} - t_i}\right)$$

Σε αντίθεση με το QHM, όπου οι τιμές της βασικής συνάρτησης ήταν παραμετρικές και στατικές, στο aQHM δεν είναι ούτε παραμετρικές, ούτε απαραίτητα στατικές.

Πιο συγκεκριμένα, στο aQHM, οι συναρτήσεις βάσης χρησιμοποιούν τις στιγμιαίες φάσεις που έχουν εκτιμηθεί αναλύοντας το σήμα εισόδου. Με αυτό τον τρόπο, δημιουργείται και η έννοια της προσαρμοστικότητας καθώς οι στιγμιαίες φάσεις είναι εξίσου προσαρμοστικές, στις εκτιμήσεις των χαρακτηριστικών της παρούσας φάσης του σήματος. Η διαδικασία των επιτυχών προσαρμοσμένων φαίνεται στο σχήμα 2.3.1. Ωστόσο, σε αυτό το μοντέλο επιτρέπεται μόνο η προσαρμογή της φάσης.



Σχήμα 2.3.1

Κατασκευή της προσαρμοσμένης συχνότητας εφαρμόζοντας παράθυρο ανάλυσης με δύο επαναλήψεις στο πεδίο του χρόνου, με εφαρμογή του μοντέλου aQHM.

2.4 Extended adaptive Quasi-Harmonic Model, eaQHM

Ωστόσο, το aQHM είναι ένα μοντέλο το οποίο δουλεύει προσαρμοστικά μόνο για τη φάση του σήματος και η αδυναμία αυτή χρήζει βελτίωσης;. Στη νέα του και τελευταία χρονικά μορφή [KPRS12], παρουσιάστηκε το επεκτάσιμο προσαρμοστικό σχεδόν Αρμονικό Μοντέλο (extended-adaptive Quasi Harmonic Model-eaQHM), στο οποίο εκτός απο την προσαρμογή της φάσης του σήματος, έχει προστεθεί και η προσαρμογή του πλάτους , ένα γεγονός που προσφέρει καλύτερο αποτέλεσμα (κοντινότερο στο αυθεντικό) στο τελικώς παραγόμενο σήμα.

Το eaQHM [KPRS12] δίνεται από την σχέση :

$$x(t) = \left(\sum_{k=-K}^K (a_k + tb_k) \hat{\alpha}_k(t) e^{j(\hat{\phi}_k(t+t_i) - \hat{\phi}_k(t_i))} \right) w(t), \quad t \in [-T, T]$$

$$\text{με} \quad \hat{\alpha}_k(t) = \frac{A_k(t + t_i)}{A_k(t_i)}$$

όπου το t_i παραμένει το κεντρο του παραθύρου ανάλυσης και $A_k(t)$ είναι το στιγμιαίο πλάτος της k -οστής συνιστώσας το οποίο είναι είτε ένα αρχικά εκτιμημένο στιγμιαίο πλάτος, είτε έχει παραχθεί απο προηγούμενο στάδιο προσαρμοστικότητας. Η αρχική εκτίμηση των παραμέτρων γίνεται όπως και στο QHM :

$$\begin{bmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{bmatrix} = (\mathbf{E}_e^H \mathbf{W}^H \mathbf{W} \mathbf{E}_e)^{-1} \mathbf{E}_e^H \mathbf{W}^H \mathbf{W} \mathbf{s}$$

Όπου \mathbf{W} και \mathbf{s} παραμένουν όπως στο QHM, ο πίνακας που περιέχει τις τιμές του παραθύρου στη διαγώνιό του και το διάνυσμα του εισαγόμενου σήματος, αντίστοιχα. Ο πίνακας \mathbf{E} μπορεί να αναπαρασταθεί απο την σχέση $\mathbf{E}_e = [\mathbf{E}_{e0} | \mathbf{E}_{e1}]$, όπου (εδώ έχουν προστεθεί πλάτη) οι υποπίνακες $\mathbf{E}_i, i=0,1$, έχουν στοιχεία που δίνονται απο τις σχέσεις:

$$(\mathbf{E}_{e0})_{n,k} = \alpha_k(t_n) e^{j(\phi_k(t_n+t_i) - \phi_k(t_i))}$$

$$(\mathbf{E}_{e1})_{n,k} = t_n \alpha_k(t_n) e^{j(\phi_k(t_n+t_i) - \phi_k(t_i))} = t_n (\mathbf{E}_{e0})_{n,k},$$

Όπως γίνεται εύκολα αντιληπτό, στις βασικές συναρτήσεις έχουν προσαρμοστεί και τα τοπικά πλάτη του σήματος. Σημαντικό στοιχείο είναι επίσης ότι το πλάτος $A_k(t)$, πριν να χρησιμοποιηθεί στις συναρτήσεις βάσης, έχει διαιρεθεί με το $A_k(t_i)$ ούτως ώστε να έχει μοναδιαία τιμή στο κέντρο του παραθύρου ανάλυσης.

Όπως ακριβώς το aQHM, έτσι και το eaQHM, χρειάζεται ένα στάδιο αρχικοποίησης. Για αυτό το σκοπό χρησιμοποιήθηκε [PRS08] το QHM, ωστόσο οποιοδήποτε μοντεό ανάλυσης μπορεί να χρησιμοποιηθεί.

Έτσι, το στάδιο της αρχικοποίησης χρησιμοποιεί τις συναρτήσεις:

$$\hat{f}_k^0(t_i) = \hat{f}_k^0(t_{i-1}) + \hat{\eta}_k$$

$$\hat{A}_k^0(t_i) = |a_k^i|, \quad \hat{\phi}_k^0(t_i) = \angle a_k^i$$

όπου το t_i είναι το κεντρο του i -οστού πλαισίου ανάλυσης.

2.4.1 Αλγόριθμος eaQHM

Η ανασύνθεση του σήματος με ημίτονα διαμορφωμένα κατά πλάτη -συχνότητες (Amplitude Modulation-Frequency Modulation , AM-FM) γίνεται με τον παρακάτω αλγόριθμο:

Αλγόριθμος extended-adapted Quasi Harmonic Model

1. Αρχικοποίηση:

Παροχή της εκτίμησης συχνότητας $f_k^0(t_1)$

ΓΙΑ τα πλαίσια $i=1,2,4, \dots, L$

- (α) Υπολόγισε τα a_{ki} , b_{ki} χρησιμοποιώντας LS
- (β) Εκτίμηση της συχνότητας και του πλάτους :

$$\hat{f}_k^0(t_i) = \hat{f}_k^0(t_{i-1}) + \hat{\eta}_k$$

$$\hat{A}_k^0(t_i) = |a_k^i|, \quad \hat{\phi}_k^0(t_i) = \angle a_k^i$$

(γ) $f_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$

ΤΕΛΟΣ_ΕΠΑΝΑΛΗΨΗΣ

Παρεμβολή των παραμέτρων : $\{\hat{A}_k^0(t), \hat{f}_k^0(t), \hat{\phi}_k^0(t)\}$

2. Προσαρμοστικότητα σε Πλάτη και Φάσεις:

ΓΙΑ την προσαρμογή $i=1,2,\dots$

ΓΙΑ τα πλαίσια $l=1,2,4, \dots, L$

- (α) Υπολογισμός a_{ki} , b_{ki} χρησιμοποιώντας $\hat{\phi}_k^{i-1}(t)$ και τη σχέση

$$x(t) = \left(\sum_{k=-K}^K (a_k + tb_k) \hat{\alpha}_k(t) e^{j(\hat{\phi}_k(t+t_i) - \hat{\phi}_k(t_i))} \right) w(t), \quad t \in [-T, T]$$

- (β) Ανανέωση της συχνότητας και του πλάτους:

$$\hat{f}_k^0(t_i) = \hat{f}_k^0(t_{i-1}) + \hat{\eta}_k$$

$$\hat{A}_k^0(t_i) = |a_k^i|, \quad \hat{\phi}_k^0(t_i) = \angle a_k^i$$

ΤΕΛΟΣ_ΕΠΑΝΑΛΗΨΗΣ

ΤΕΛΟΣ_ΕΠΑΝΑΛΗΨΗΣ

Παρεμβολή των παραμέτρων: $\{\hat{A}_k^i(t), \hat{f}_k^i(t), \hat{\phi}_k^i(t)\}$

Το κριτήριο σύγκλισης μεταξύ των μοντέλων aQHM και eaQHM επιλέχθηκε να δίδεται από την ανισότητα:

$$\frac{SRRER^{i-1} - SRRER^i}{SRRER^{i-1}} < \epsilon$$

με $SRRER_i$ να ορίζεται ως ο λόγος του αυθεντικού σήματος προς το σφάλμα ανακατασκευής (Signal-to-Reconstruction-Error Ratio):

$$SRRER = 20 \log_{10} \frac{\sigma_{x(t)}}{\sigma_{x(t) - \hat{x}(t)}}$$

στην i -οστή προσαρμογή όπου το σ_x δηλώνει την τυπική απόκλιση του $x(t)$, $\hat{x}(t)$ είναι το αυθεντικό σήμα και $\hat{x}(t)$ το ανακατασκευασμένο σήμα στην i -οστή προσαρμογή και τέλος, ϵ είναι ένα όριο όπου για τα παραδείγματά μας θα έχει την τιμή 0.02. Σαν τελευταίο βήμα αυτού του αλγορίθμου, το σήμα μπορεί να αναπαρασταθεί τελικά ως το άθροισμα των AM-FM συνιστωσών του:

$$\hat{x}(t) = \sum_{k=1}^K \hat{A}_k(t) e^{j\hat{\phi}_k(t)}$$

Συμπερασματικά, στο [KPRS12], γίνεται σαφής η υπεροχή του eaQHM σε σχέση με το aQHM σε συνθετικά σήματα καθώς επίσης και σε έμφωνα σήματα φωνής, χρησιμοποιώντας σύγκριση των SRRER αλλά και του Mean Absolute Error (MAE) το οποίο δίνεται από τον τύπο:

$$MAE\{\hat{\theta}\} = \frac{1}{M} \sum_{i=1}^M |\hat{\theta}^{(i)} - \theta|$$

Όπου $\theta(i)$, η εκτιμώμενη παράμετρος στην i -οστή προσομοίωση (πλάτος, συχνότητα), και M ο αριθμός των προσομοιώσεων Monte Carlo.

2.5 Πλήρους εύρους (full-band) eaQHM

Με αφετηρία την ύπαρξη του καινοτόμου και εξελιγμένου μοντέλου eaQHM, και την ένδειξη βελτιωμένων αποτελεσμάτων στην ανάλυση και στην ανασύνθεση σημάτων, δημιουργήθηκε η ανάγκη για εφαρμογή του μοντέλου σε πραγματικά σήματα φωνής [KPOS12]. Ωστόσο, θα πρέπει να δωθεί ιδιαίτερη προσοχή στις λεπτομέρειες της υλοποίησης του συστήματος. Η σταθερότητα, η ευρωστία και οι δυνατότητες του μοντέλου πρέπει να είναι εξαιρετικές όχι μόνο στην ανάλυση και σύνθεση του σήματος, αλλά και σε πιθανές τροποποιήσεις και ανωμαλίες που δεν μπορούν να προβλεφθούν σε ένα πραγματικό σήμα, όπως απότομες αυξομειώσεις στην ένταση και “τρέμουλο” φωνής. Στή συνέχεια θα παρουσιαστούν οι λεπτομέρειες για τη λειτουργία του full-band eaQHM.

Η βασική ιδέα για το full-band eaQHM είναι αρχικά μια καθαρά αρμονική προσέγγιση του σήματος, όπου μετά από προσεγγίσεις και επαναλαμβανόμενες διορθώσεις του πλάτους και

της συχνότητας μέσω της διαδικασίας της προσαρμογής, τελικά επιτυγχάνει να συγκλίνει σε μία σχεδόν αρμονική αναπαράσταση. Το full-band σήμα αναπαρίσταται σαν μία AM-FM αποσύνθεση που δίδεται από τη σχέση:

$$d(t) = \sum_{k=-K}^K A_k(t) e^{j\phi_k(t)}$$

Με $A_k(t)$ το στιγμιαίο πλάτος και $\phi(t)$ η στιγμιαία φάση της k -οστής συνιστώσας, διαδοχικά. Η στιγμιαία φάση αναπαρίσταται από τη σχέση:

$$\phi_k(t) = \phi_k(t_i) + \int_{t_i}^t 2\pi f_k(u) du$$

με $\phi_k(t_i)$ την τιμή της στιγμιαίας φάσης την χρονική στιγμή t_i και $f_k(t)$ η στιγμιαία συχνότητα στην k -οστή συνιστώσα.

2.5.1 Ανάλυση

Στο πρώτο στάδιο, υπολογίζεται μία αρχική και συνεχής εκτίμηση της θεμελιώδους συχνότητας f_0 για κάθε ένα από τα πλαίσια του σήματος. Αν και δεν υπάρχει f_0 σε άφωνα (unvoiced) κομμάτια, μπορεί να χρησιμοποιηθεί για την αρχικοποίηση της, μια «άτακτη» εκτίμηση. Στη συνέχεια, αποσκοπώντας στην απόκτηση μίας αρχικής εκτίμησης και για τα στιγμιαία πλάτη όλων των αρμονικών, χρησιμοποιείται μια full-band αρμονικότητα. Εφαρμόζοντας ένα παράθυρο Blackman $w(t)$ μεγέθους $2T$ στα όρια $[t_i - T, T + t_i]$, δημιουργείται ένα τμήμα σήματος φωνής το οποίο μοντελοποιείται ως:

$$d(t) = \left(\sum_{k=-L}^L a_k e^{j2\pi \hat{f}_k t} \right) w(t)$$

Όπου a_k είναι το μιγαδικό πλάτος της k -οστής αρμονικής, $\hat{f}_k = k\hat{f}_0$ αναπαριστούν τις συχνότητες ανάλυσης, και L ο αριθμός των αρμονικών που υπάρχουν στο φάσμα της συχνότητας Nyquist [1]. Η εκτίμηση των παραμέτρων παρέχεται με μία μέθοδο ελάχιστων τετραγώνων (LS) [PTRS10]. Το iQHM δεν εφαρμόστηκε στο σύστημα του eaQHM καθώς αυξάνει την πολυπλοκότητα χωρίς να αυξάνει ιδιαίτερα την ακρίβεια υπολογισμού παραμέτρων. Στο τελικό στάδιο της ανάλυσης, το συνολικό σήμα μπορεί να συνθετηθεί παρεμβάλλοντας τις τιμές των $|a_k|$ και \hat{f}_k πάνω σε κατάλληλα στιγμιότυπα χρόνου ανάλυσης t_i :

$$\hat{d}(t) = \sum_{k=-L}^L \hat{A}_k(t) e^{j\hat{\phi}_k(t)}$$

με

$$\hat{A}_k(t) = |a_k(t)|$$

$$\hat{\phi}_k(t_i) = \angle a_k(t_i)$$

και

$$\hat{\phi}_k(t) = \hat{\phi}_k(t_i) + \int_{t_i}^t (2\pi k \hat{f}_0(u) + c(u)) du$$

2.5.2 Προσαρμοστικότητα

Το μοντέλο που παρουσιάστηκε, είναι αρμονικό και στατικό μέσα σε ένα πλαίσιο ανάλυσης. Με σκοπό την σύμπτυξη της σχεδόν-αρμονικότητας στο σύστημα και την απαλειφή της στατικότητας, προτάθηκε [KPOS12] η προβολή του σήματος σε ένα σύνολο απο προσαρμοστικές συναρτήσεις βάσης. Έτσι παράχθηκε το eaQHM με μορφή:

$$d(t) = \left(\sum_{k=-L}^L (a_k + tb_k) (\hat{A}_k(t) e^{j\hat{\phi}_k(t)}) \right) w(t)$$

όπου

$$\hat{A}_k(t) = \frac{\hat{A}_k(t + t_i)}{\hat{A}_k(t_i)}$$

Σε αυτό το μοντέλο, a_k , b_k είναι το μιγαδικό πλάτος και η μιγαδική κλίση της k -οστής συνιστώσας,

ενώ $\hat{A}_k(t)$, $\hat{f}_k(t)$, $\hat{\phi}_k(t)$ οι εκτιμήσεις του στιγμιαίου πλάτους, συχνότητας και φάσης της k -οστής συνιστώσα, διαδοχικά, απο το προηγούμενο βήμα ανάλυσης. Οι παράμετροι a_k , b_k υπολογίζονται από τη μέθοδο LS. Είναι σημαντικό να γίνει σαφές ότι οι συναρτήσεις βάσης στις οποίες το σήμα προβάλεται είναι χρονικά μεταβαλλόμενες. Η προσαρμοστικότητα ολοκληρώνεται χρησιμοποιώντας τον μηχανισμό διόρθωσης συχνότητας [PRS08], και εκθέτει μια εκτίμηση της διαφοράς ανάμεσα στην πραγματική και την υπολογισμένη συχνότητα στην k -οστή συνιστώσα, ορισμένη ως $\eta_k = f_k - \hat{f}_k$, και δίδεται απο τη σχέση:

$$\hat{\eta}_k = \frac{1}{2\pi} \frac{\Re\{a_k\}\Im\{b_k\} - \Im\{a_k\}\Re\{b_k\}}{|a_k|^2}$$

Έτσι, κατα την πρώτη προσαρμογή, για τη χρονική στιγμή ανάλυσης t_i , οι στιγμιαίες συχνότητες είναι:

$$\hat{f}_k(t_i) = k\hat{f}_0(t_i) + \hat{\eta}_k(t_i)$$

και οι στιγμιαίες φάσεις αναπαρίστανται ως:

$$\hat{\phi}_k(t) = \hat{\phi}_k(t_i) + \int_{t_i}^t (2\pi\hat{f}_k(u) + c(u))du$$

Με το $c(u)$ να αναπαριστάται όπως στο Adaptive Quasi Harmonic Model(2.3).

Στη συνέχεια, μια LS λύση για τα α_k, b_k (χρησιμοποιεί τις βελτιωμένες τιμές των συχνοτήτων και των φάσεων) οδηγεί σε μία καλύτερη εκτίμηση των στιγμιαίων πλατών και των όρων $\hat{\eta}_k$. Προσθέτοντας επαναλαμβανόμενα τον όρο $\hat{\eta}_k$ της παρούσας προσαρμογής, στην k -οστή συχνότητα της πηγούμενης προσαρμογής, τα κομμάτια συχνότητας παρεκκλίνουν από την αυστηρή αρμονικότητα και απεικονίζουν τις συχνότητες καλύτερα.

Όλα τα παραπάνω δεδομένα έχουν ως αποτέλεσμα το σύστημα αυτό να παρέχει όλο και περισσότερο αξιόπιστα δεδομένα μετά απο κάθε προσαρμογή και οι εκτιμώμενες συχνότητες να πλησιάζουν τις τιμές των συχνοτήτων του αυθεντικού σήματος. Η διόρθωση συχνότητας για την k -οστή συνιστώσα, δίδεται απο τη σχέση:

$$|\hat{\eta}_k(t_i)| \leq \frac{\hat{f}_0(t_i)}{m+1}$$

με το $m \in \{1, \dots, M\}$ να είναι ο αριθμός της παρούσας προσαρμογής. Επίσης, M είναι ο μέγιστος αριθμός προσαρμογών. Χρειάζεται ο ορισμός ενός τέτοιου αριθμού καθώς μετά απο πεπερασμένο αριθμό προσαρμογών υπάρχει πιθανότητα να χαθεί η αξιοπιστία των αποτελεσμάτων.

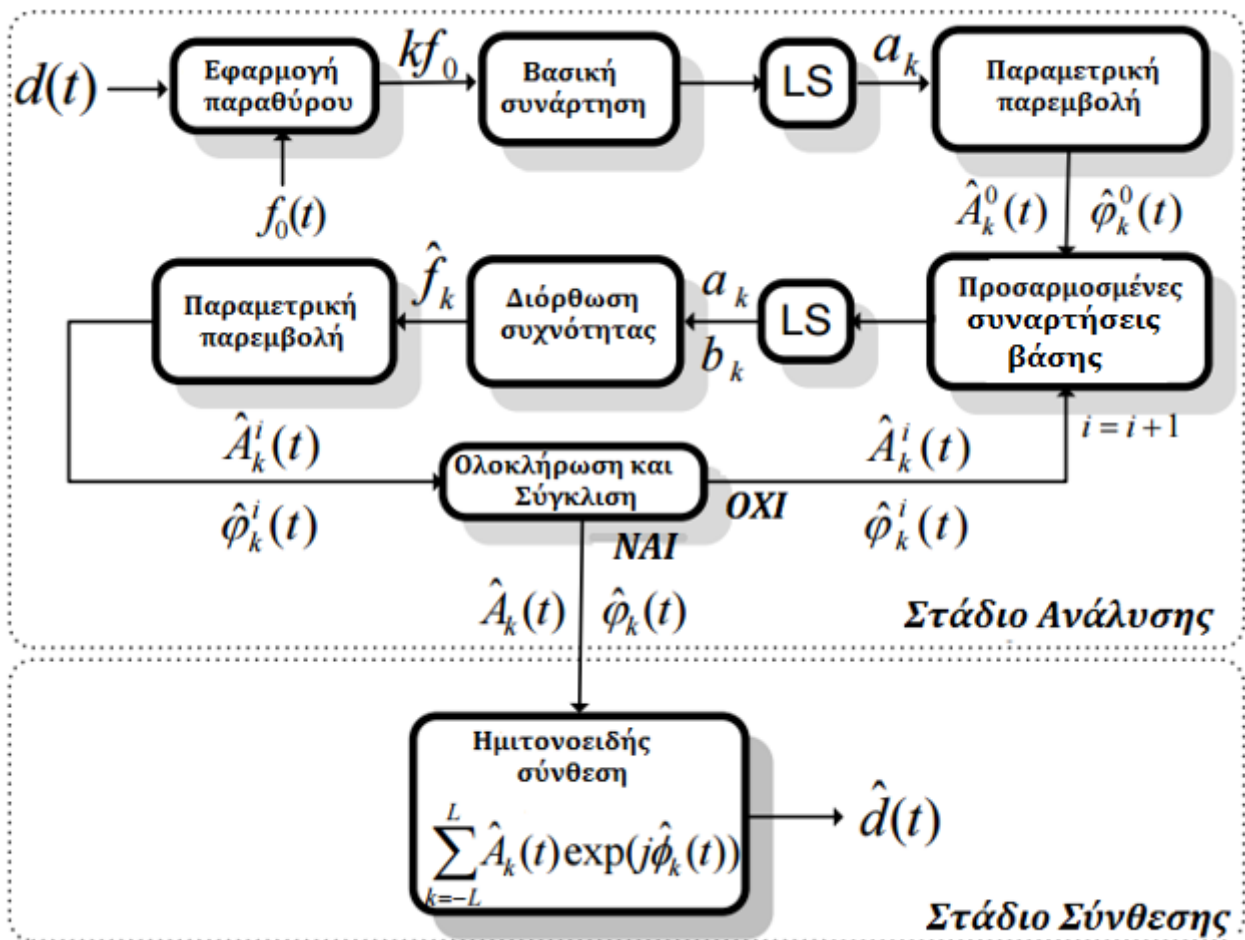
2.5.3 Σύνθεση

Στο στάδιο της σύνθεσης, το k -οστό στιγμιαίο πλάτος υπολογίζεται είτε ως γραμμική είτε ως πολυωνυμική παρεμβολή, έχοντας ως δεδομένα επιτυχείς εκτιμήσεις της τελευταίας προσαρμογής. Η k -οστή στιγμιαία συχνότητα, $f_k(t)$, υπολογίζεται επίσης με πολυωνυμική παρεμβολή. Ένα εξίσου σημαντικό χαρακτηριστικό του συστήματος ανάλυσης, είναι ότι ο μηχανισμός αντιπαραβολής συχνοτήτων είναι ανούσιος καθώς οι αναλυόμενες συχνότητες είναι ακέραια πολλαπλάσια μιας θεμελιώδους συχνότητας και ο αριθμός των συνιστωσών είναι σχεδόν σταθερός.

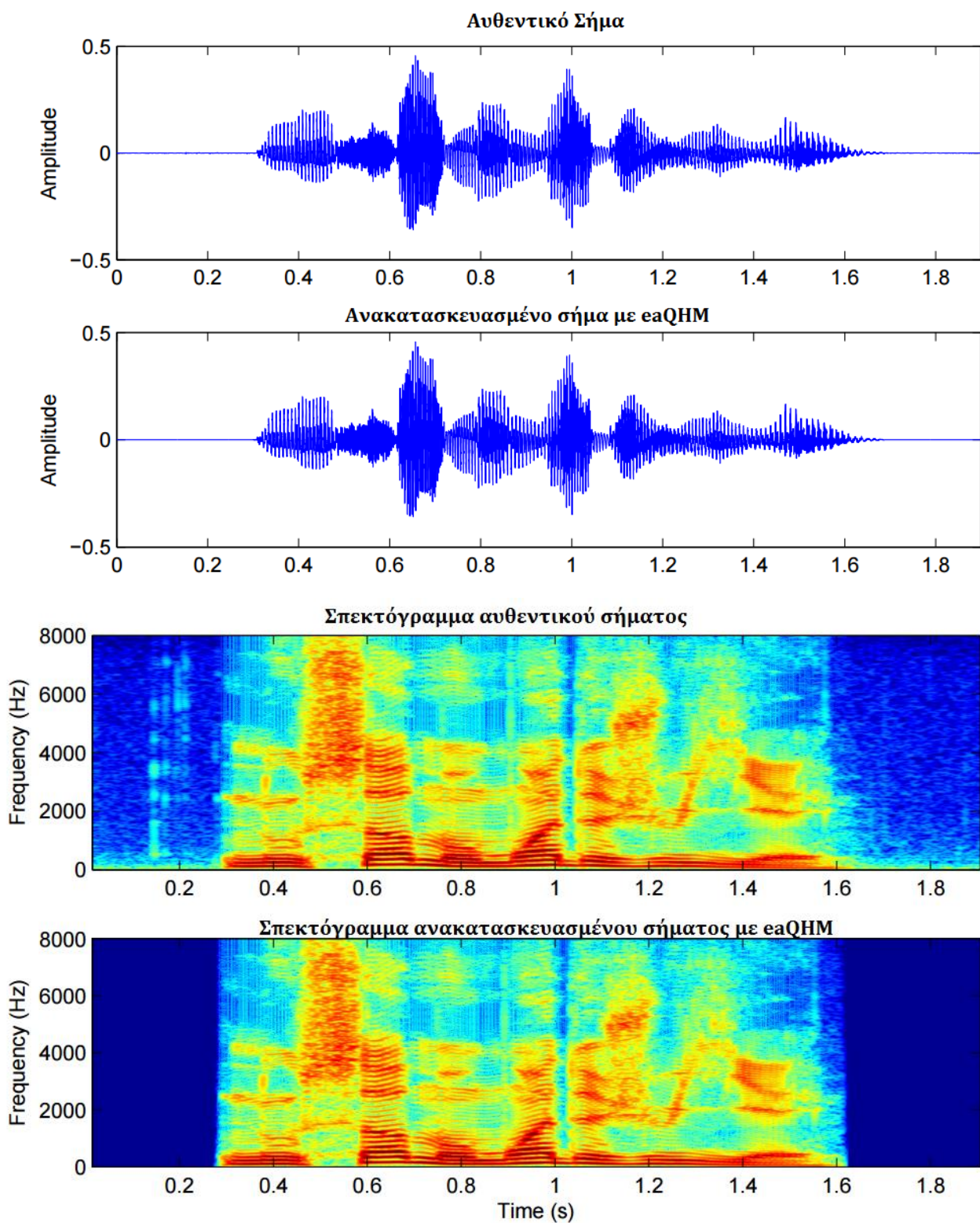
Όσον αφορά την k -οστή στιγμιαία φάση, χρησιμοποιείται η μη παραμετρική προσέγγιση, βασισμένη στο ολοκλήρωμα της στιγμιαίας συχνότητας. Τελικά, το σήμα φωνής μπορεί να υπολογισθεί απο τις χρονικά μεταβαλλόμενες συνιστώσες του απο τη σχέση:

$$\hat{d}(t) = \sum_{k=-L}^L \hat{A}_k(t) e^{j\hat{\phi}_k(t)}$$

Παρακάτω παρατίθεται ένα σχήμα για τον τρόπο που λειτουργεί το full-band σύστημα του eaQHM.



2.5.4 Παράδειγμα ανακατασκευής σήματος



Σχήμα 2.5.41

Η αναπαράσταση ενός σήματος ανδρικής φωνής στην αρχική του μορφή και στην ανακατασκευασμένη του μορφή, χρησιμοποιώντας το μοντέλο eaQHM.

3 Πειραματικό Μέρος

3.1 Συνθετικά πειράματα

3.1.1 Εισαγωγή

Στο πρώτο στάδιο αυτής της εργασίας, και αφού αποσαφηνίστηκαν οι λειτουργίες των σχεδόν-αρμονικών μοντέλων, δημιουργήθηκε η ανάγκη να εξετάσουμε την απόδοση κάθε πλαισίου ανάλυσης ξεχωριστά, με στόχο την βελτιστοποίηση, την σταθεροποίηση και την δυναμικότητα των μοντέλων σε διαφοροποιημένα σήματα ομιλίας. Η ανακατασκευή του σήματος μέσα σε ένα παράθυρο ανάλυσης έχει μεγάλη σχέση με το είδος και το μέγεθος του εφαρμοζόμενου παραθύρου, $w(t)$, σε αυτό το τμήμα σήματος.

Με αφετηρία λοιπόν το [PRS08], χρησιμοποιήσαμε το QHM για τις ανάγκες των πειραμάτων μας. Κατά την μελέτη επιλογής ενός παραθύρου το οποίο θα έδινε τα βέλτιστα δυνατά απο-τελέσματα σε σήματα φωνής βασικότατο ρόλο διαδραμάτισαν έννοιες όπως το Bias, Mean Square Error(MSE), Signal to Noise Ratio(SNR) και Bandwidth. Στη συνέχεια αυτής της ενότητας θα παρουσιαστούν αυτές οι έννοιες, καθώς και οι διάφοροι τύποι παραθύρων που χρησιμοποιήθηκαν στην διάρκεια της παρούσας εργασίας. Τέλος, θα διευκρινιστούν μέσω συγκριτικών παραδειγμάτων, τα είδη των παραθύρων που είναι αποδοτικότερα και σταθερότερα κατά την ανάλυση σημάτων με QHM.

3.1.2 Παράθυρα Ανάλυσης

Όπως σε κάθε πρόβλημα εκτίμησης συχνότητας, το παράθυρο ανάλυσης πρέπει να είναι αρκετά μεγάλο ώστε να έχουμε υψηλή ποιότητα στον υπολογισμό των παραμέτρων. Ωστόσο το QHM είναι ένα στατικό μοντέλο και η ανάλυση σημάτων φωνής χρήζει το μέγεθος του παραθύρου να είναι τόσο μικρό ώστε να ικανοποιήσει τα μη-στατικά χαρακτηριστικά του φυσικού σήματος φωνής.

Τύπος παραθύρου	Συνάρτηση παραθύρου
Rectangular	$w(n) = 1$
Bartlett	$w(n) = 1 - \frac{2 n - \frac{M}{2} }{M}$
Hanning	$w(n) = 0.5 - 0.5 \cos\left(\frac{2\pi n}{M}\right)$
Hamming	$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{M}\right)$
Blackman	$w(n) = 0.42 - 0.5 \cos\left(\frac{2\pi n}{M}\right) + 0.08 \cos\left(\frac{4\pi n}{M}\right)$

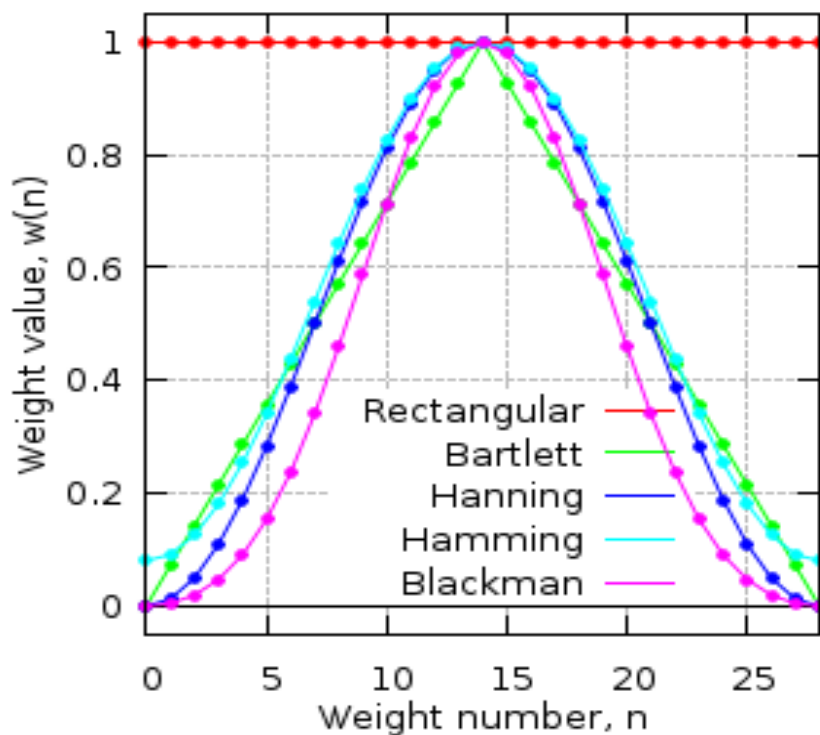
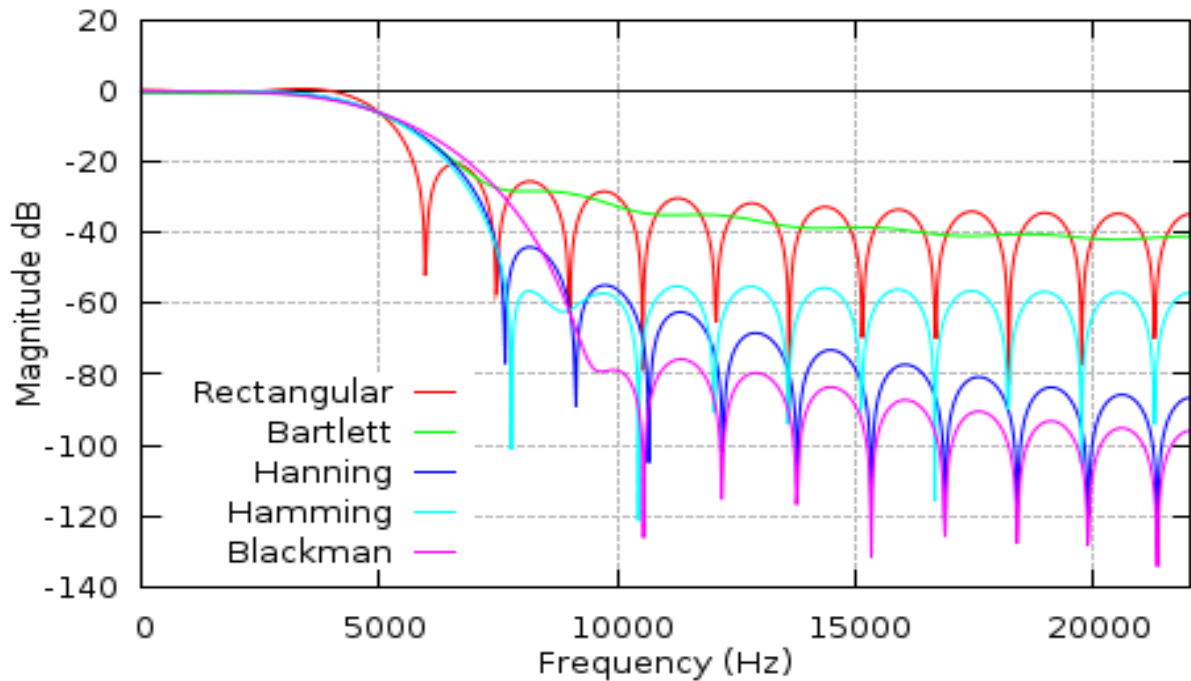
Πίνακας 3.1.21 [EIK1]

Τύποι παραθύρου και οι συναρτήσεις τους.

Ανάμεσα σε ένα μεγάλο εύρος παραθύρων, στο [KPRS12] επιλέγεται το παράθυρο hamming για να εφαρμοστεί στην ανάλυση του σήματος, καθώς έχει υψηλό bandwidth και συνεπώς

δίνει αρκετά χαμηλή τιμή στο $W''(f-f_k)$. Ως bandwidth ορίζεται η απόσταση (σε Hz) μεταξύ των δύο πρώτων μηδενισμών (εκατέρωθεν του $f=0$) στο φασμα πλάτους του παραθύρου [2]

Στον πίνακα 3.1.21 και στα γραφήματα 3.1.22_(α,β) παρουσιάζονται κάποια από τα παράθυρα ανάλυσης που χρησιμοποιήθηκαν σε αυτό το πείραμα, σε αναζήτηση βέλτιστων αποτελεσμάτων.



Σχήμα 3.1.22_(α,β) [EIK2]

Πάνω: Η συχνότητα απόκρισης πέντε διαφορετικών παραθύρων ανάλυσης. Κάτω: Πέντε παράθυρα ανάλυσης όπως απεικονίζονται από τις συναρτήσεις τους και εφαρμόζονται στα σήματα φωνής.

3.1.3 Bias

Το πρώτο μέτρο σύγκρισης που χρησιμοποιήθηκε για την εύρεση του καταλληλότερου παραθύρου ανάλυσης είναι το συστηματικό σφάλμα (bias). Το bias μετρά την αντοχή του κάθε παραθύρου στην διόρθωση συχνοτήτων. Πιο συγκεκριμένα, δείχνει το εύρος των συχνοτήτων που μπορεί να διορθώσει το QHM καθώς σε όποιες τιμές το bias έχει μηδενική τιμή, σημαίνει ότι το μοντέλο παρέχει την βέλτιστη διόρθωση στην συχνότητα αυτή.

Με σκοπό τον υπολογισμό του bias για εκτίμηση συχνότητας σε QHM χρησιμοποιήθηκε στα πειράματα ένα σήμα με μία συνιστώσα:

$$x(t) = \alpha e^{j(2\pi ft + \eta t)}$$

Καθώς και το αντίστοιχο QHM μοντέλο:

$$s(t) = (a + tb)e^{j(2\pi ft)} \quad -T \leq t \leq T$$

Στο σημείο αυτό, πραγματοποιήθηκαν εναλλαγές στον τύπο του εφαρμοζόμενου παραθύρου, $w(t)$, με τα παράθυρα που επιλέχθηκαν να είναι τα : Bartlett, Blackman, Flattop, Gauss, Hamming, Hanning, Kaiser, Rectangular, Taylor, Triangle, Tukey.

Για τις ανάγκες του πειράματος κρατήθηκε σταθερό μέγεθος για όλους του τύπους παραθύρων μεγέθους $2T=16\text{ms}$. Έτσι με την εφαρμογή του παραθύρου στο σήμα προκύπτει η LS λύση με :

$$a = \alpha \frac{\sin(\eta T)}{\eta T}$$
$$b = \alpha 3j \left(\frac{\sin(\eta T)}{\eta^2 T^3} - \frac{\cos(\eta T)}{\eta T^2} \right)$$

και ο συντελεστής ρ_2 από το QHM (2.2) γίνεται:

$$\rho_2 = 3 \left(\frac{1}{\eta T^2} - \frac{\cot(\eta T)}{T} \right)$$

ή με άλλα λόγια από το QHM προκύπτει:

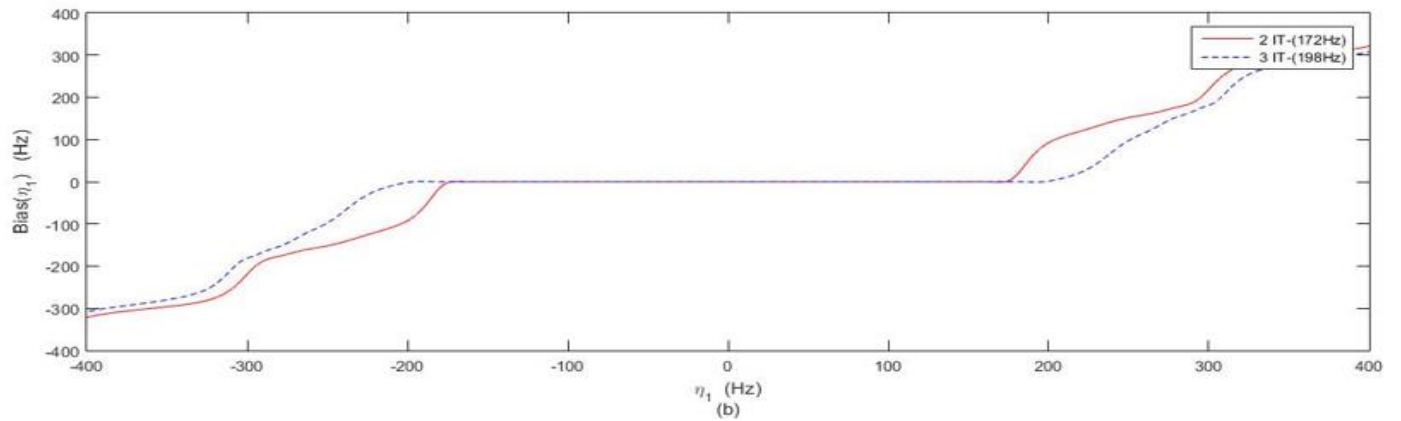
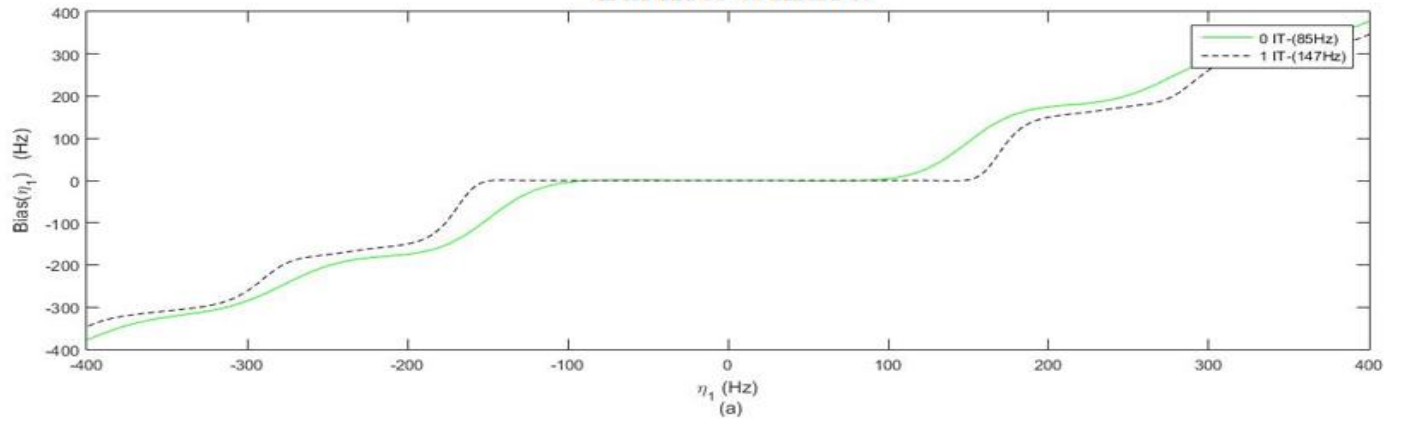
$$\hat{\eta} = 3 \left(\frac{1}{\eta T^2} - \frac{\cot(\eta T)}{T} \right)$$

Τέλος ο υπολογισμός του bias της εκτίμησης αυτής, δίδεται από τον τύπο:

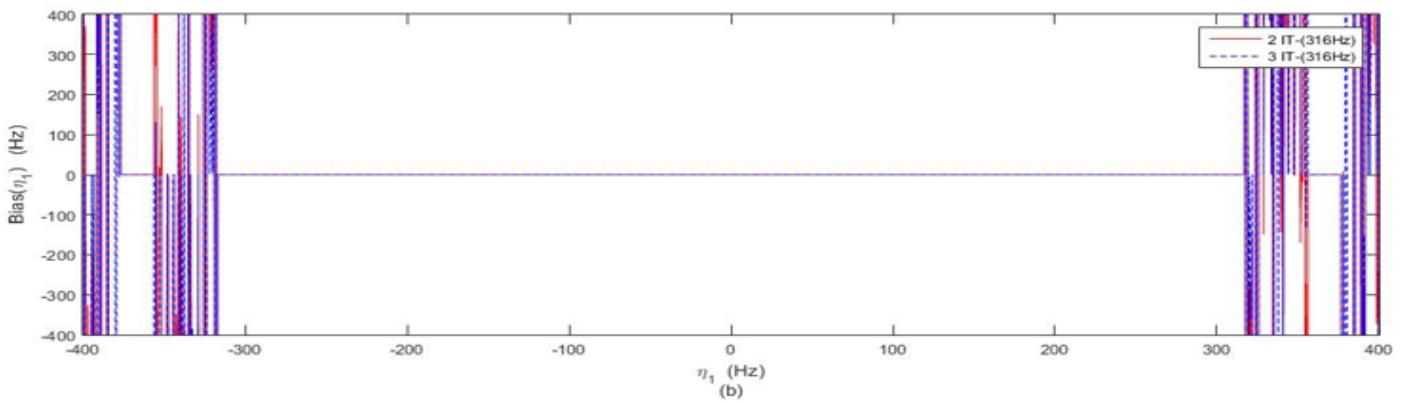
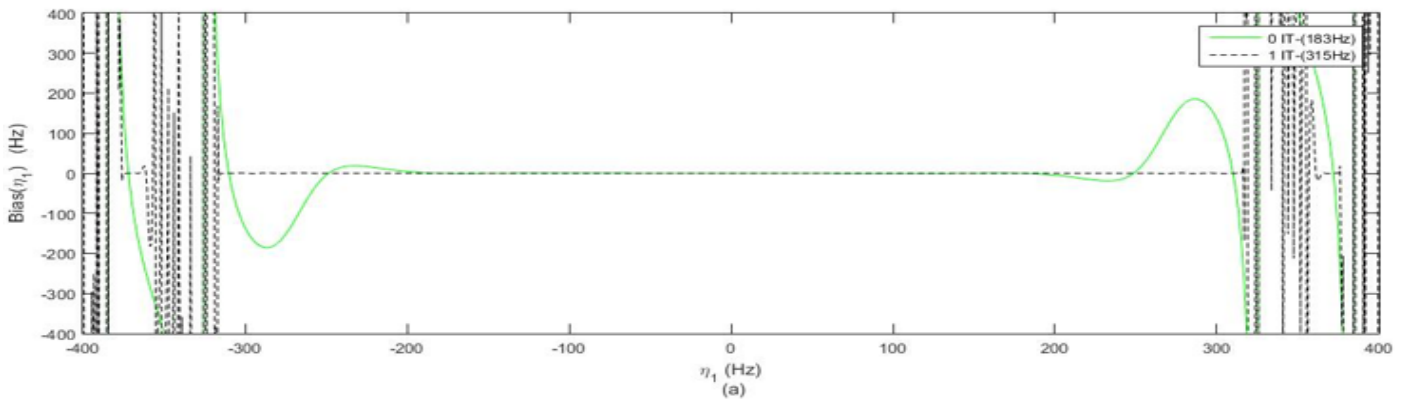
$$\text{Bias}(\eta) = \eta - \hat{\eta}$$

Τον τελικό παράγοντα στην εκτέλεση του συγκεκριμένου πειράματος έπαιξε η εφαρμογή μηχανισμού διόρθωσης συχνότητας.

Bartlett Window



Gauss Window



Γράφημα 3.1.31_(α,β)
 Το bias του παραθύρου Gauss(πάνω) και του παραθύρου Bartlett (κάτω)
 για επαναληπτικές διορθώσεις(iterations) $-IT=\{0,1,2,3\}$

Στο πείραμα της εργασίας αυτής υπολογίσθηκε το bias του κάθε παραθύρου συνολικά τέσσερις φορές. Αρχικά χωρίς διόρθωση, ενώ στη συνέχεια προστέθηκε μία επαναληπτική διόρθωση, έπειτα δύο και τελικά τρεις.

Στα γραφήματα 3.1.31_(α,β) φαίνεται το bias και το εύρος στο οποίο αυτό μηδενίζεται. Δίδονται ενδεικτικά τα γραφήματα για τα παράθυρα Gauss και Bartlett ενώ συγκεντρωτικά γραφήματα για όλα τα παράθυρα δίδονται στο Παράρτημα Α.

Αναλυτικότερα, οι τιμές των συχνοτήτων που μπορούν να διορθωθούν από το κάθε παράθυρο για επαναληπτικές διορθώσεις (IT={0,1,2,3}) δίδεται από τον συγκεντρωτικό πίνακα 3.1.33. Μελετώντας τα αριθμητικά αποτελέσματα από την εφαρμογή διαφόρων παραθύρων, παρατηρείται ιδιαίτερα υψηλή διόρθωση της συχνότητας με χρήση των παραθύρων Flattop και Gauss. Στη συνέχεια, ακολουθεί το παράθυρο Blackman το οποίο παρουσιάζει και την μεγαλύτερη βελτίωση καθώς μετά από 3 επαναληπτικές διορθώσεις το εύρος του μηδενικού bias έχει αυξηθεί κατά 380%.

WINDOW	NO ITERATION	ITERATION 1	ITERATION 2	ITERATION 3
Bartlett	±85	±147	±172	±198
Blackman	±75	±234	±275	±287
Flattop	±200	±355	±390	±435
Gauss	±183	±315	±316	±316
Hamming	±38	±133	±160	±166
Hanning	±34	±120	±155	±160
Kaiser	±14	±38	±47	±49
Rectangular	±14	±38	±46	±48
Taylor	±30	±154	±165	±166
Triangle	±85	±146	±172	±197
Tukey	±15	±54	±68	±71

Πίνακας 3.1.33

Το εύρος της συχνότητας σε Hertz για κάθε παράθυρο, όπου το bias είναι μηδενικό με iterations={0,1,2,3}

3.1.4 Mean Square Error (MSE)

Το επόμενο μέτρο σύγκρισης μεταξύ των τύπων παραθύρων ανάλυσης είναι το μέσο τετραγωνικό σφάλμα (Mean Square Error – MSE) των πλατών και των συχνοτήτων του σήματος. Χρησιμοποιώντας ξανά το [PRS08], τώρα οι αποδόσεις του μοντέλου QHM δοκιμάζονται σε σήματα στα οποία έχει προστεθεί θόρυβος.

Η υπόθεση λοιπόν είναι ότι υπάρχει ένα σήμα με τέσσερις συνιστώσες στο οποίο έχει προστεθεί λευκός Gaussian θόρυβος :

$$y(t) = \sum_{k=1}^4 a_k e^{j2\pi f_k t} + v(t)$$

Στον πίνακα 3.1.41, παραθέτονται οι πληροφορίες για τα πλάτη και τις συχνότητες για κάθε έναν από τις τέσσερις συνιστώσες του σήματος. Όπως στο πείραμα υπολογισμού του bias έτσι και εδώ, για την ανάλυση του σήματος, χρησιμοποιήθηκαν τα παράθυρα : Bartlett, Blackman, Flattop, Gauss, Hamming, Hanning, Kaiser, Rectangular, Taylor, Triangle, Tukey. Το μέγεθος των παραθύρων ήταν για όλα τα παράθυρα 17ms και η συχνότητα δειγματοληψίας σταθερή στα 8000 Hz. Η τελευταία γραμμή του πίνακα περιέχει το μεσοδιάστημα της επιτρεπούσας διαφοράς στη συχνότητα ανά συνιστώσα (mismatch). Η τιμή του mismatch πρέπει να είναι ανάλογη με την απόσταση της συχνότητας ανάμεσα στις συνιστώσες.

Για την εκτίμηση της δυνατότητας του QHM στην εκτίμηση συχνότητας, χρησιμοποιήθηκε προσομοίωση Monte Carlo. Για κάθε προσομοίωση, το mismatch της συχνότητας της εκάστοτε συνιστώσας, παράχθηκε από μία ομοιόμορφη κατανομή του αντίστοιχου μεσοδιαστήματος.

Sinusoid	1st	2nd	3rd	4th
Frequency (Hz)	100	200	1000	2000
Amplitude	$e^{j\pi/10}$	$e^{j\pi/4}$	$e^{j\pi/3}$	$e^{j\pi/5}$
Freq. Mismatch (Hz)	± 10	± 10	± 100	± 100

Στη συνέχεια, υποθέτοντας ότι η ισχύς της φασματικής πυκνότητας του θορύβου είναι $V(f)$, προκύπτει ο λόγος του σήματος προς τον θόρυβο (Signal to Noise Ratio – SNR). Το SNR για την k -οστή συνιστώσα του σήματος και με δεδομένες τις συναρτήσεις του QHM, δίδεται από τη σχέση:

$$SNR_k \approx 10 \log_{10} \frac{(2N + 1)|a_k|^2}{V(f_k)}$$

Επιπρόσθετα, ένας σημαντικότερος όρος για την εκπόνηση του MSE πειράματος είναι το βέλτιστο κάτω όριο, γνωστό ως Cramer-Rao lower bound (CRLB), για την συχνότητα και για το πλάτος της κάθε συνιστώσας. Το όριο αυτό μπορεί να μας δείξει την βέλτιστη (χαμηλότερη) απόδοση ενός εκτιμητή παραμέτρων και έτσι μπορούμε εύκολα να συγκρίνουμε τους τύπους παραθύρων, ώστε να βρούμε αυτά με την καλύτερη εκτίμηση παραμέτρων. Το CRLB για τη συχνότητα της k -οστής συνιστώσας δίδεται από τη σχέση:

$$CRLB\{f_k\} = \frac{3V(f_k)}{|a_k|^2 N(N + 1)(2N + 1)}$$

Ενώ για το CRLB του πλάτους δίδεται από τη σχέση:

$$CRLB\{a_k\} = \frac{V(f_k)}{2N + 1}$$

Όπου $2N+1$ είναι το μέγεθος του παραθύρου ανάλυσης σε δείγματα. Τέλος, το MSE επιτέλεσε τον σημαντικότερο παράγοντα του πειράματος εύρεσης του αποδοτικότερου τύπου παραθύρου. Το MSE προκύπτει από τον μέσο όρο των ($M=10^4$ για το πείραμά μας) Monte Carlo simulations, όπου για τις συχνότητες δίδεται από τη σχέση :

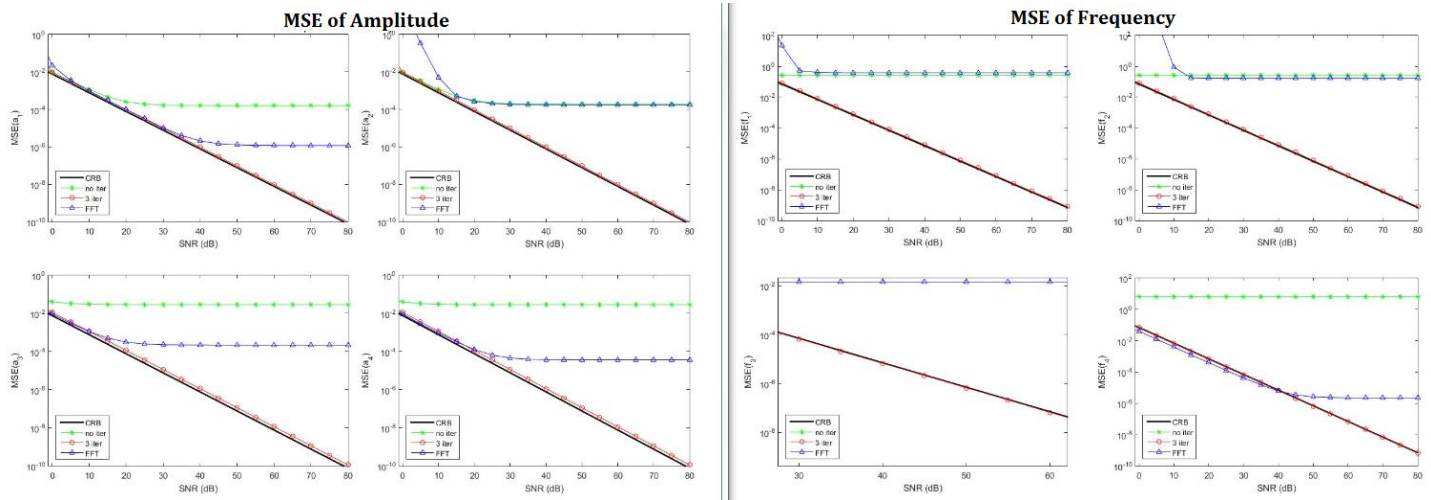
$$MSE\{\hat{f}_k\} = \frac{1}{M} \sum_{i=1}^M |\hat{f}_k(i) - f_k|^2$$

$$MSE\{\hat{a}_k\} = \frac{1}{M} \sum_{i=1}^M |\hat{a}_k(i) - a_k|^2$$

και για τα πλάτη:

Για λόγους σύγκρισης, χρησιμοποιήθηκε εκτίμηση συχνότητας και πλάτους με μέθοδο peak-picking όπως στα κλασικά ημιτονοειδή μοντέλα και αναφέρεται στα γραφήματα ως FFT.

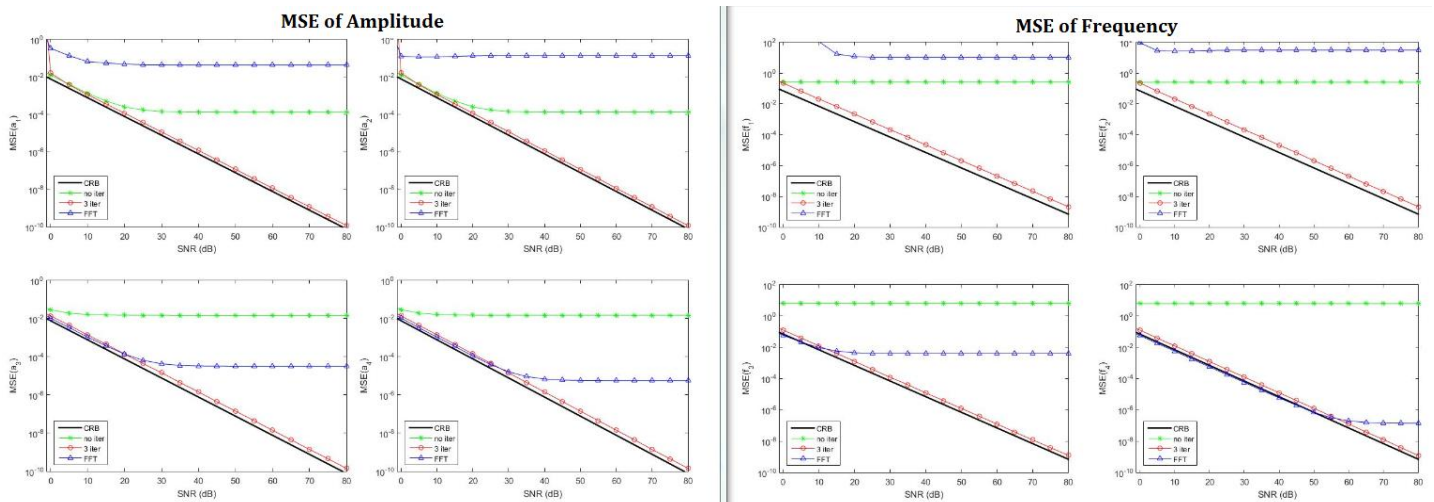
TAYLOR WINDOW



Γράφηματα 3.1.41_{αβ}

MSE για τα πλάτη (αριστερά) και τις συχνότητες (δεξιά) των 4 διαφορετικών συνιστωσών εφαρμόζοντας παράθυρο Taylor.

HAMMING WINDOW



Γράφηματα 3.1.42_{αβ}

MSE για τα πλάτη (αριστερά) και τις συχνότητες (δεξιά) των 4 διαφορετικών συνιστωσών εφαρμόζοντας παράθυρο Hamming.

Στα γραφήματα 3.1.41_{αβ} παρουσιάζεται το MSE του QHM σε πλάτη και συχνότητες, για κάθε μία από τις 4 συνιστώσες, εφαρμόζοντας το παράθυρο Taylor. Σε κάθε ένα γράφημα παρουσιάζεται η τιμή του MSE για την ανάλογη τιμή του SNR(σε decibels). Επίσης στο κάθε

γράφημα φαίνεται η απόδοση για το CRLB(μαύρη γραμμή), για το FFT(μπλε με Δ στη γραμμή) ,για QHM (πράσινο με * στη γραμμή) και για QHM με 2 iterations(κόκκινο με ο στη γραμμή). Επιπρόσθετα, στα γραφήματα 3.1.42_{α,β} παρουσιάζεται το αντίστοιχο παράδειγμα για εφαρμογή του παραθύρου hamming. Αναλυτικότερα, γραφήματα για τα MSE συχνοτήτων και πλατών για όλα τα παράθυρα παραθέτονται στο παράρτημα [B].

Μετά από εφαρμογή όλων των παραθύρων στο συγκεκριμένο πείραμα, το παράθυρο Taylor παρουσίασε τα βέλτιστα αποτελέσματα, καθώς μετά απο 2 iterations του QHM, φτάνει στα όρια του CRLB (βέλτιστη απόδοση). Το παράθυρο Hamming παρουσιάζει και αυτό μεγάλη αξιοπιστία μετά από 2 iterations και σχεδόν εφάπτεται με το CRLB.

Ενδεικτικά, δίδονται οι πίνακες 3.1.43 και 3.1.44 στους οποίους παρουσιάζονται οι τιμές των πλατών και των συχνοτήτων αντίστοιχα, για κάθε συνιστώσα, για κάθε ένα από τα έξι πιο αποδοτικά παράθυρα που χρησιμοποιήθηκαν. Καλό θα ήταν να αναφερθεί ότι όσο μικρότερο το MSE τόσο αποδοτικότερο είναι το παράθυρο . Επίσης, για την καλύτερη οπτική κατανόηση, επιλέχθηκε το φόντο να σκουραίνει όσο καλύτερα αποτελέσματα έχει το MSE.

Πίνακας 3.1.43

MSE για τα πλάτη, με τις τιμές του SNR στα 0, 40, 80 decibels αντίστοιχα, των 4 διαφορετικών συνιστωσών, εφαρμόζοντας τα 6 παράθυρα με το μικρότερο MSE. (Darkest=Best)

Comp – SNR	Bartlett	Gauss	Hamming	Taylor	Triangle	Tukey
1 st 0db	0.011802	0.016579	0.016155	0.009795	0.011115	0.011925
1 st 40db	1.1416e-06	1.2912e-06	1.1963e-06	9.7350e-07	1.0963e-06	1.1428e-06
1 st 80db	9.0248e-02	1.3110e-10	1.2060e-10	9.9283e-11	1.1139e-10	1.1334e-10
2 nd 0db	0.011656	0.015765	0.016189	0.009535	0.011546	0.011517
2 nd 40db	1.1477e-06	1.2790e-06	1.1783e-06	9.7479e-07	1.1194e-06	1.1066e-06
2 nd 80db	1.1373e-10	1.3065e-10	1.1689e-10	9.7564e-11	1.1043e-10	1.1482e-10
3 rd 0db	0.014074	0.026425	0.013955	0.011492	0.013909	0.010568
3 rd 40db	1.4076e-06	1.5304e-06	1.4489e-06	1.1583e-06	1.3799e-06	1.0360e-06
3 rd 80db	1.4078e-10	1.5659e-10	1.4180e-10	1.1638e-10	1.3488e-10	1.0444e-10
4 th 0db	0.014020	0.015673	0.014115	0.011522	0.013632	0.010414
4 th 40db	1.3924e-06	1.5803e-06	1.4175e-06	1.1377e-06	1.3838e-06	1.0684e-06
4 th 80db	1.4327e-10	1.5975e-10	1.4368e-10	1.1662e-10	1.4074e-10	1.0441e-10

Γίνεται, επομένως, εύκολα αντιληπτό μέσω των αριθμητικών αποτελεσμάτων, ότι οι δυνατότητες του παραθύρου Taylor είναι εξαιρετικά υψηλές. Στο MSE της συχνότητας υπερτερεί σε όλες τις τιμές του SNR και σε όλες τις συνιστώσες. Χάνει την πρωτοκαθεδρία του μόνο στο MSE του πλάτους, στον τρίτο και στην τέταρτη συνιστώσα. Ωστόσο, το Taylor, ακόμα και σε αυτές τις κατηγορίες παραμένει σε υψηλές αποδόσεις, δεύτερο, πίσω απο το παράθυρο Tukey.

Αναλυτικότερα, τα αποτελέσματα από όλα τα πειράματα MSE που παράχθηκαν για τις ανάγκες της παρούσας εργασίας, παρουσιάζονται σε λεπτομερείς πίνακες στο παράρτημα [Γ].

Comp – SNR	Bartlett	Gauss	Hamming	Taylor	Triangle	Tukey
1 st 0db	0.199765	0.291904	0.221478	0.081603	0.186391	0.190454
1 st 40db	1.9494e-05	2.8769e-05	2.1621e-05	8.1901e-06	1.6980e-05	1.8670e-05
1 st 80db	1.87890e-09	2.7775e-09	2.1550e-09	7.8630e-10	1.7616e-09	1.8069e-09
2 nd 0db	0.202876	0.295418	0.218701	0.081970	0.189351	0.192748
2 nd 40db	1.9438e-05	2.9073e-05	2.0710e-05	8.0539e-06	1.7478e-05	1.8825e-05
2 nd 80db	1.9230e-09	2.8676e-09	2.1633e-09	7.9494e-10	1.7842e-09	1.7984e-09
3 rd 0db	0.103323	0.150339	0.127900	0.067945	0.099369	0.081760
3 rd 40db	1.0104e-05	1.4815e-05	1.2449e-05	6.6677e-06	9.5976e-06	8.1468e-06
3 rd 80db	1.0407e-09	1.5409e-09	1.2518e-09	6.7201e-10	9.8355e-10	8.4992e-10
4 th 0db	0.102691	0.152781	0.122223	0.067700	0.095392	0.084525
4 th 40db	1.0367e-05	1.5084e-05	1.2258e-05	6.6735e-06	9.6652e-06	8.2642e-06
4 th 80db	1.0163e-09	1.4954e-09	1.2219e-09	6.6555e-10	9.9997e-10	8.0826e-10

Πίνακας 3.1.44

MSE για τις συχνότητες, με τις τιμές του SNR στα 0, 40, 80 decibels αντίστοιχα, των 4 διαφορετικών συνιστωσών, εφαρμόζοντας τα 6 παράθυρα με το μικρότερο MSE. (Darkest=Best)

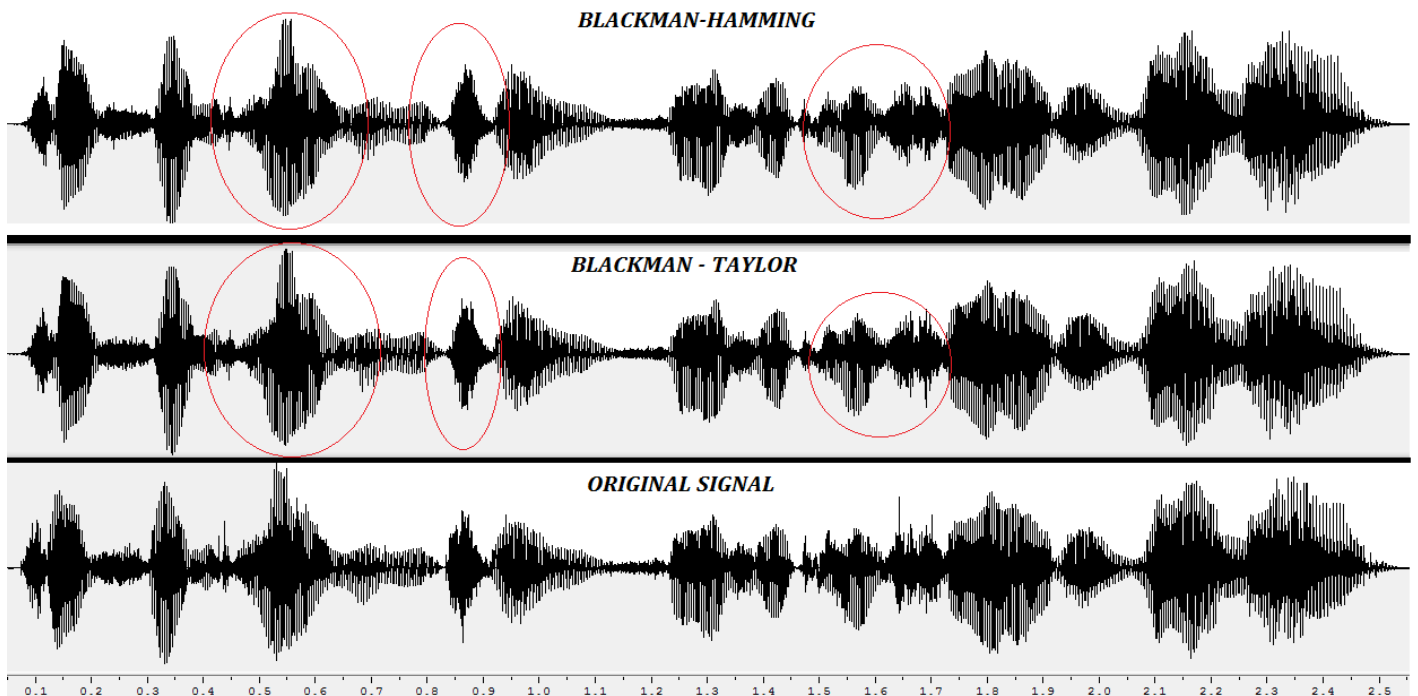
3.1.5 Συμπεράσματα

Μετά την εκτέλεση των συνθετικών πειραμάτων σχετικά με το bias και το MSE, και αφού έγινε αναλυτική σύγκριση απόδοσης μεταξύ διαφόρων τύπων παραθύρων και στα δύο πειράματα, τα αριθμητικά αποτελέσματα κατέδειξαν το παράθυρο που αναζητούσαμε ως βέλτιστο. Το παράθυρο Taylor παρουσίασε ικανοποιητικά αποτελέσματα στο πείραμα εύρους μηδενικού bias, ωστόσο στο πείραμα MSE οι επιδόσεις του έκαναν την διαφορά ανάμεσα στα άλλα παράθυρα ανάλυσης, φθάνοντας οριακά σε τέλεια νούμερα. Το γεγονός αυτό, το έχρησε ως το πλέον ελπιδοφόρο παράθυρο για εφαρμογή σε πραγματικά σήματα φωνής.

3.2 Πειράματα σε πραγματικά σήματα

3.2.1 Εισαγωγή

Στο δεύτερο μέρος των πειραμάτων της παρούσας εργασίας, διαδραματίστηκαν πειράματα σε πραγματικά σήματα φωνής με όσες ιδιαιτερότητες, ιδιομορφίες και απρόβλεπτες ανακατατάξεις πλατών και συχνοτήτων μπορεί αυτά να παρουσιάζουν. Το μοντέλο που χρησιμοποιήθηκε για την ανάλυση και την ανακατασκευή των πραγματικών σημάτων φωνής είναι το full-band eaQHM, όπως αυτό αναλύθηκε στην ενότητα 2.5, καθώς αποτελεί μέχρι πρότινος την πιο αξιόπιστη και συνάμα πιο καινοτόμα μέθοδο ανάλυσης σημάτων φωνής.



Γραφήματα 3.2.11_{α,β,γ}

Η κυματομορφή του αυθεντικού σήματος φωνής (κάτω) και των αντίστοιχων ανασυνθέντων σημάτων (πάνω), αρχικά με χρήση παραθύρου Blackman για την αρχική εκτίμηση των παραμέτρων και στη συνέχεια με χρήση των παραθύρων Taylor και Hamming, εναλλάξ.

Στο σχήμα εντοπίζονται επίσης και σημεία στα οποία η ανασύνθεση του σήματος περιέχει σημαντικές διαφορές σε σχέση με το αυθεντικό σήμα.

Με αφετηρία λοιπόν τα συμπεράσματα από την εκτέλεση των συνθετικών πειραμάτων, και αφού τα αριθμητικά αποτελέσματα του παραθύρου Taylor παρέπεμπαν σε ένα παράθυρο που θα βελτίωνε την απόδοση των σχεδον-αρμονικών μοντέλων, ήταν αναγκαία η έρευνα και απόδειξη της δυναμικότητάς του σε πραγματικά σήματα φωνής.

Στην αρχική έκδοση του μοντέλου [KOS14], εφαρμόζεται το παράθυρο Blackman στην πρώτη ανακατασκευή του σήματος (χωρίς προσαρμοστικότητα). Στη συνέχεια, για τις πεπερασμένες προσαρμογές ανακατασκευής των ημιτονοειδών παραμέτρων, εφαρμόζεται το παράθυρο Hamming. Στα γραφήματα 3.2.11_{α,β,γ}, παρουσιάζονται οι κυματομορφές ενός σήματος, αρχικά το αυθεντικό και στη συνέχεια τα δύο ανακατασκευασμένα σήματα χρησιμοποιώντας παράθυρο Taylor και Blackman, αντίστοιχα.

3.2.2 Βάση Δεδομένων

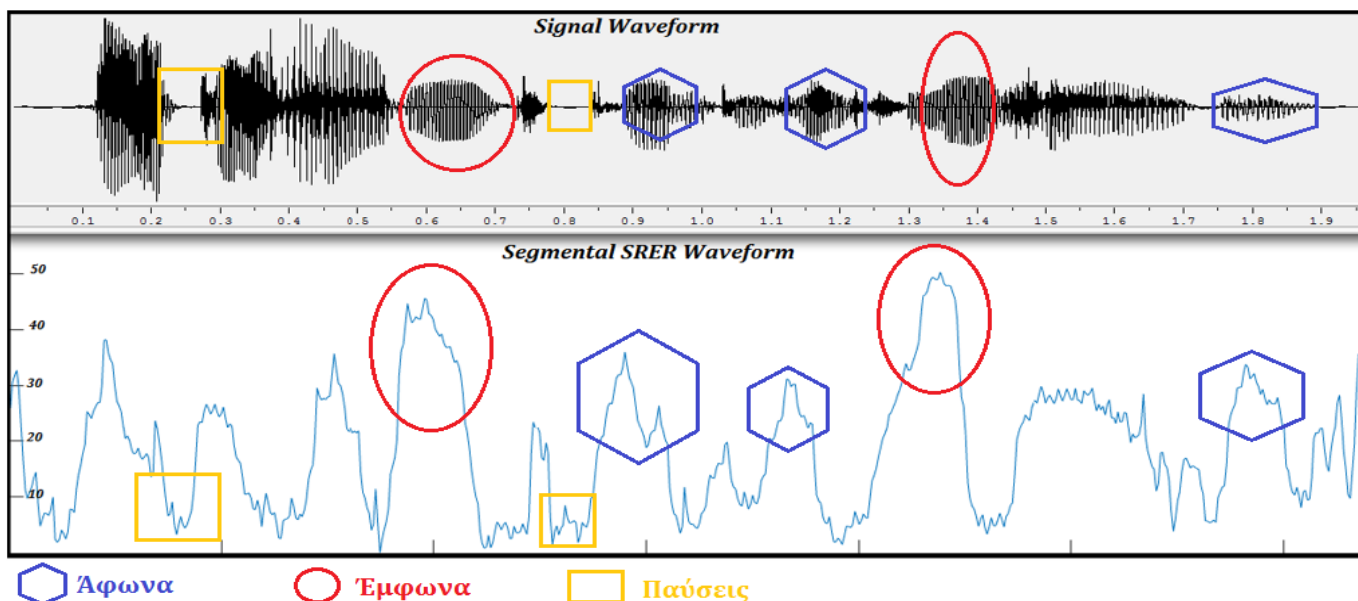
Για τις ανάγκες του πειράματος πραγματικών σημάτων, χρησιμοποιήθηκε σαν είσοδος αυθεντικών σημάτων, βάση δεδομένων με πραγματικά σήματα φωνής. Η βάση δεδομένων περιείχε φωνές ανδρών και γυναικών από 16 διαφορετικές γλώσσες και διαλέκτους, και πιο συγκεκριμένα: Αγγλικά, Ιαπωνικά, Γερμανικά, Ιταλικά, Γαλλικά, Ελληνικά, Ισπανικά, Βάσκικα, Ισλανδικά, Ινδικά, Τούρκικα, Φινλανδικά, Αραβικά, Κινέζικα, Κορεάτικα και Ρώσικα.

Για κάθε μία απο τις παραπάνω γλώσσες περιέχεται μία αντρική και μία γυναικεία ηχογράφηση διάρκειας δύο έως τεσσάρων δευτερολέπτων αποθηκευμένα στη βάση σαν Waveform Audio File Format (γνωστά ως .wav αρχεία) και για το πείραμα χρησιμοποιήθηκε συχνότητα δειγματοληψίας 16KHz.

3.2.3 Δομή Πειράματος

Με σκοπό την ανάδειξη του παραθύρου με την καλύτερη ανακατασκευή σήματος, χρησιμοποιήσαμε το full band eaQHM [KOS11], εφαρμόζοντας τα πραγματικά σήματα φωνής (3.2.2) και εναλλάσσοντας το είδος των παραθύρων ανάλυσης καθώς επίσης και το βήμα ανάλυσης στα 5ms , 7.5 ms , 10 ms, 12,5 και 15 ms , εναλλάξ για κάθε παράθυρο και κάθε σήμα φωνής. Επίσης, όπως και στο [KOS11], χρησιμοποιήθηκε η τιμή του SRER (όλου του σήματος) για την εκτίμηση του υπολοίπου της «χαμένης» ενέργειας για κάθε παράθυρο. Όσο μεγαλύτερη είναι η τιμή του SRER, τόσο μεγαλύτερη πληροφορία συλλαμβάνεται από το μοντέλο και το εφαρμοσμένο παράθυρο.

Ωστόσο, στο συγκεκριμένο πείραμα, εκτός απο το ολικό SRER, χρησιμοποιήθηκε και η τιμή του τμηματικού SRER (segmental SRER), για τη ακριβέστερη καταγραφή του βέλτιστου παραθύρου ανάλυσης. Για τη δημιουργία του τμηματικού SRER, χρησιμοποιήθηκε το αυθεντικό SRER σε τμήματα των 18.6 ms τα οποία εκτείνονταν σε ολόκληρο το σήμα, με βήμα 3.1 ms και τελικά ο μέσος όρος όλων αυτών των τιμών SRER, δίνει την τιμή του Segmental SRER.



Γράφημα 3.2.31

Η κυματομορφή του γυναικείου Γερμανόφωνου σήματος και της αντίστοιχης του τμηματικού SRER εφαρμόζοντας παράθυρο Taylor, με το βήμα ανάλυσης παραθύρου στα 12,5 ms.

Τέλος, υπολογίσθηκε η τυπική απόκλιση ανάμεσα στις τιμές του SRER κάθε τμήματος στο σύνολο του σήματος. Μεγάλη τιμή στην τυπική απόκλιση συνεπάγεται ότι το παράθυρο σε άλλα κομμάτια (εμφωνά ή άφωνα) κάνει καλή εκτίμηση των παραμέτρων, ενώ σε άλλα κομμάτια του ίδιου σήματος, κάνει χειρότερη εκτίμηση. Ενδεικτικά, στο σχήμα 3.2.31 παρουσιάζεται ένα σήμα φωνής και οι αντίστοιχες τιμές του τμηματικού SRER κατά την διάρκεια του σήματος. Έτσι απεικονίζονται οι περιοχές του σήματος όπου η συλλογή πληροφορίας από το μοντέλο έχουν υψηλή ή χαμηλή απόδοση. Όπως φαίνεται και στο σχήμα, τα έμφωνα τμήματα έχουν το μεγαλύτερο SRER, δηλαδή την καλύτερη εκτίμηση των παραμέτρων τους. Τα άφωνα τμήματα έχουν λίγο χειρότερη απόδοση στις εκτιμήσεις, ενώ τα τμήματα που υπάρχει παύση, δεν γίνεται καλή εκτίμηση, με το SRER να πέφτει πολλές φορές κάτω από τους 10 βαθμούς.

3.2.4 Αποτελέσματα πειράματος

Κατά την διάρκεια εκτέλεσης του συγκεκριμένου πειράματος, λήφθηκαν σημαντικά αποτελέσματα και σε μεγάλο βαθμό έχρισαν το παράθυρο Taylor καλύτερο (τουλάχιστον αριθμητικά) έναντι του παραθύρου Hamming. Ωστόσο, στην πορεία και αφού ακούστηκαν τα παραχθέντα σήματα από το full band eaQHM, τα αριθμητικά αποτελέσματα των SRER και Segmental SRER σε κάποιες περιπτώσεις δε συμβάδιζαν με την ποιότητα ήχου των παραγόμενων σημάτων.

KOREAN M.	5 ms	7.5 ms	10 ms	12.5 ms	15 ms
1ST SRER	19.2698	16.0962	13.3686	10.9963	10.0026
SRER	29.5945(-0.6)	20.9391(+0.2)	16.7196(+1.0)	14.9557(+0.8)	13.3390(+0.7)
SEGM. SRER	27.6001(10.9)	22.0059(9.8)	17.2500(8.8)	15.6141(8.4)	13.4547(7.9)
ADAPTATIONS	6	4	2	5	4
KOREAN FEM.	5 ms	7.5 ms	10 ms	12.5 ms	15 ms
1ST SRER	16.3813	13.6791	11.9463	10.6016	9.8648
SRER	20.0278(+0.4)	15.8461	14.5484(+0.5)	13.7053(+0.6)	12.8591(+0.7)
SEGM. SRER	22.8083(9.8)	17.7220(9.4)	16.4166(9.7)	15.5440(9.7)	14.4743(9.6)
ADAPTATIONS	4	2	3	5	4
RUSSIAN M.	5 ms	7.5 ms	10 ms	12.5 ms	15 ms
1ST SRER	20.2448	17.0169	14.8653	13.3074	11.9451
SRER	28.7623(+0.1)	23.8187(+0.4)	20.2982(+0.3)	17.0140(+0.4)	15.5108(+0.4)
SEGM. SRER	27.6600(11.2)	23.4790(11.7)	18.6808(10.4)	15.1947(9.4)	13.4103(9.2)
ADAPTATIONS	3	4	5	3	3
RUSSIAN FEM.	5 ms	7.5 ms	10 ms	12.5 ms	15ms
1ST SRER	15.8164	12.2771	10.4241	8.9532	7.8970
SRER	22.2008(+1.1)	14.4225(+0.1)	13.2123(+0.7)	12.3085(+0.9)	10.5794(+0.5)
SEGM. SRER	21.7117(11.4)	14.9928(9.5)	13.1914(9.3)	12.2621(9.0)	11.3432(8.9)
ADAPTATIONS	8 (!!!)	2	2	3	2

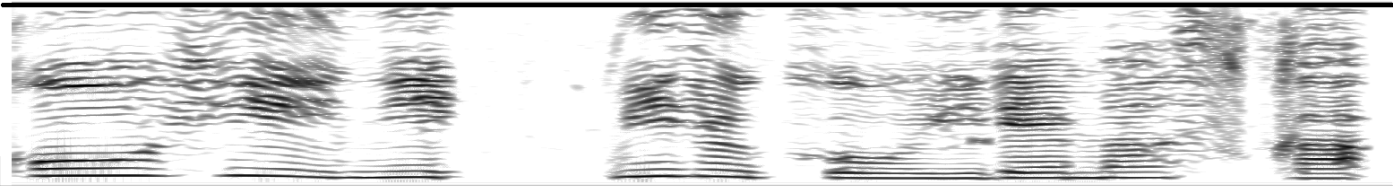
Πίνακας 3.2.41

Καταγραφή του SRER με χρήση παραθύρου Taylor (σε παρένθεση η διαφορά του από το αντίστοιχο Hamming) και του Segmental SRER(σε παρένθεση η τυπική απόκλιση των τιμών του), καθώς και ο αριθμός των adaptations που αυτά χρειάστηκαν για να ολοκληρώσουν την ανασύνθεση του σήματος, για 4 σήματα(Κορεάτικα-Ρώσικα άνδρας και γυναίκα), με το βήμα ανάλυσης στα 5, 7.5, 10, 12.5 και 15ms.

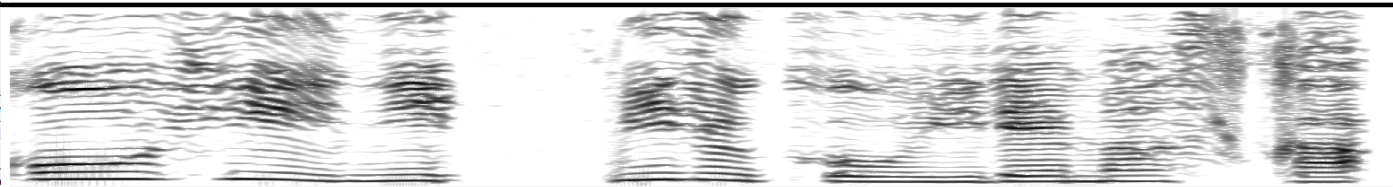
Αναλυτικότερα, με την παρατήρηση των πρώτων αριθμητικών αποτελεσμάτων, το παράθυρο Taylor είναι καλύτερο σε σχέση με το Hamming σχεδόν σε όλη την βάση δεδομένων (Στον Πίνακα 3.2.41 δίδονται αποτελέσματα από 4 φωνές, ενώ στο παράρτημα [Δ] δίδονται όλοι οι πίνακες αποτελεσμάτων του συγκεκριμένου πειράματος).

Παρατηρείται ότι όταν το βήμα ανάλυσης του παραθύρου είναι μικρό, οπότε η ανάλυση είναι υψηλής ποιότητας και το SRER είναι μεγαλύτερο του 20, το παράθυρο Taylor έχει λίγο χειρότερα αριθμητικά αποτελέσματα από το Hamming. Όμως, αυτή η διαφορά είναι αμελητέα καθώς η ποιότητα του ανασυνθέντος σήματος είναι πολύ υψηλή για τιμές του SRER μεγαλύτερες των 15 μονάδων και δεν παρατηρείται κάποια διαφορά από το ανθρώπινο αυτί. Η παρατήρηση αυτή μας οδήγησε στην ανάγκη να εξετάσουμε τα ανασυνθέντα σήματα με SRER μικρότερο του 15. Τέτοια σήματα ανασυνθέθηκαν όταν το βήμα ανάλυσης ήταν μεγαλύτερο ή ίσο των 12.5 ms.

Σπεκτόγραμμα εφαρμόζοντας παράθυρο Taylor στο eaQHΜ.



Σπεκτόγραμμα εφαρμόζοντας παράθυρο Hamming στο eaQHΜ.



Σχήμα 3.2.42_{α,β}

Σπεκτόγραμμα για το σήμα της Ιαπωνικής γυναικείας φωνής με χρήση παραθύρου Taylor(α) και Hamming (β), με βήμα ανάλυσης παραθύρου τα 15ms.

Είναι φανερή η ύπαρξη θορύβου στο σπεκτόγραμμα του παραθύρου Taylor, σε αντίθεση με το αντίστοιχο του παραθύρου Hamming.

Πιο συγκεκριμένα, στις Κινέζικες και Κορεάτικες φωνές, ανδρών και γυναικών, με το βήμα ανάλυσης στα 12.5ms και στα 15ms, εκτός από αριθμητική υπεροχή του παραθύρου Taylor, υπάρχει και σαφώς καλύτερη ποιότητα στο ανακατα-σκευασμένο σήμα. Διακρίνεται κάποιου είδους θόρυβος και το SRER είναι χαμηλό, πολλές φορές μικρότερο των 10 βαθμών. Κάποια εξίσου σημαντικά πλεονεκτήματα του παραθύρου Taylor παρατηρήθηκαν στην Ιταλική φωνή (αντρική και γυναικεία - στα 12.5ms βήματος ανάλυσης), καθώς και στην Αραβική γυναικεία φωνή (στα 12.5ms και 15ms), στις οποίες το ανασυνθεμένο σήμα του παραθύρου Hamming δημιουργούσε κάποιο μικροφωνισμό και θόρυβο σε σχέση με το αντίστοιχο Taylor (Σχήμα 3.2.42_{α,β}).

Ενώ όλα έδειχναν ότι θα οδηγηθούμε σε ένα νέο ιδανικό παράθυρο, με δυναμική και αντόχη ακόμη και σε μεγάλα βήματα ανάλυσης, κάποια από τα αποτελέσματα δημιούργησαν ενδοιασμούς σχετικά με τις δυνατότητές του. Παρατηρήθηκε αρχικά μια αδυναμία του παραθύρου Taylor (και ενώ πραγματοποιούσε κατά μέσο όρο περισσότερα adaptions από το Hamming) να ανέβει σε πολύ υψηλούς αριθμούς SRER, όπως το Hamming που προσέγγιζε τα 30 dB SRER όταν το βήμα ανάλυσης ήταν στα 5ms. Μολαταύτα, όπως αναφέρθηκε και

νωρίτερα, όταν το ανασυνθεμένο σήμα έχει SRER μεγαλύτερο από τους 15 βαθμούς δεν παρουσιάζει κάποια ιδιαίτερη ακουστική ανωμαλία σε σχέση με το αυθεντικό σήμα.

Ωστόσο, ακούγοντας εξονυχιστικά και σχολαστικά κάποια από τα παραχθέντα σήματα από το παράθυρο Taylor, και ενώ το SRER ήταν αρκετά υψηλό, διαπιστώθηκε ότι κατά την ανασύνθεση τους είχε προστεθεί θόρυβος τόσο σε χαμηλές όσο και σε υψηλές συχνότητες. Χαρακτηριστικά παραδείγματα ανασυνθεμένων σημάτων που έχει προστεθεί θόρυβος σε μικρό βήμα ανάλυσης του παραθύρου Taylor, είναι στην Φινλανδική ανδρική φωνή, καθώς και στην Κινέζικη ανδρική φωνή σε βήμα ανάλυσης μόλις 5 ms.

3.2.5 Συμπεράσματα

Συνοψίζοντας, μπορούμε να επισημάνουμε, ότι το παράθυρο Taylor σε κάποιες περιπτώσεις κάνει εξαιρετική ανασύνθεση φωνής ακόμα και με μεγάλο βήμα ανάλυσης παραθύρου. Στην πλειονότητα των παραδειγμάτων έδειξε υπεροχή σε σχέση με το καθιερωμένο (από την αρχική έκδοση του eaQHM) παράθυρο Hamming.

Ωστόσο, κάποιες αστάθειες που παρουσίασε το παράθυρο Taylor με την δημιουργία θορύβου σε ανύποπτα σημεία στα οποία θα έπρεπε να έχουμε σχεδόν τέλεια ανασύνθεση, μας αναγκάζει να αντιμετωπίσουμε με επιφυλακτικότητα την αρχική αριθμητική και ακουστική υπεροχή του.

Το σημαντικότερο ίσως συμπέρασμα είναι ότι αποδείχθηκε πως το μοντέλο eaQHM μπορεί να δώσει πολύ καλύτερα ανασυνθεμένα σήματα όταν επιλεγθεί, ανάλογα με την περίπτωση, το κατάλληλο παράθυρο ανάλυσης.

4 Συμπέρασμα και μελλοντική εργασία

Όπως είδαμε στην παράγραφο 3.1, το παράθυρο Taylor υπερτερεί στα συνθετικά πειράματα του Bias και του MSE, των υπολοίπων παραθύρων ανάλυσης, παρουσιάζοντας σχεδόν τέλεια αποτελέσματα. Δε συμβαίνει όμως το ίδιο σε όλο το εύρος των πραγματικών σημάτων. Οι ιδιαιτερότητες και οι λεπτομέρειες που κρύβει η ανθρώπινη φωνή, την καθιστούν πολλές φορές απρόβλεπτη και τα θεωρητικά αποτελέσματα δεν έχουν καθολική συνέπεια εφαρμογής σε ένα πραγματικό σήμα φωνής.

Ο θόρυβος που προστέθηκε σε ορισμένες περιπτώσεις με το παράθυρο Taylor μας δημιούργησαν μία βάση δεδομένων ανασυνθεμένων σημάτων φωνής, όπου όταν το βήμα ανάλυσης ήταν μικρό, τα παραχθέντα σήματα από το παράθυρο Hamming έχουν καλύτερη ποιότητα και μηδενικό θόρυβο. Όμως όσο το βήμα ανάλυσης μεγάλωνε, το παράθυρο Taylor έδειχνε σταθερότητα και υπερτερούσε σε ποιότητα και SRER του παραθύρου Hamming.

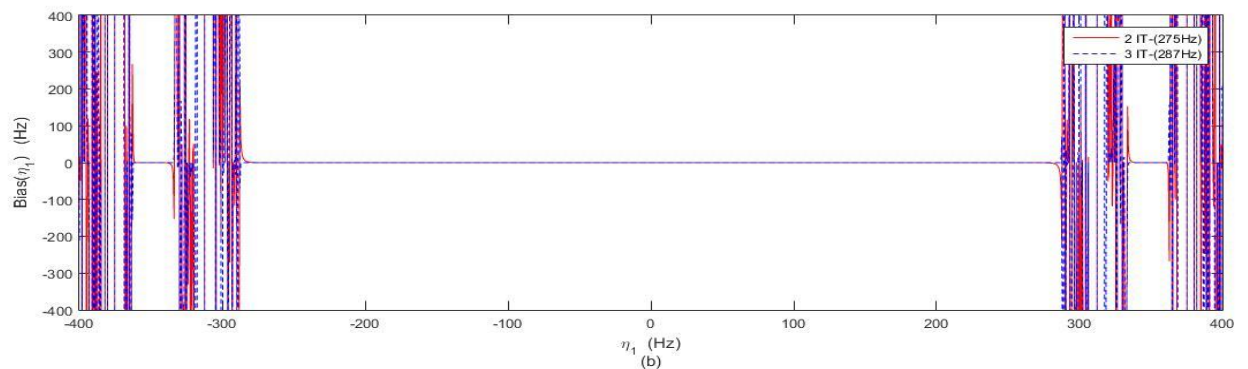
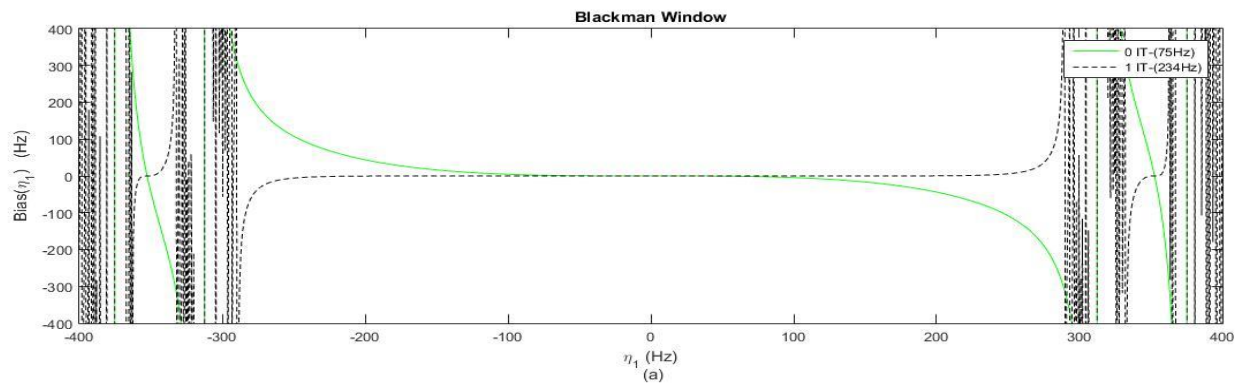
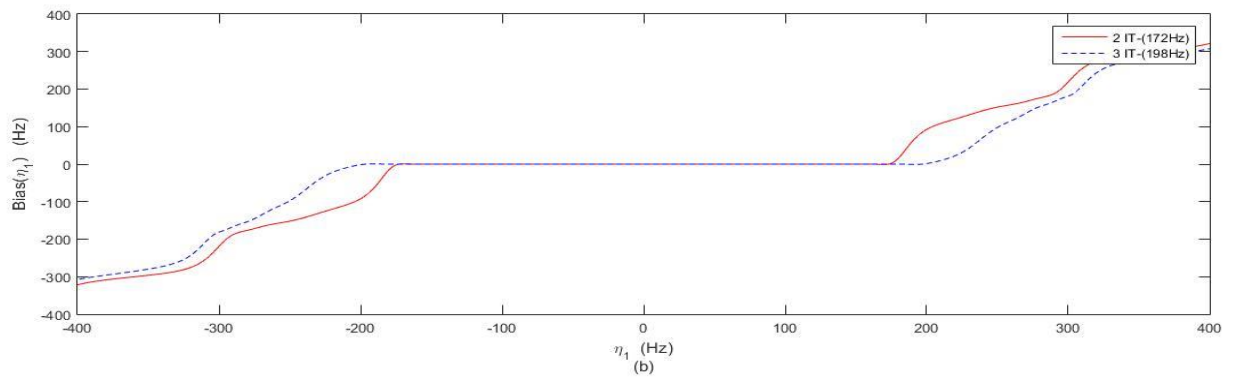
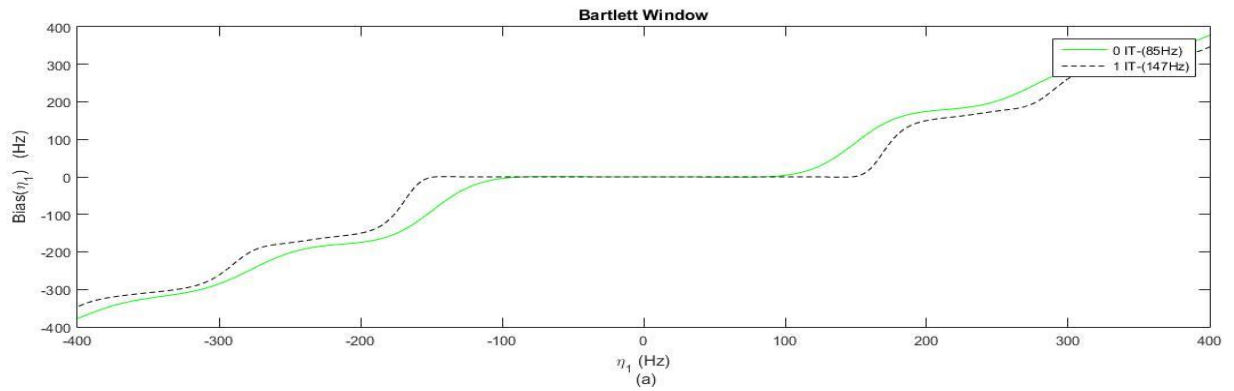
Με δεδομένη την ανομοιομορφία και την αστάθεια των ανασυνθεμένων σημάτων, αποτράπηκε η δημιουργία ενός listening test στο οποίο θα μπορούσε να αναδειχθεί το παράθυρο ανάλυσης με την αποδοτικότερη ανασύνθεση φωνής.

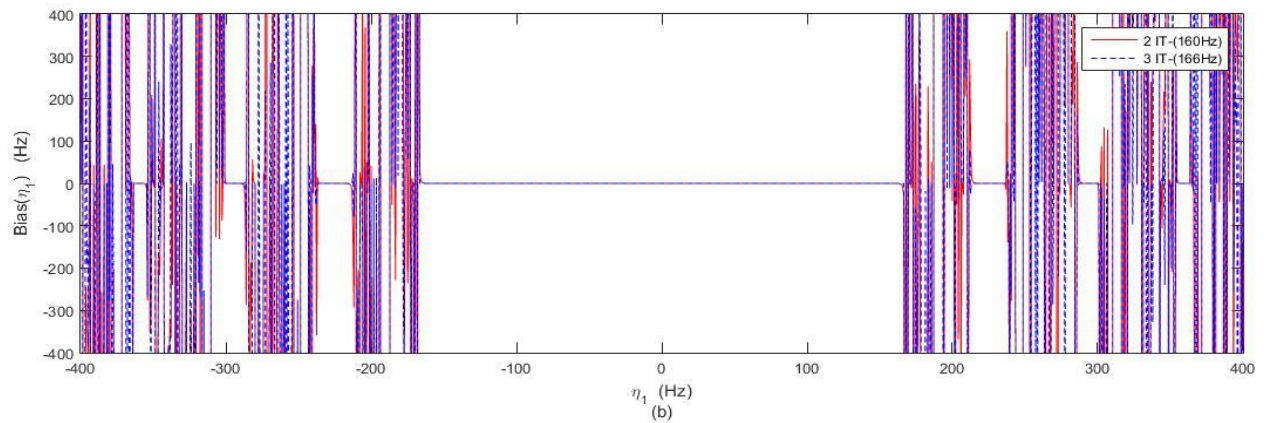
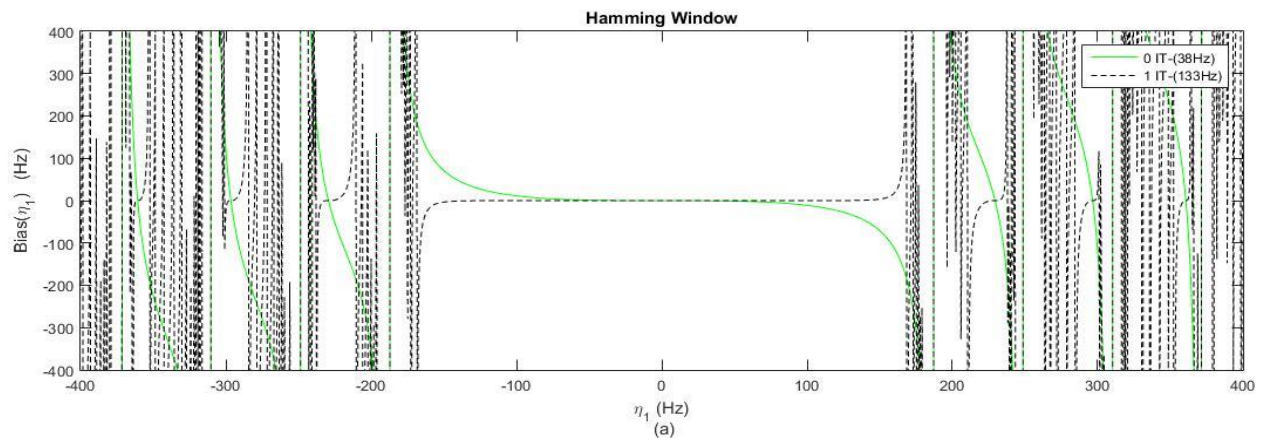
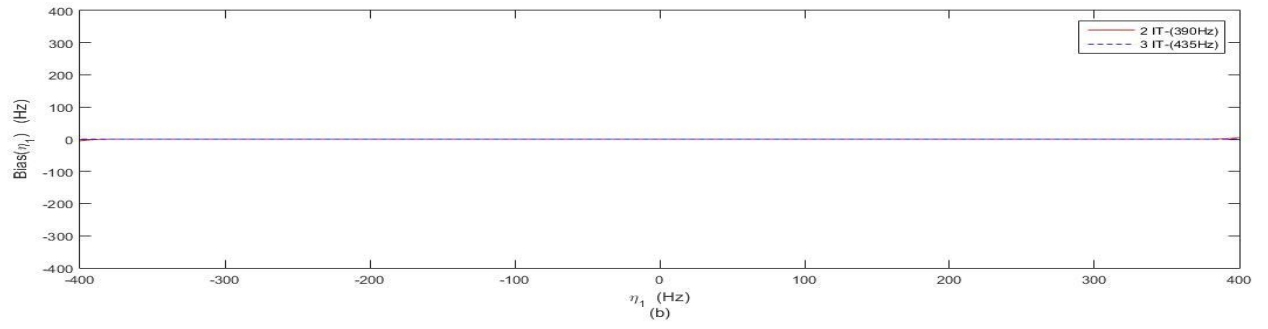
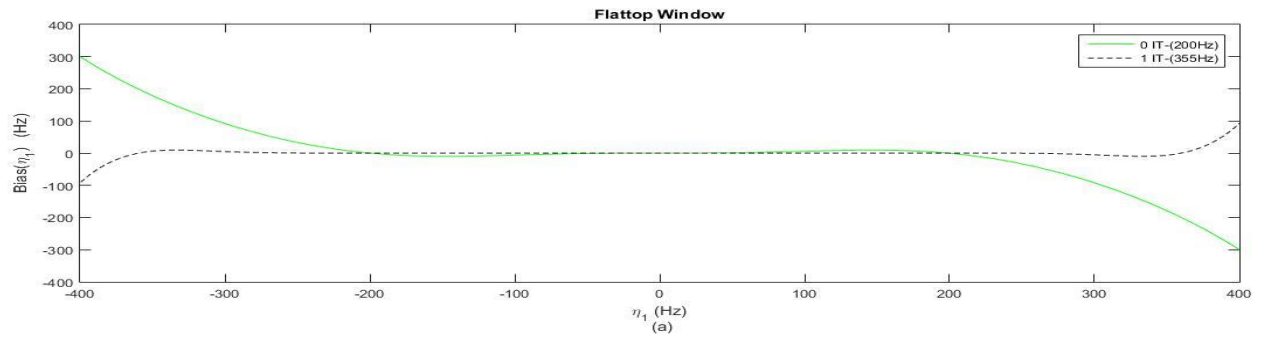
Η παρούσα εργασία κλείνει αφήνοντας αρκετά μέτωπα ανοιχτά και πολλά ερωτήματα να αναζητούν μία απάντηση. Το κυριότερο από αυτά είναι ο λόγος που δημιουργεί θόρυβο το μοντέλο eaQHM, σε μικρό βήμα ανάλυσης, με εφαρμογή του παραθύρου Taylor. Το υψηλό SRER, που ισοδυναμεί με λίγη χαμένη ενέργεια σε σχέση με το αυθεντικό σήμα, παραμένει μόνο ένας αριθμός, ο οποίος σε κάποιες περιπτώσεις δεν συμβαδίζει με την ποιότητα του παραγόμενου σήματος.

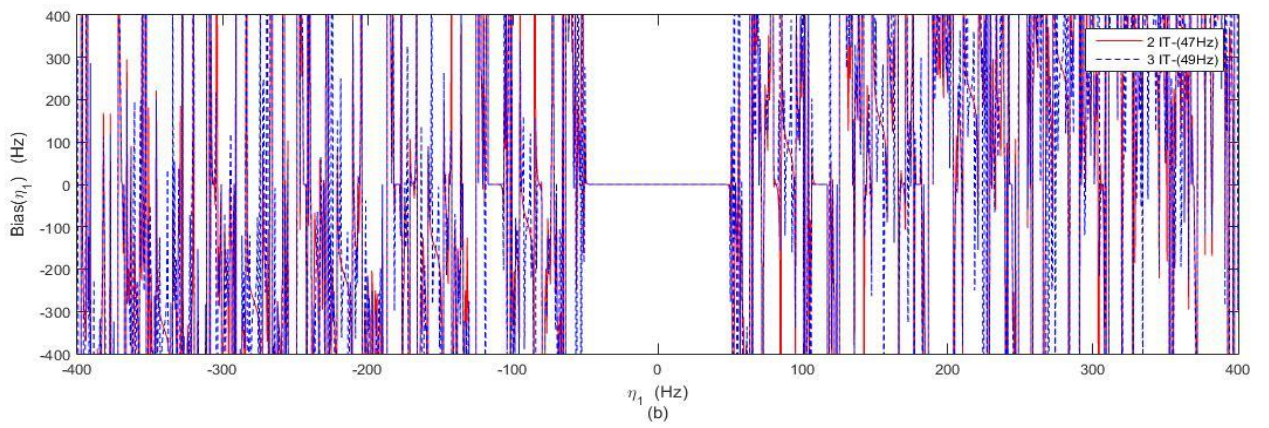
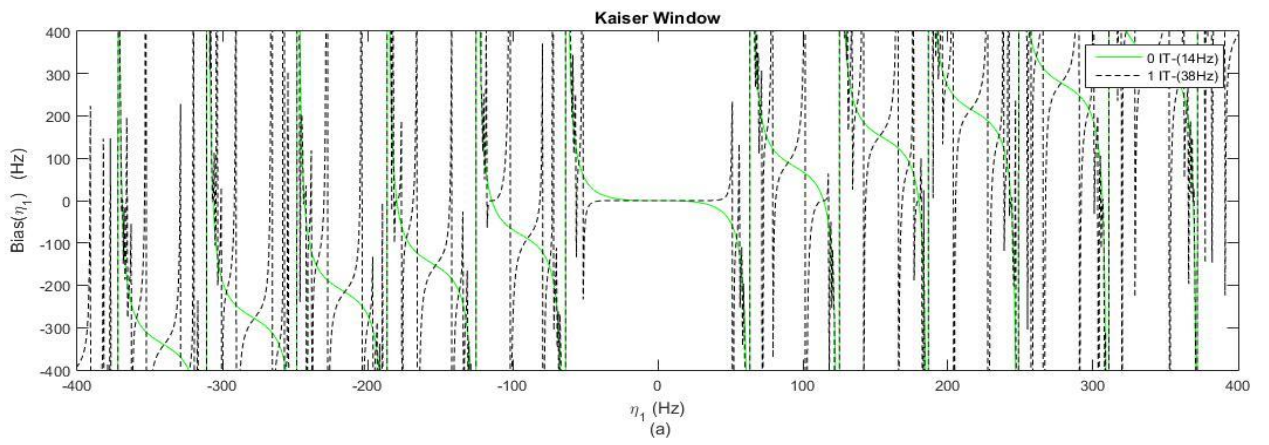
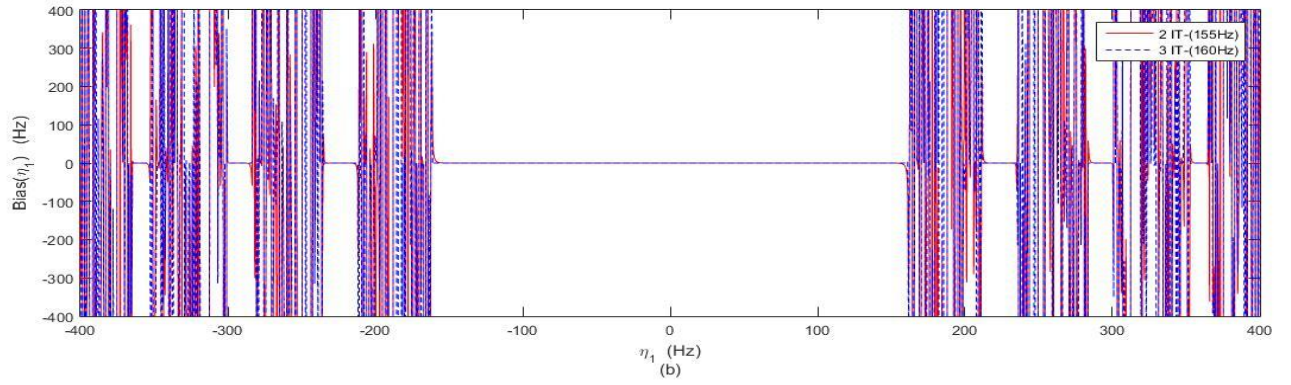
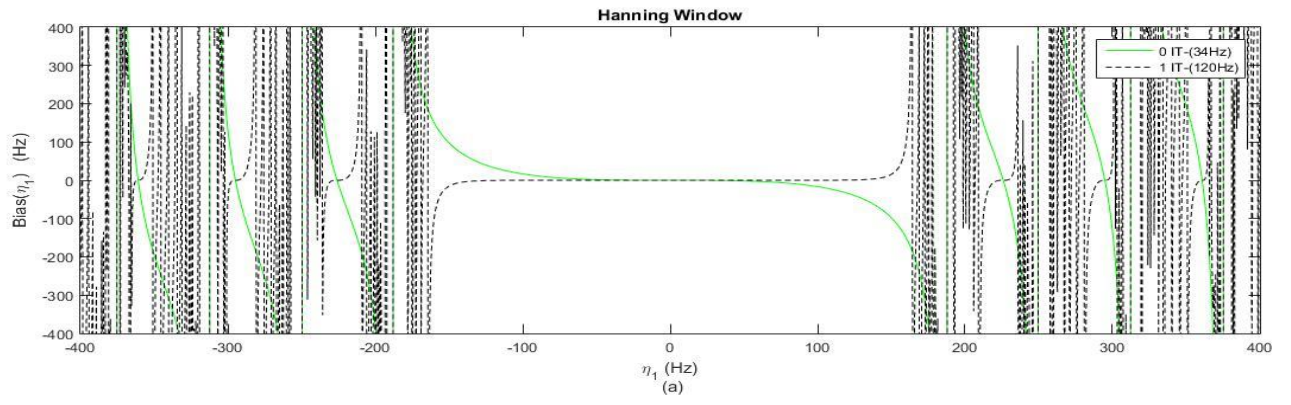
Ένας εξίσου σημαντικός παράγοντας που χρήζει ιδιαίτερης έρευνας, αποτελεί η τμηματική ανασκόπηση των ανασυνθεμένων σημάτων. Η εξονυχιστική στατιστική έρευνα των κυματομορφών των Segmental SRER σε σχέση με τις κυματομορφές των σημάτων και η κατάληξη σε συμπεράσματα σχετικά με την ποιότητα ανασύνθεσης των σημάτων σε έμφωνα και άφωνα τμήματα.

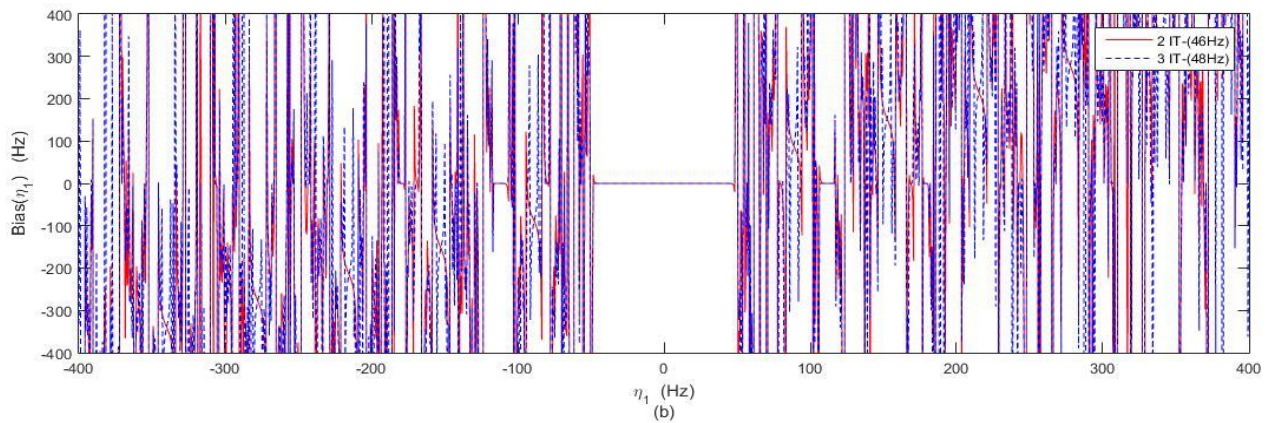
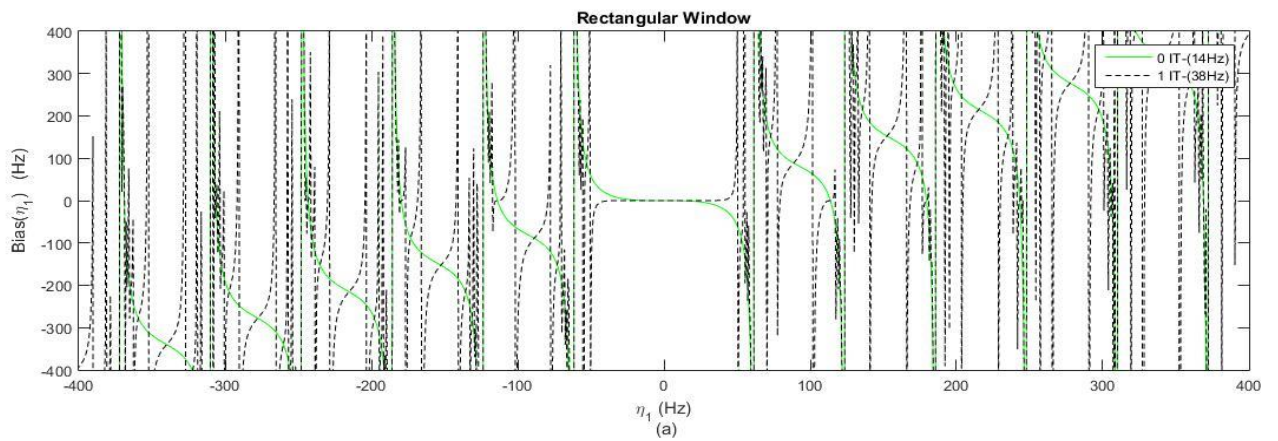
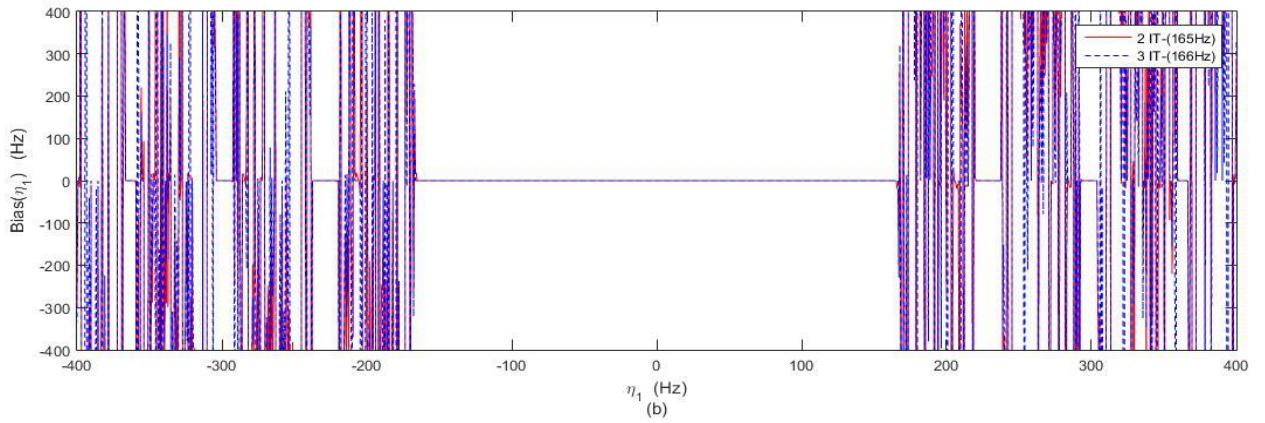
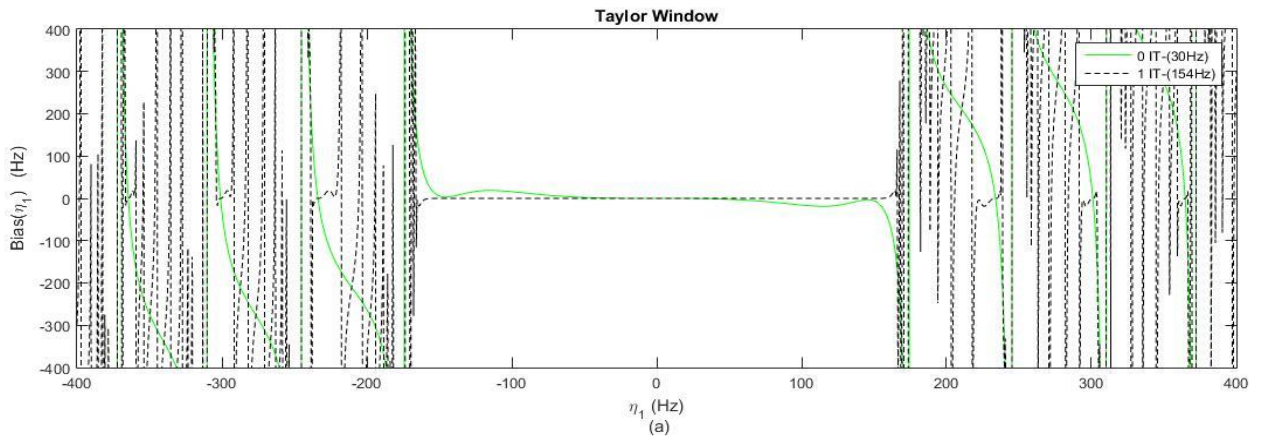
5 Παράρτημα

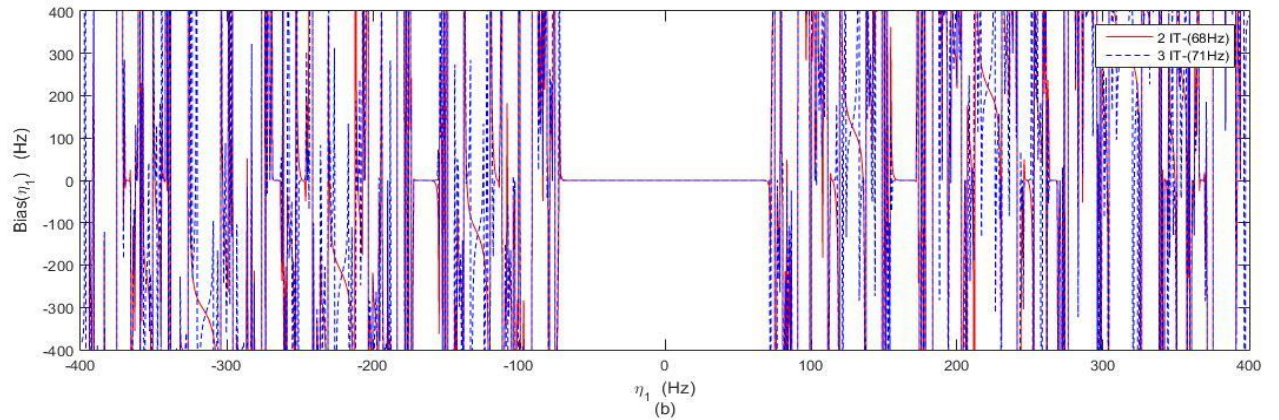
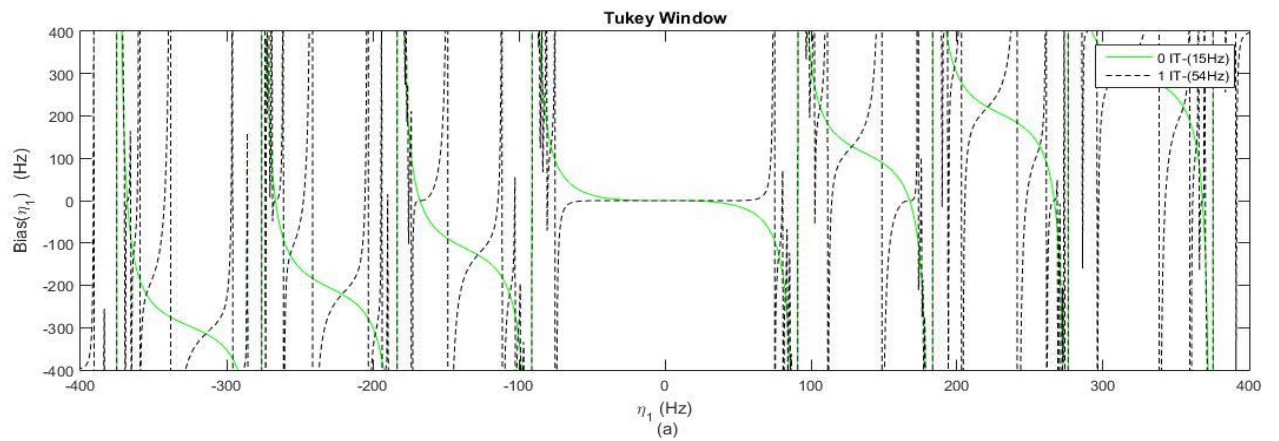
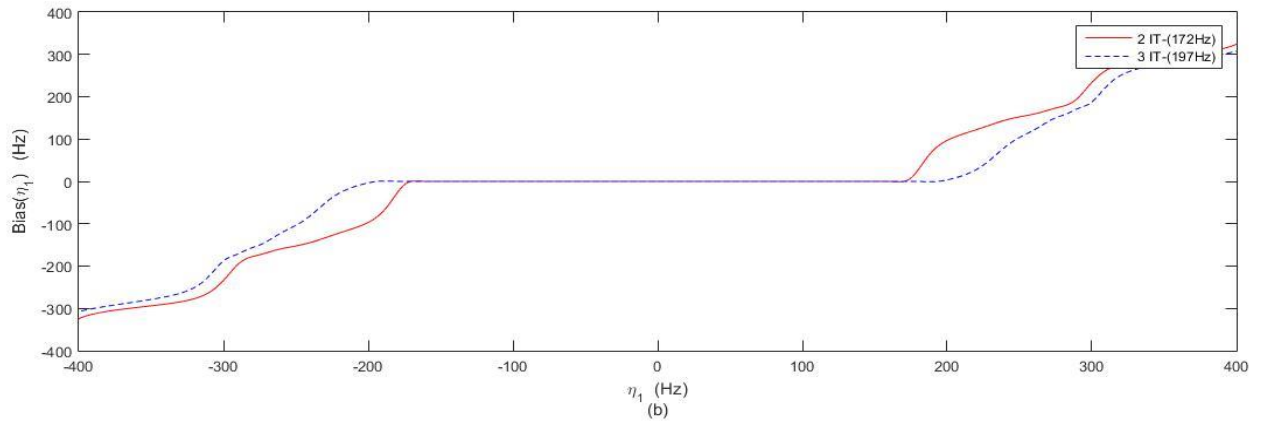
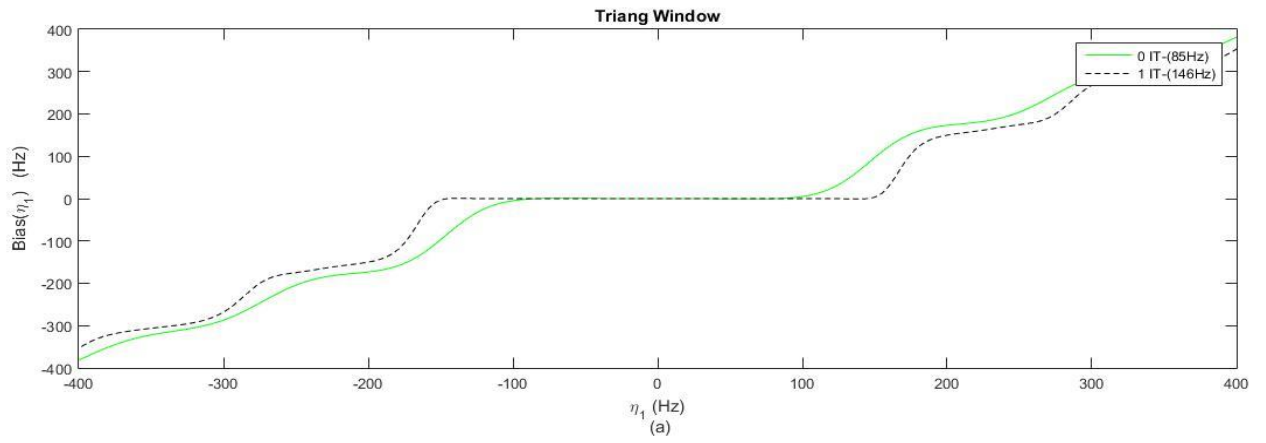
A. Γραφήματα για το bias για τα διάφορα παράθυρα.



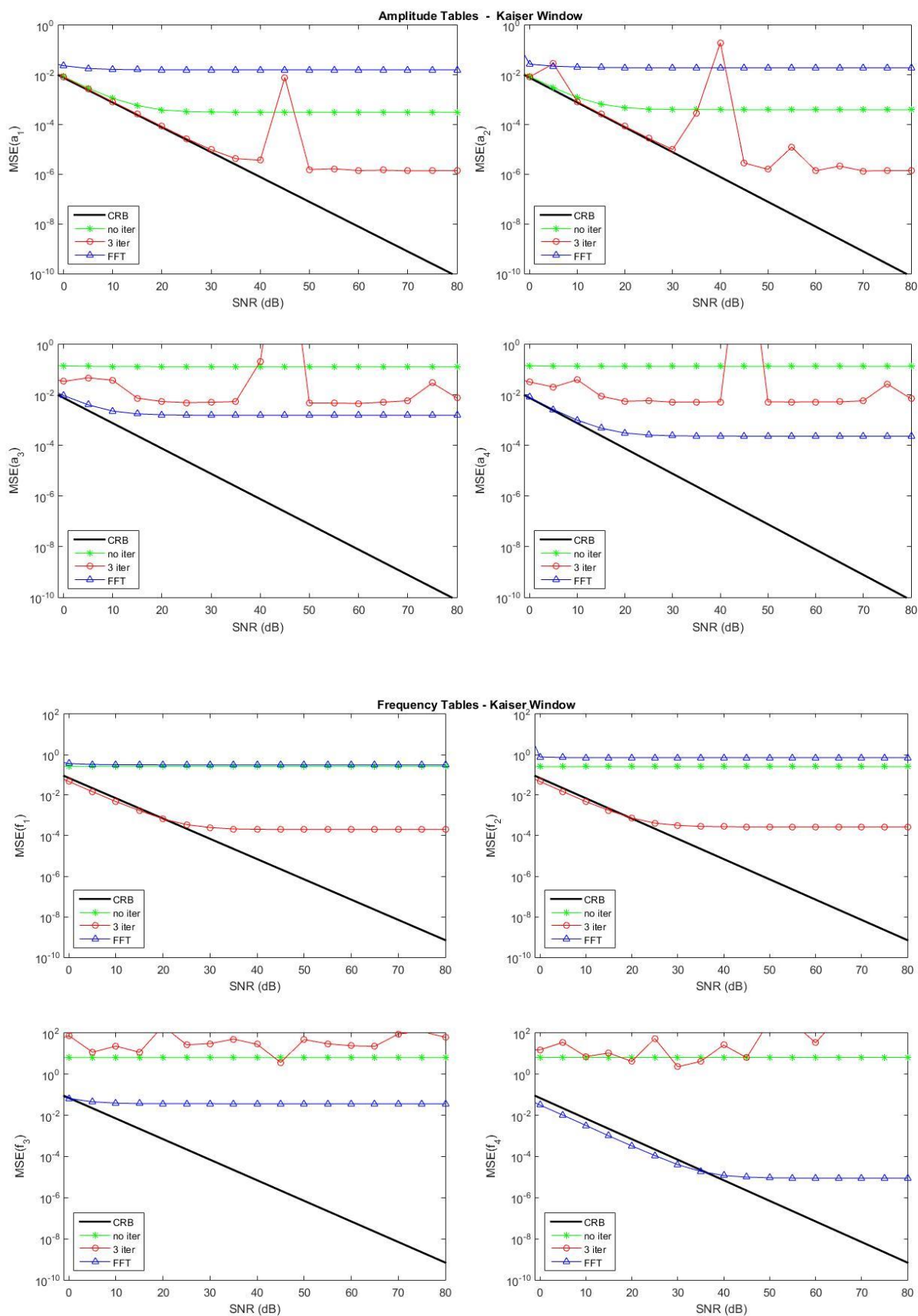




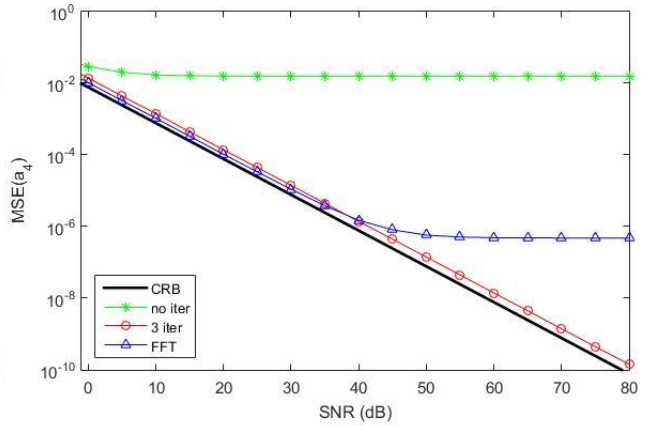
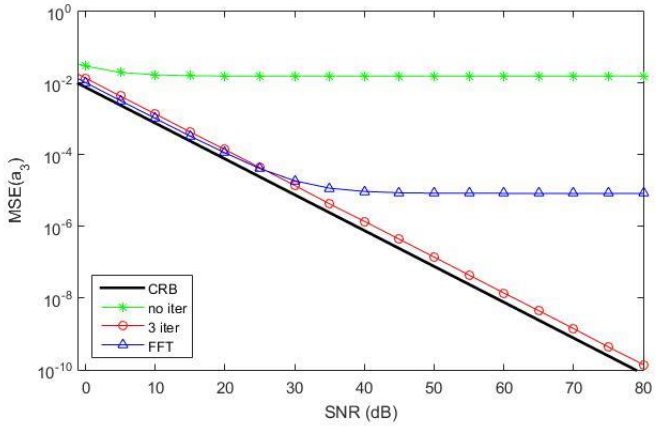
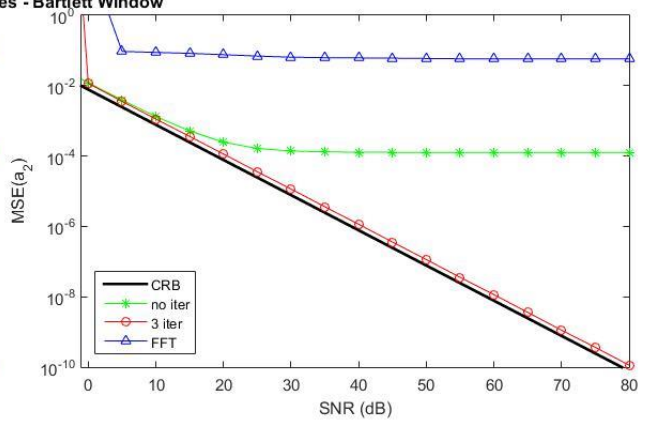
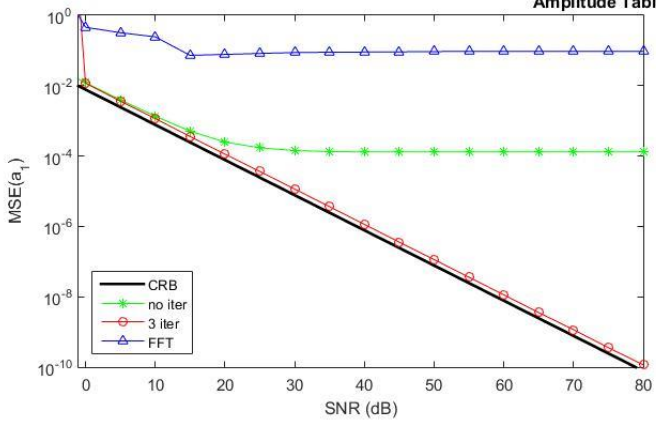




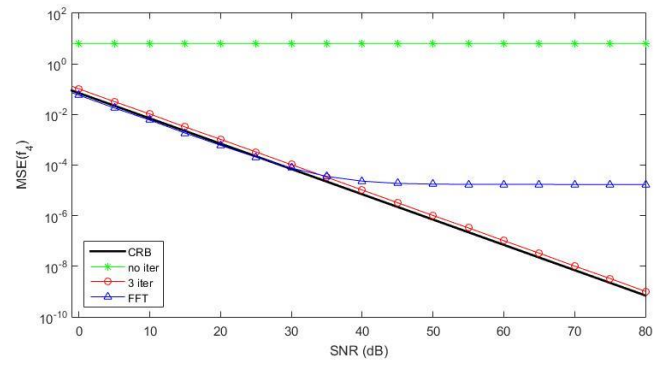
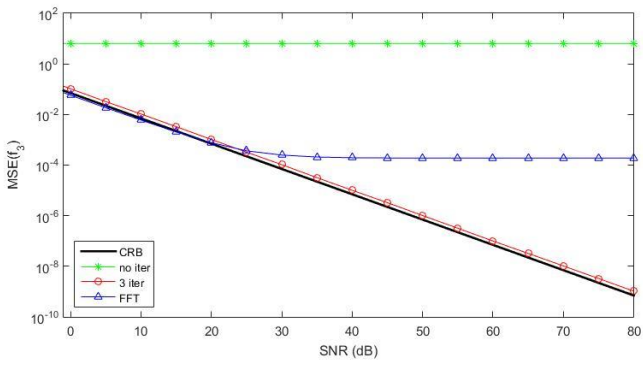
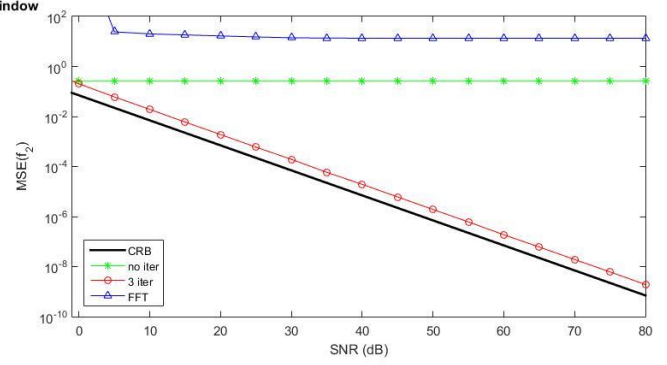
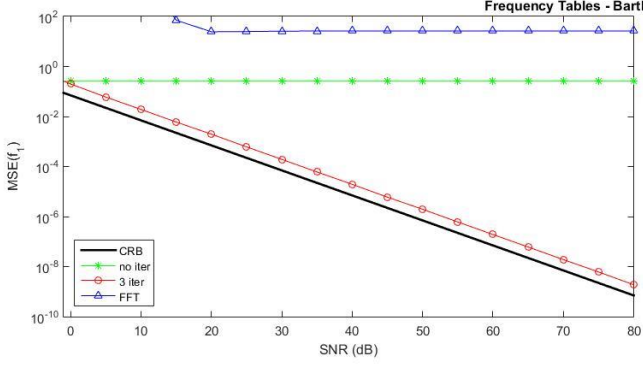
B. Γραφήματα για το MSE πλάτων και συχνοτήτων για τα διάφορα παράθυρα



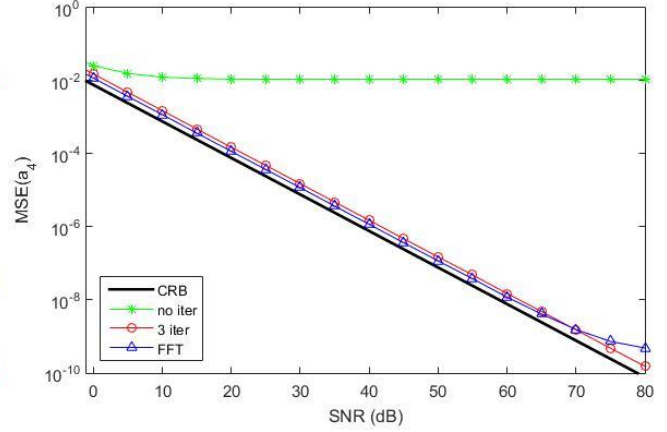
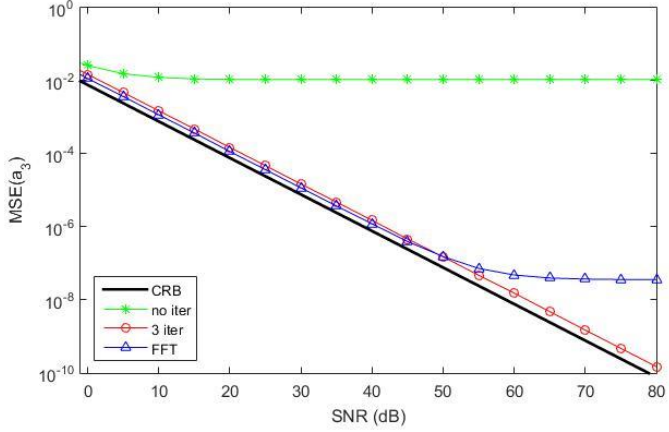
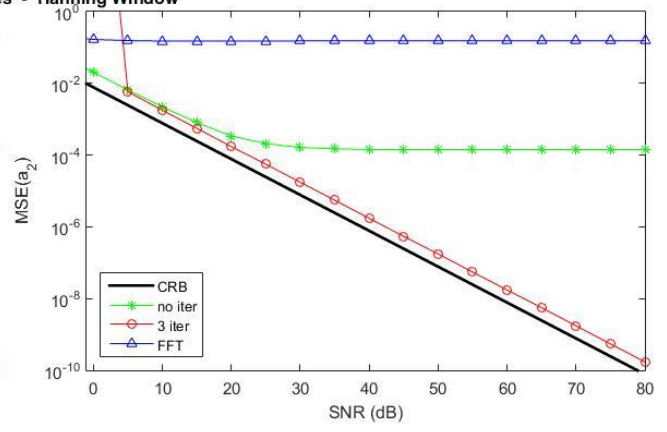
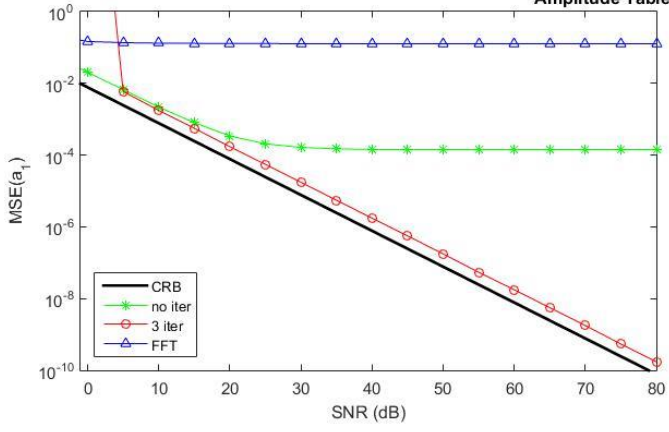
Amplitude Tables - Bartlett Window



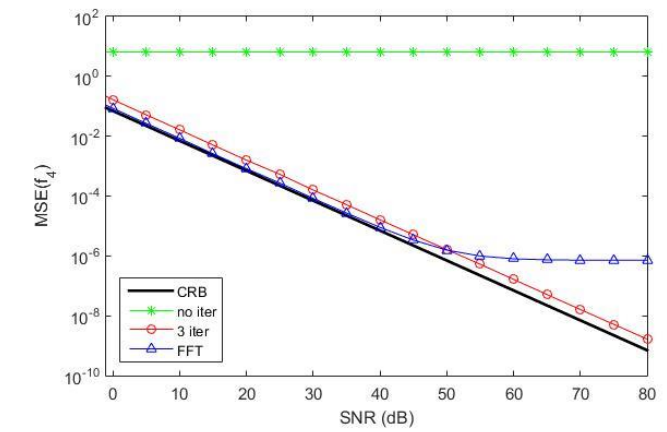
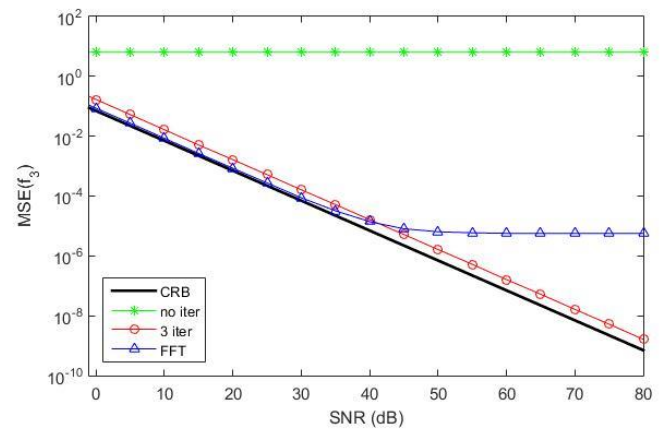
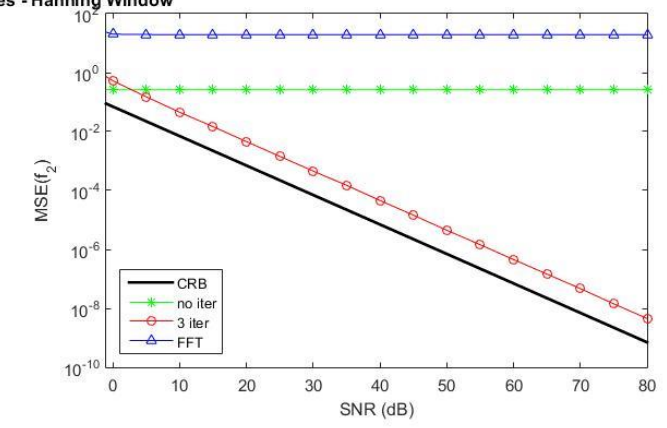
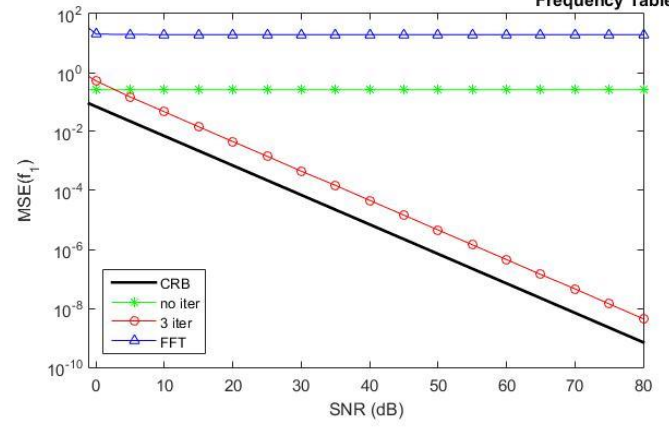
Frequency Tables - Bartlett Window



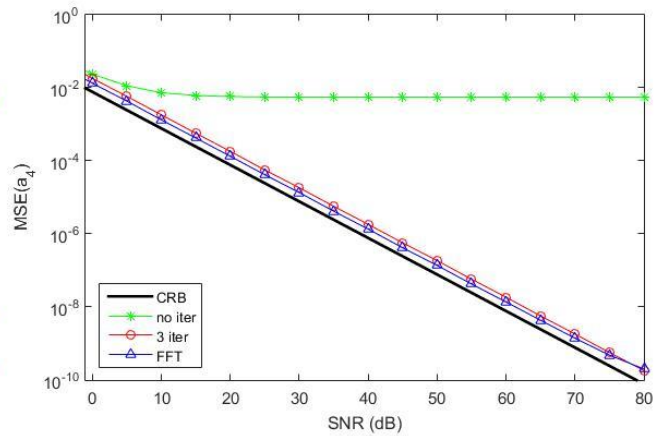
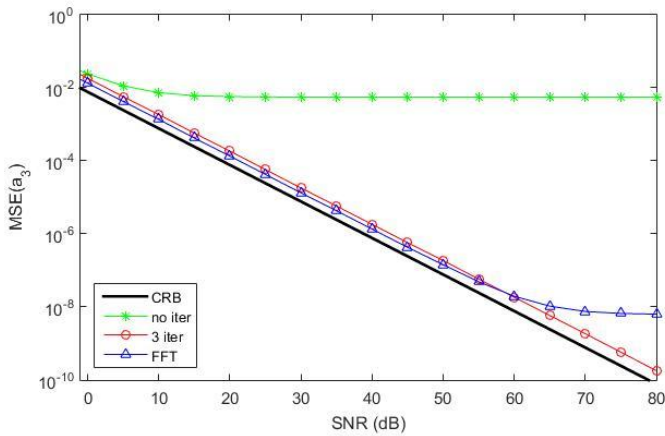
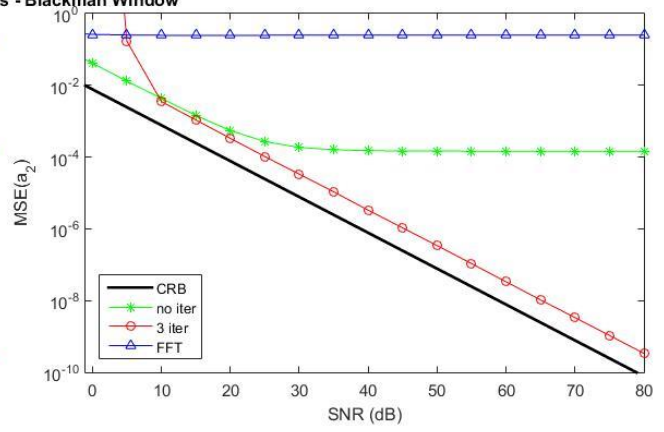
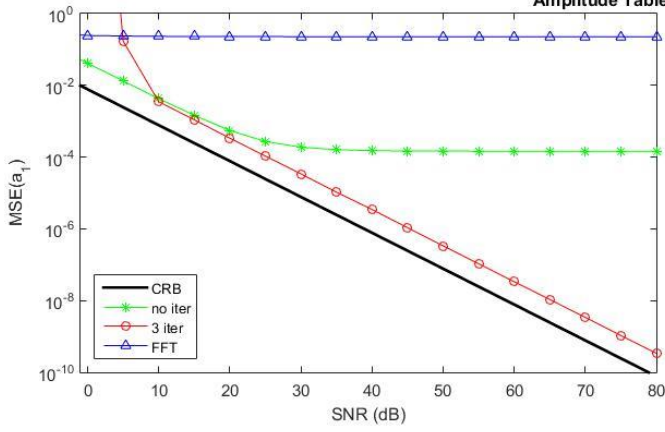
Amplitude Tables - Hanning Window



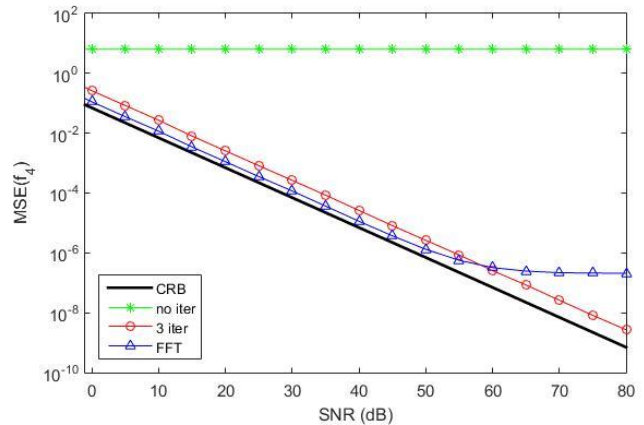
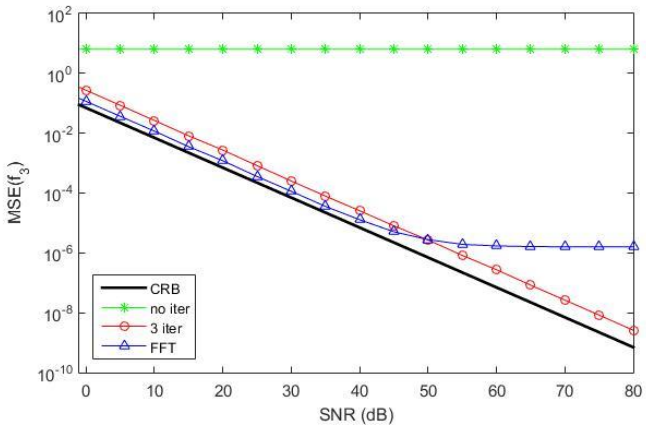
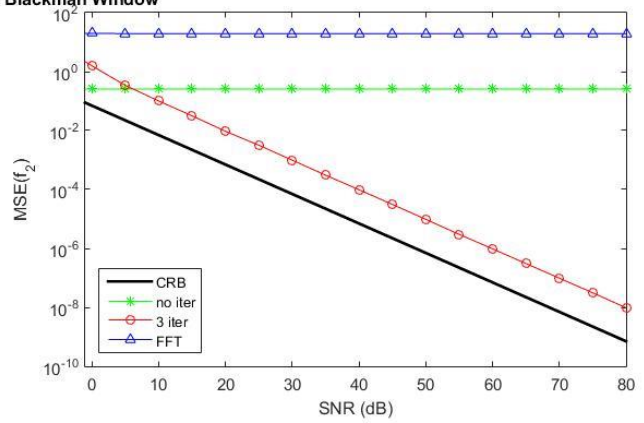
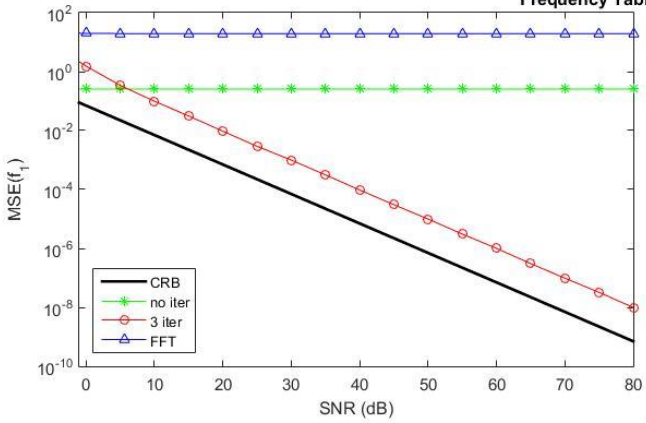
Frequency Tables - Hanning Window



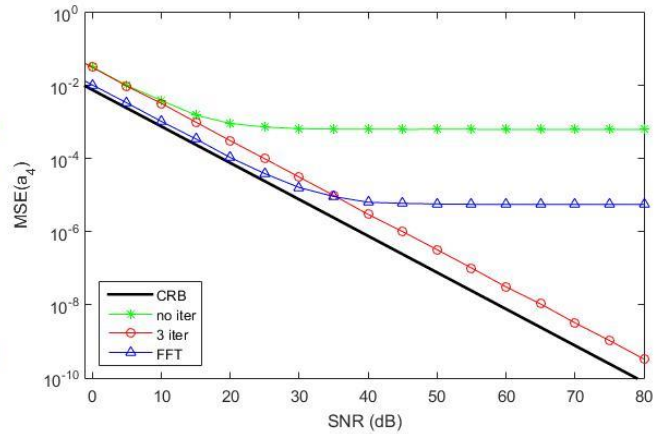
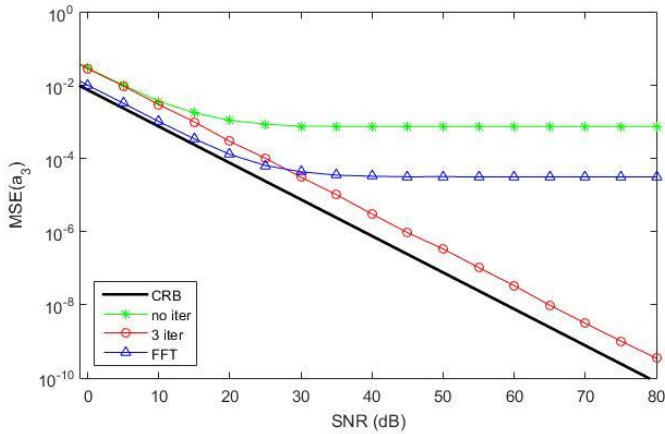
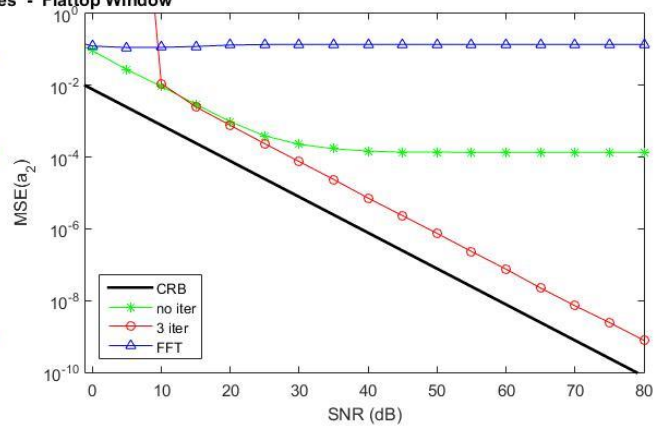
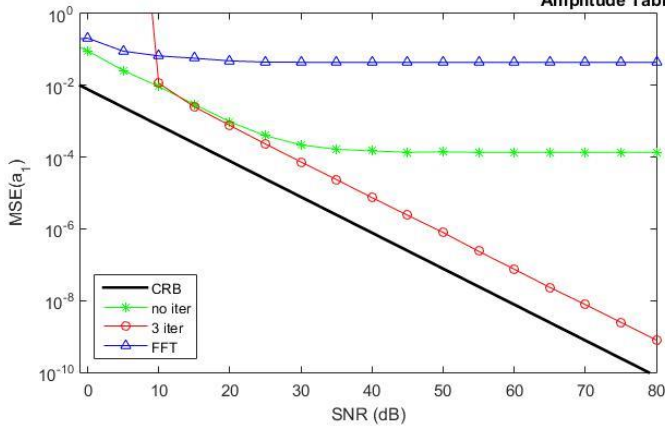
Amplitude Tables - Blackman Window



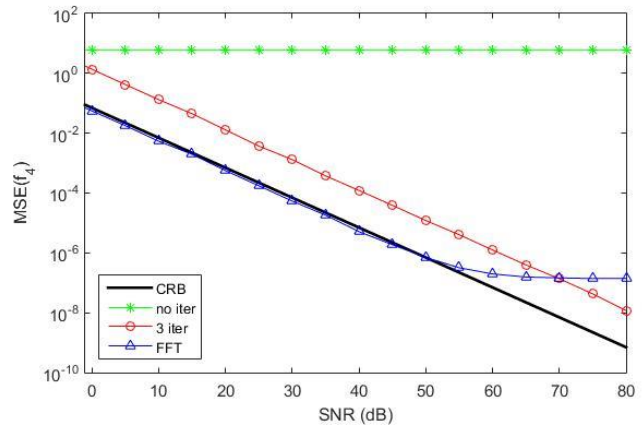
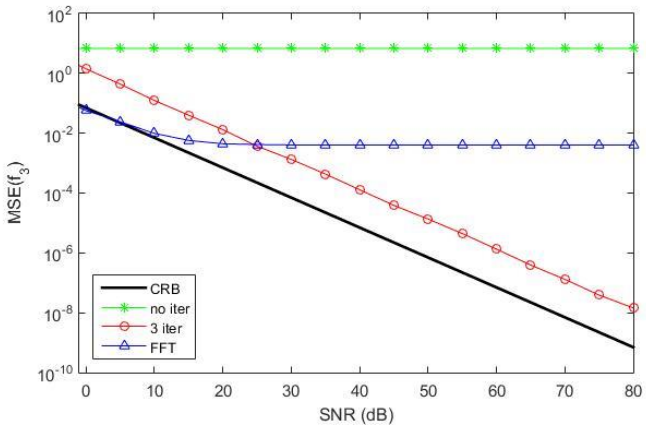
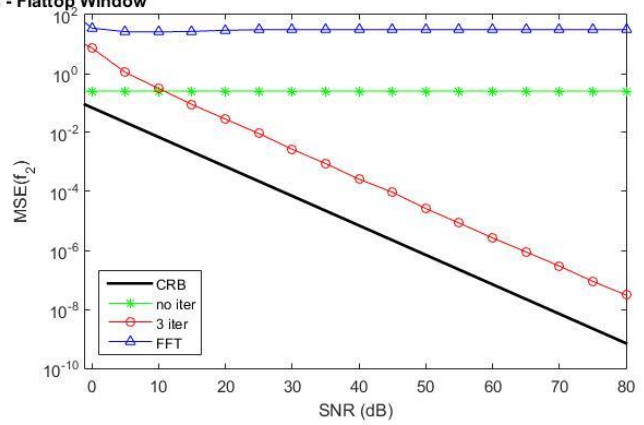
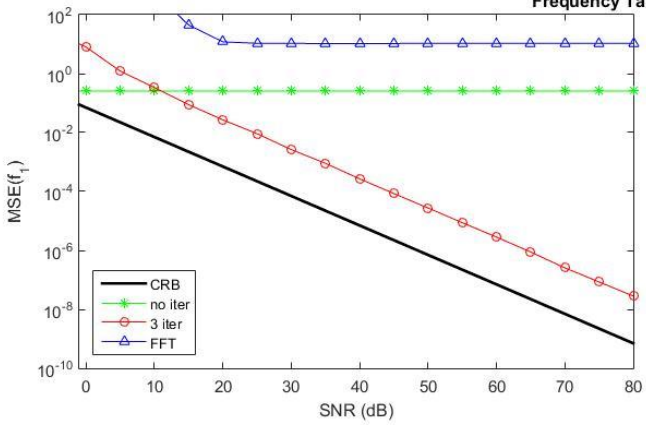
Frequency Tables - Blackman Window



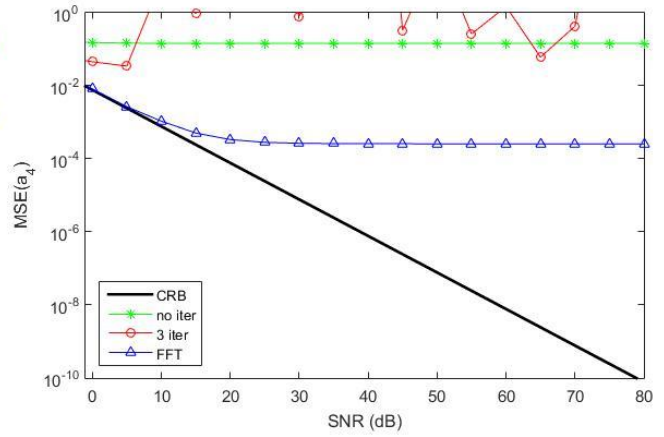
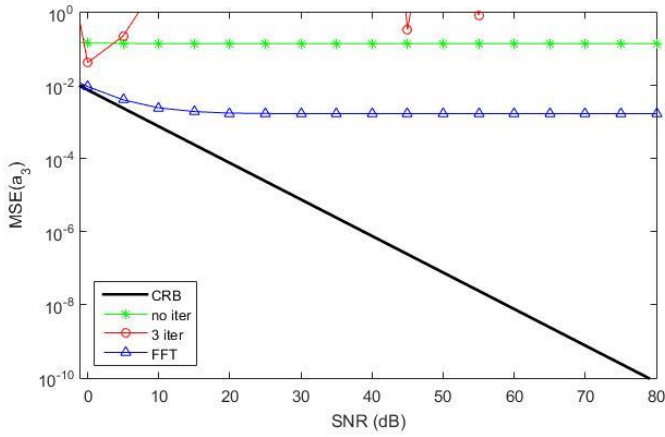
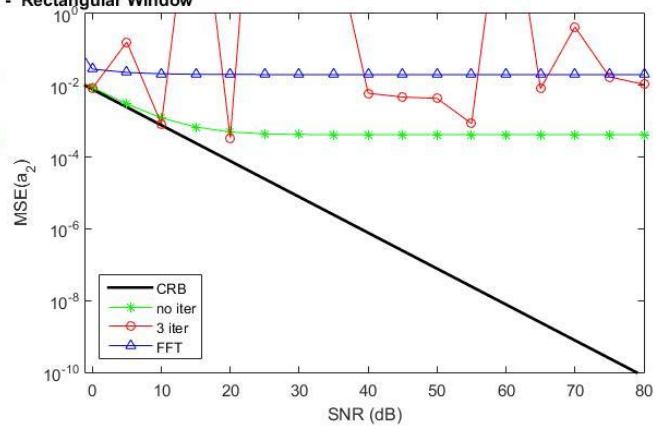
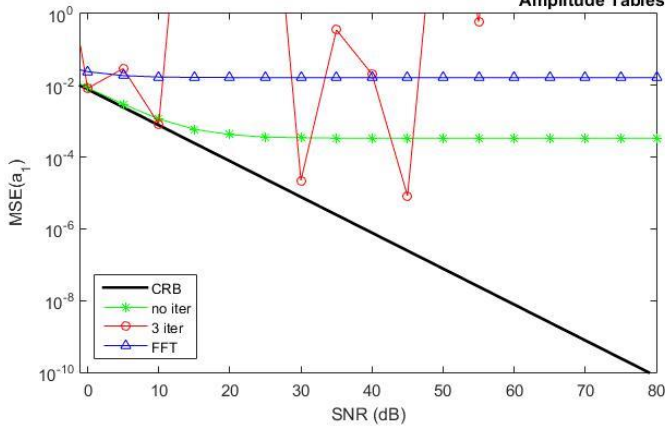
Amplitude Tables - Flattop Window



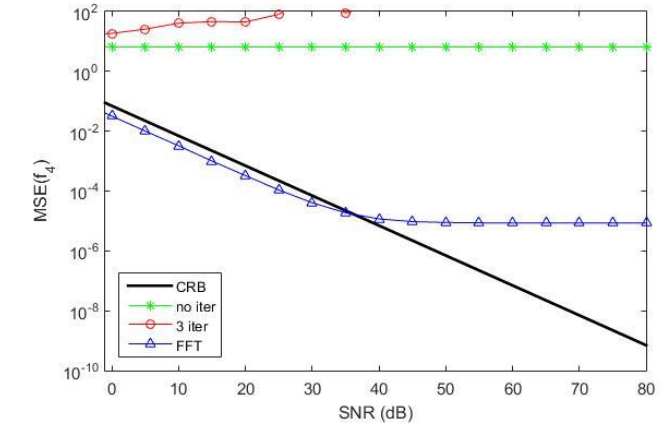
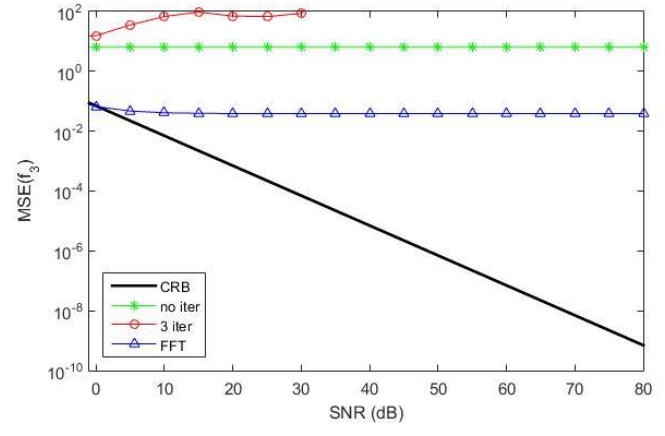
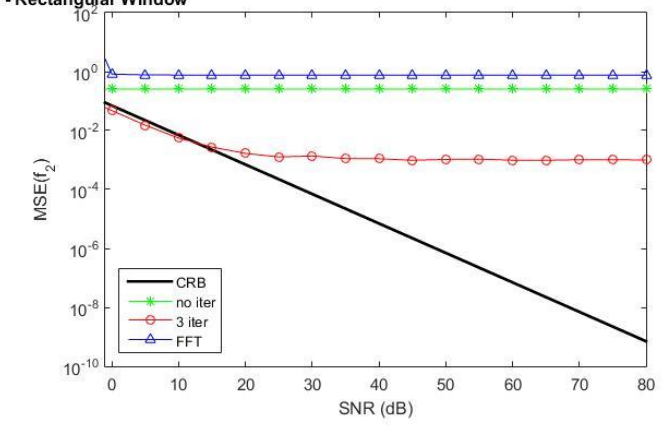
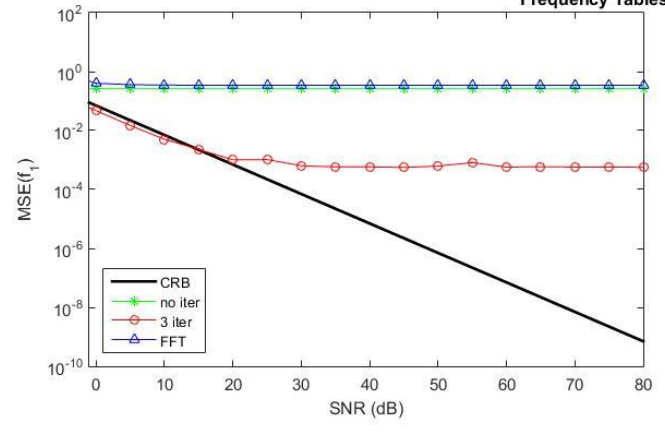
Frequency Tables - Flattop Window



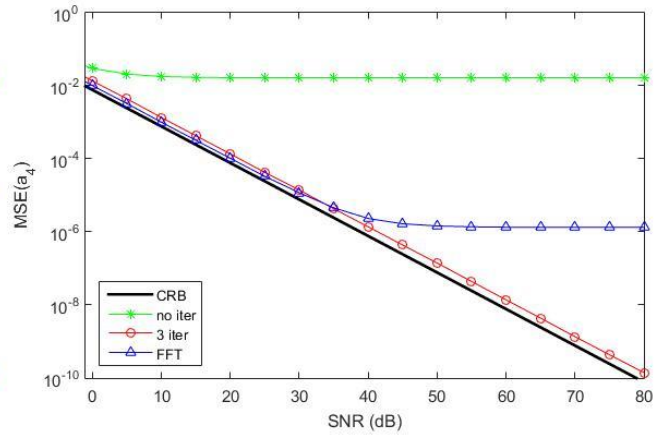
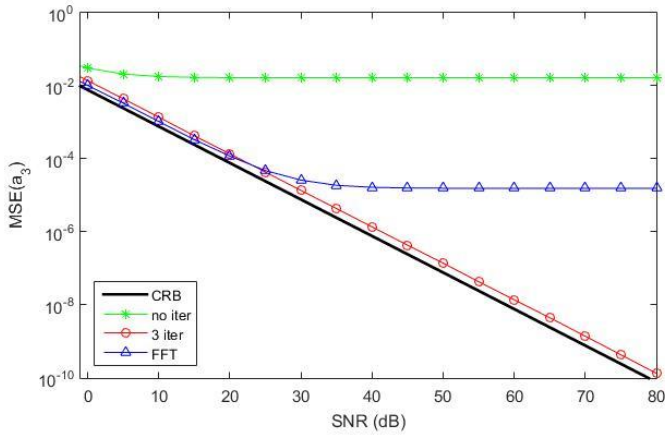
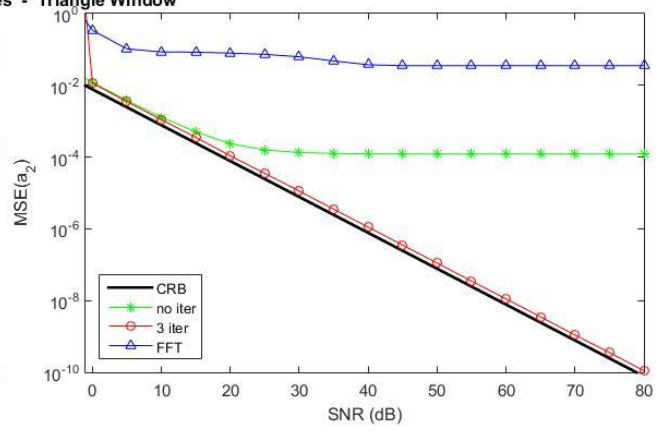
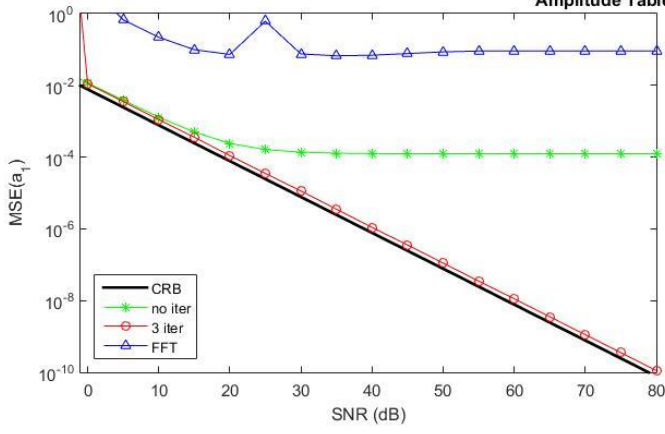
Amplitude Tables - Rectangular Window



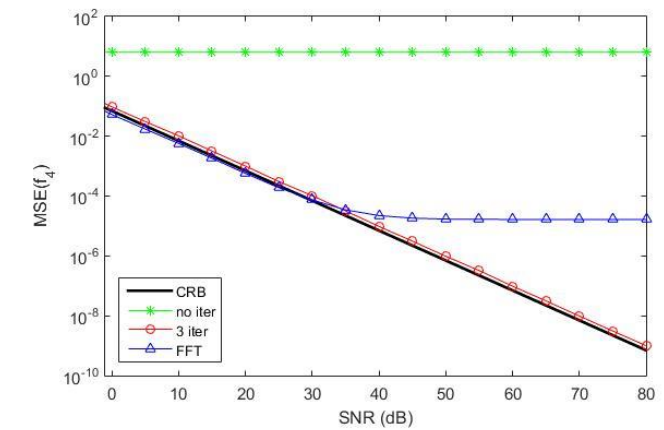
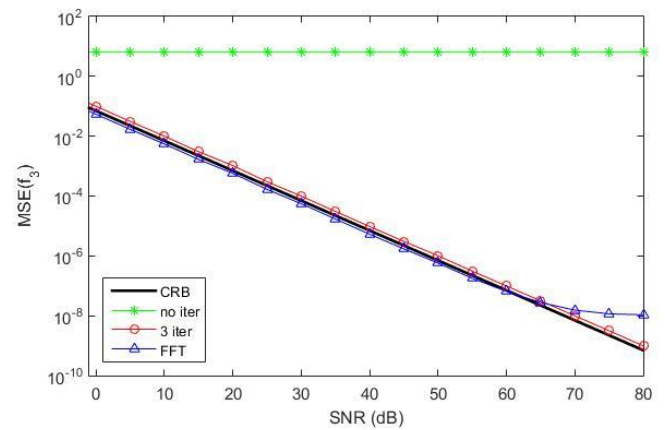
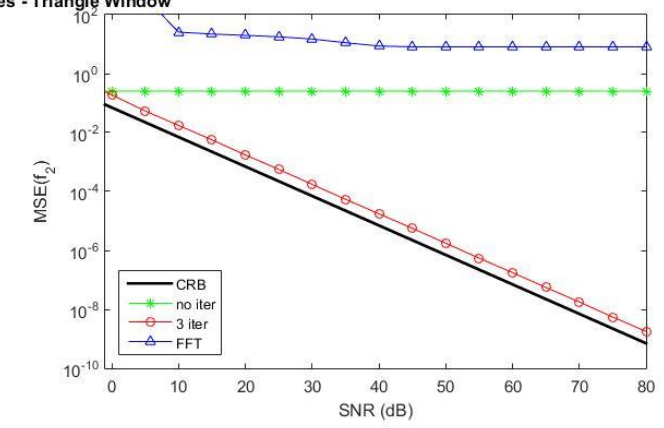
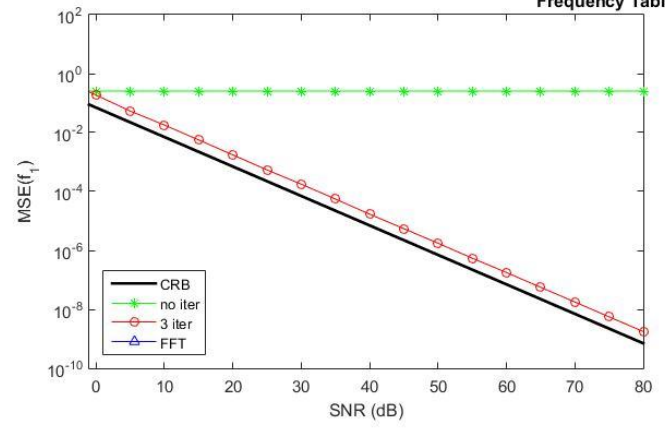
Frequency Tables - Rectangular Window



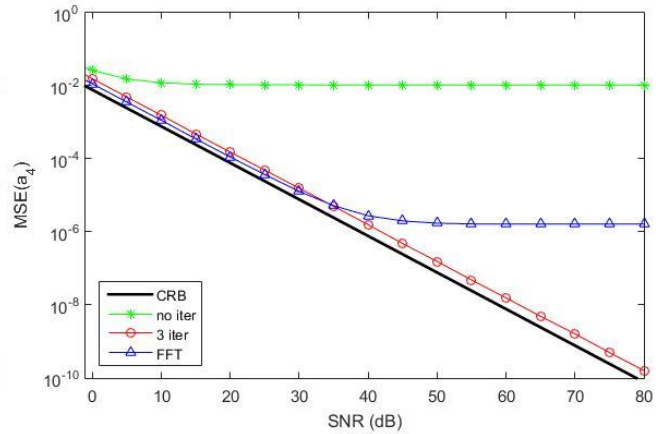
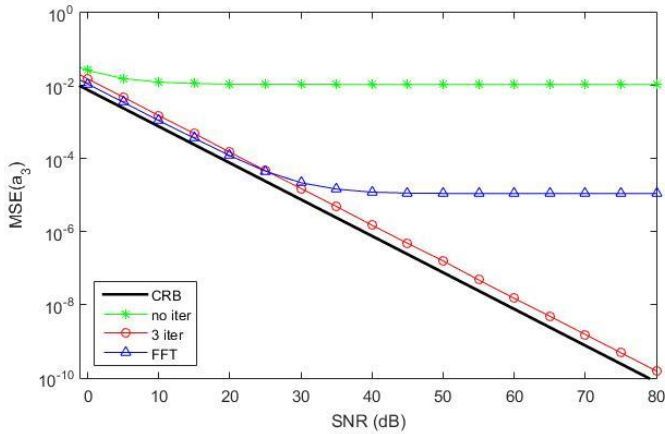
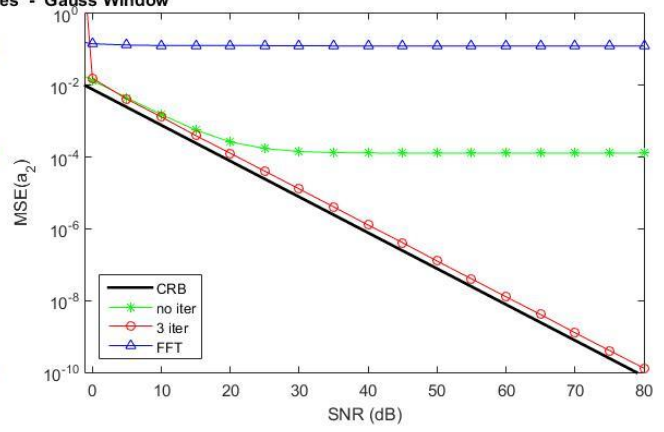
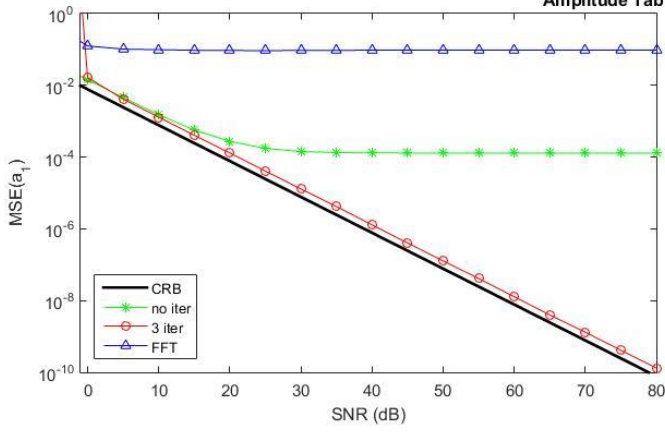
Amplitude Tables - Triangle Window



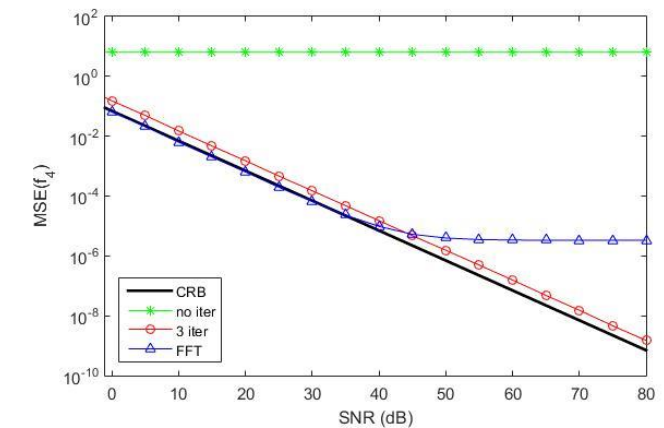
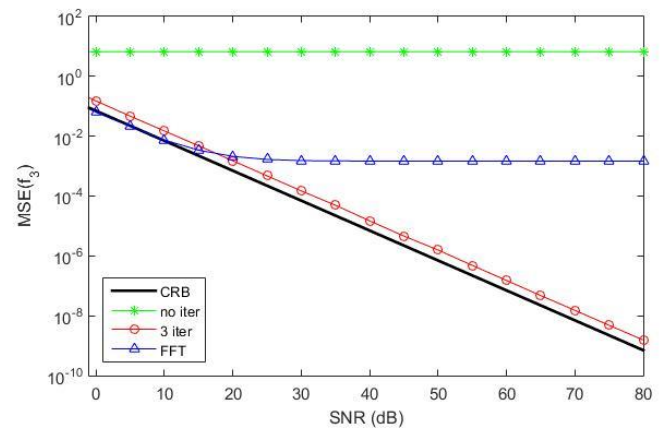
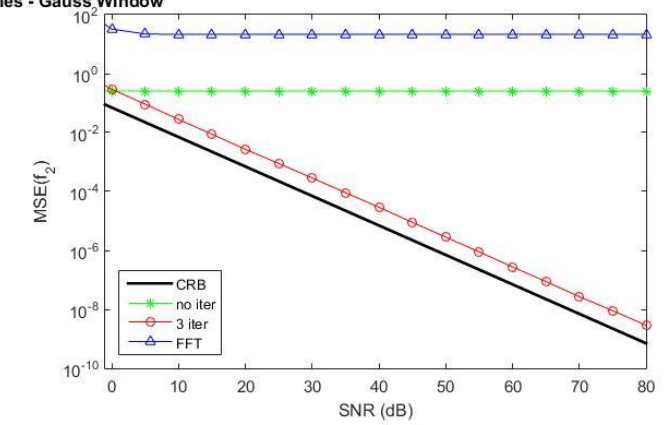
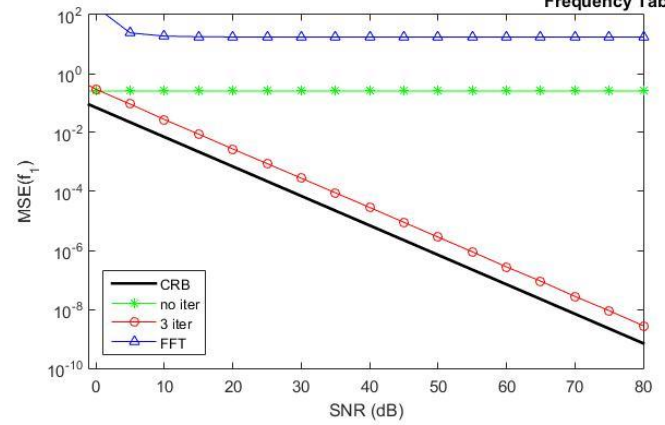
Frequency Tables - Triangle Window



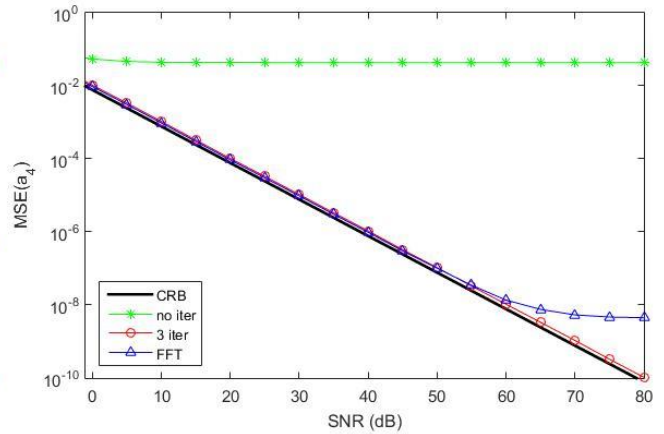
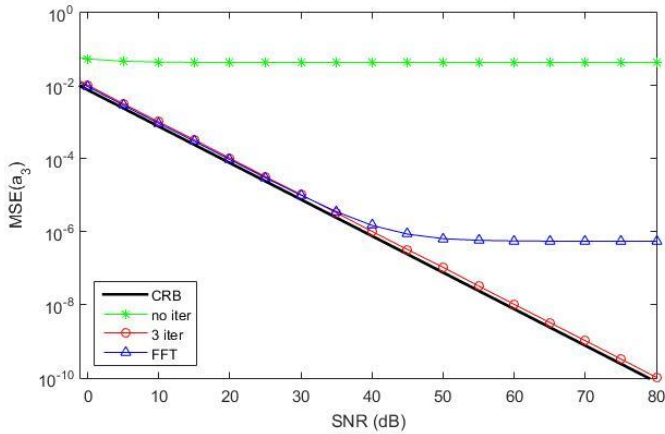
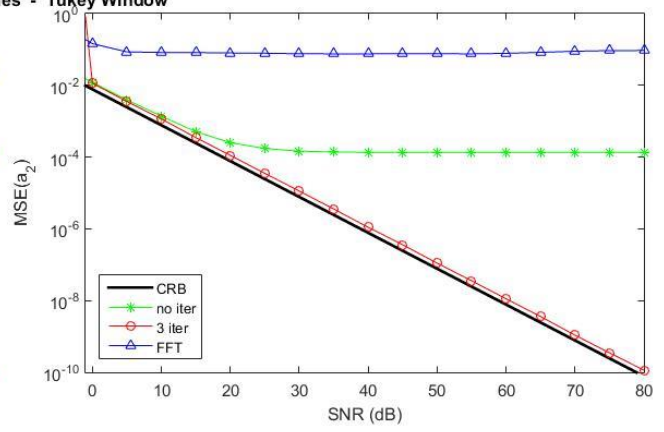
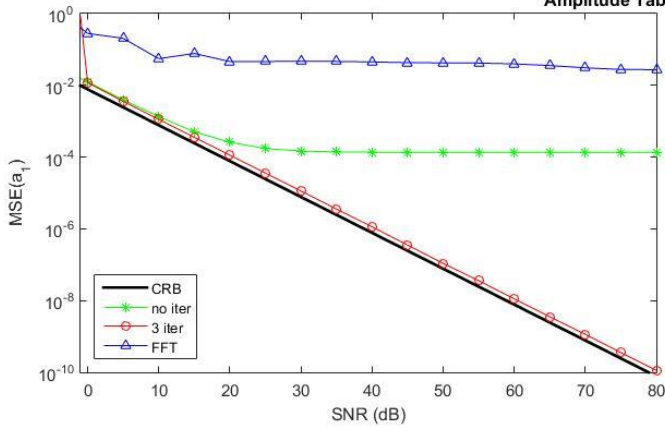
Amplitude Tables - Gauss Window



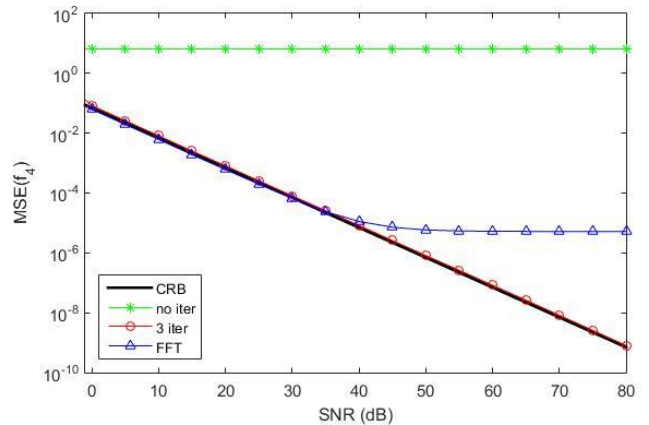
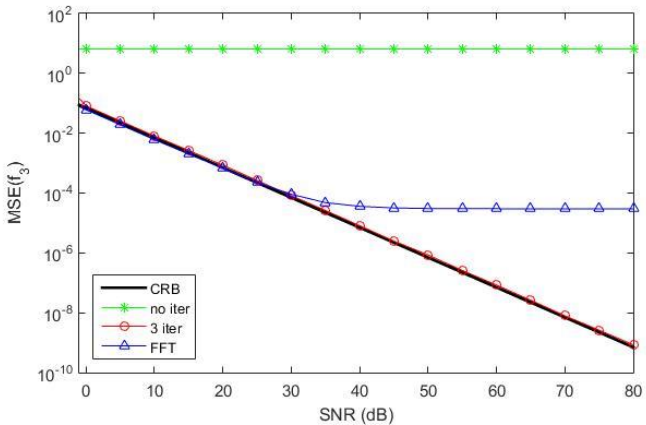
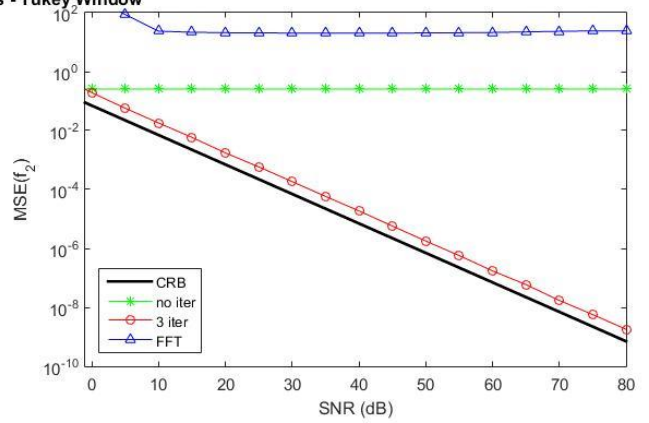
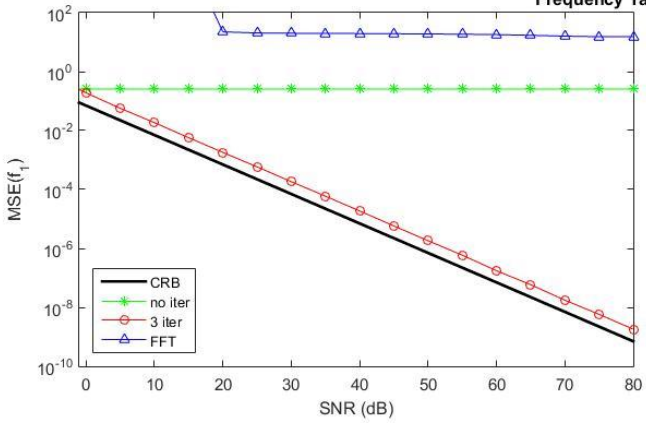
Frequency Tables - Gauss Window



Amplitude Tables - Tukey Window



Frequency Tables - Tukey Window



Γ. Αναλυτικοί πίνακες για MSE ,συχνοτήτων και πλατών, των παραθύρων.

Hamming Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012509	0.016155	0.337786
40 db	1.325161e-04	1.196399e-06	4.378978e-02
80 db	1.313031e-04	1.206083e-10	4.378810e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012470	0.016189	0.126426
40 db	1.336856e-04	1.178389e-06	1.381743e-01
80 db	1.321460e-04	1.168984e-10	1.381668e-01
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.029184	0.013955	0.010549
40 db	1.495862e-02	1.448945e-06	3.318568e-05
80 db	1.495896e-02	1.418096e-10	3.224093e-05
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.028957	0.014115	0.010724
40 db	1.499749e-02	1.417584e-06	6.851507e-06
80 db	1.500043e-02	1.436819e-10	5.691655e-06

Triangle Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.011685	0.011115	6.705646
40 db	1.254940e-04	1.096358e-06	6.920528e-02
80 db	1.242548e-04	1.113932e-10	9.117131e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012099	0.011546	0.343576
40 db	1.231713e-04	1.119448e-06	3.753045e-02
80 db	1.224630e-04	1.104337e-10	3.529418e-02
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.030763	0.013909	0.010303
40 db	1.667907e-02	1.379935e-06	1.667270e-05
80 db	1.668064e-02	1.348815e-10	1.566324e-05
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.030145	0.013632	0.010163
40 db	1.661338e-02	1.383876e-06	2.339853e-06
80 db	1.661267e-02	1.407406e-10	1.345090e-06

Hamming Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.255931	0.221478	5937.67441
40 db	2.559310e-01	2.162156e-05	1.055064e+01
80 db	2.559310e-01	2.155000e-09	1.055236e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.253613	0.218701	89.211355
40 db	2.536131e-01	2.071010e-05	3.087297e+01
80 db	2.536131e-01	2.163350e-09	3.086758e+01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.622879	0.127900	0.059663
40 db	6.622879e+00	1.244945e-05	4.021747e-03
80 db	6.622879e+00	1.251827e-09	4.017271e-03
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.600909	0.122223	0.056704
40 db	6.600909e+00	1.225830e-05	5.687099e-06
80 db	6.600909e+00	1.221952e-09	1.400331e-07

Triangle Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.255193	0.186391	44692.202141
40 db	2.551930e-01	1.698068e-05	1.392608e+02
80 db	2.551930e-01	1.761624e-09	1.766802e+02
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.254928	0.189351	1892.955507
40 db	2.549278e-01	1.747866e-05	8.647731e+00
80 db	2.549278e-01	1.784256e-09	8.105225e+00
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.407853	0.099369	0.053900
40 db	6.407853e+00	9.597603e-06	5.294139e-06
80 db	6.407853e+00	9.835526e-10	1.074127e-08
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.357637	0.095392	0.053035
40 db	6.357637e+00	9.665214e-06	2.222166e-05
80 db	6.357637e+00	9.999754e-10	1.653442e-05

Blackman Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.041095	22182074279.818279	0.245656
40 db	1.508493e-04	3.460692e-06	2.271070e-01
80 db	1.465001e-04	3.396777e-10	2.271184e-01
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.041685	22182075107.637436	0.263508
40 db	1.504560e-04	3.309400e-06	2.500836e-01
80 db	5.582513e-03	3.370298e-10	2.500801e-01
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.024161	0.018320	0.013418
40 db	5.584536e-03	1.834868e-06	1.363392e-06
80 db	1.585631e-02	1.793082e-10	6.368486e-09
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.023726	0.018122	0.013339
40 db	5.560700e-03	1.834764e-06	1.357559e-06
80 db	5.557933e-03	1.812970e-10	2.006815e-10

Bartlett Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012193	0.011802	0.457496
40 db	1.311875e-04	1.141681e-06	9.024885e-02
80 db	1.298293e-04	9.024885e-02	9.495223e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012130	0.011656	47.093451
40 db	1.283464e-04	1.147743e-06	6.151723e-02
80 db	1.272487e-04	1.137391e-10	5.768208e-02
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.030100	0.014074	0.010506
40 db	1.585508e-02	1.407666e-06	9.583098e-06
80 db	1.585631e-02	1.407814e-10	8.480665e-06
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.029544	0.014020	0.010413
40 db	1.578930e-02	1.392414e-06	9.583098e-06
80 db	1.578932e-02	1.432753e-10	4.802046e-07

Blackman Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.256778	1.530648	19.783396
40 db	2.567777e-01	9.605781e-05	1.938510e+01
80 db	2.567777e-01	9.931747e-09	1.938460e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.255990	1.575897	19.790987
40 db	2.559900e-01	9.824376e-05	1.937283e+01
80 db	2.559900e-01	9.651358e-09	1.937325e+01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.438390	0.272092	0.115903
40 db	6.438390e+00	2.560065e-05	1.285393e-05
80 db	6.438390e+00	2.606010e-09	1.591505e-06
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.428210	0.262369	0.112080
40 db	6.428210e+00	2.625697e-05	1.141844e-05
80 db	6.428210e+00	2.700507e-09	2.075887e-07

Bartlett Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.258944	0.199765	8114.438271
40 db	2.589444e-01	1.949448e-05	2.631518e+01
80 db	2.589444e-01	1.878905e-09	2.643969e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.256325	0.202876	142098.610848
40 db	2.563255e-01	1.943853e-05	1.351095e+01
80 db	2.563255e-01	1.923064e-09	1.341278e+01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.399321	0.103323	0.057685
40 db	6.399321e+00	1.010485e-05	1.956416e-04
80 db	6.399321e+00	1.040777e-09	1.896538e-04
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.376947	0.102691	0.057898
40 db	6.376947e+00	1.036738e-05	2.320689e-05
80 db	6.376947e+00	1.016384e-09	1.715534e-05

Rectangular Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.008585	0.008278	0.024023
40 db	3.370637e-04	2.047456e-02	1.642509e-02
80 db	3.362010e-04	1.367697e+06	1.642209e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.008741	0.008325	0.028265
40 db	4.200655e-04	6.026974e-03	2.000890e-02
80 db	4.193209e-04	1.061452e-02	2.000899e-02
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.150472	0.043361	0.009582
40 db	1.426915e-01	2.391098e+09	1.724254e-03
80 db	1.426818e-01	1.367681e+06	1.723810e-03
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.153585	0.045494	0.008289
40 db	1.451754e-01	2.391096e+09	2.600352e-04
80 db	1.451764e-01	8.614397e+00	2.592633e-04

Taylor Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.009963	0.009795	0.021520
40 db	1.605641e-04	9.735083e-07	2.105946e-06
80 db	1.593143e-04	9.928390e-11	1.181964e-06
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.009772	0.009535	38.438166
40 db	1.933080e-04	9.747969e-07	1.783790e-04
80 db	1.920788e-04	9.756466e-11	1.775989e-04
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.041136	0.011492	0.009390
40 db	2.971147e-02	1.158377e-06	2.227669e-04
80 db	2.971159e-02	1.163816e-10	2.218042e-04
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.040530	0.011522	0.009179
40 db	2.951242e-02	1.137715e-06	3.772363e-05
80 db	2.951178e-02	1.166253e-10	3.687407e-05

Rectangular Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.261678	0.047339	0.401759
40 db	2.616782e-01	5.843994e-04	3.428483e-01
80 db	2.616782e-01	5.659118e-04	3.428218e-01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.257793	0.047729	0.819886
40 db	2.577935e-01	1.103210e-03	7.586330e-01
80 db	2.577935e-01	1.006923e-03	7.586759e-01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.410848	15.591482	0.066920
40 db	6.410848e+00	3.335609e+02	3.871841e-02
80 db	6.410848e+00	1.257115e+02	3.870948e-02
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.470544	18.557721	0.032023
40 db	6.470544e+00	1.792602e+02	1.169233e-05
80 db	6.470544e+00	5.183324e+02	8.567178e-06

Taylor Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.258386	0.081603	22.290992
40 db	2.583864e-01	8.190190e-06	3.826595e-01
80 db	2.583864e-01	7.863043e-10	3.825849e-01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.256911	0.081970	57468.485194
40 db	2.569110e-01	8.053962e-06	1.721775e-01
80 db	2.569110e-01	7.949404e-10	1.721146e-01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.412898	0.067945	0.052899
40 db	6.412898e+00	6.667732e-06	1.421243e-02
80 db	6.412898e+00	6.720166e-10	1.421544e-02
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.395529	0.067700	0.040158
40 db	6.395529e+00	6.673516e-06	6.168587e-06
80 db	6.395529e+00	6.655521e-10	2.256019e-06

Hanning Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.020349	26013519.820706	0.146800
40 db	1.438260e-04	1.745801e-06	1.538925e-01
80 db	1.419360e-04	1.734068e-10	1.260006e-01
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.020489	26013522.819584	0.167216
40 db	1.442130e-04	1.722558e-06	1.381743e-01
80 db	1.420581e-04	1.714066e-10	1.538901e-01
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.026222	0.015034	0.011918
40 db	1.083652e-02	1.520751e-06	1.181605e-06
80 db	1.083344e-02	1.181605e-06	3.620341e-08
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.026169	0.015459	0.011918
40 db	1.105429e-02	1.526023e-06	1.181605e-06
80 db	1.105402e-02	1.535910e-10	4.838243e-10

Kaiser Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.008651	0.008352	0.023468
40 db	3.129127e-04	3.684482e-06	1.600359e-02
80 db	3.121447e-04	1.405362e-06	1.600175e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.008710	0.008371	0.027411
40 db	3.975758e-04	1.980364e-01	1.954328e-02
80 db	3.970749e-04	1.404618e-06	1.954194e-02
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.142597	0.034684	0.009606
40 db	1.349220e-01	2.119801e-01	1.583574e-03
80 db	1.349139e-01	7.667847e-03	1.583045e-03
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.144183	0.032954	0.008258
40 db	1.365713e-01	5.375226e-03	2.387116e-04
80 db	1.365751e-01	7.171345e-03	2.377958e-04

Hanning Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.260237	0.523530	20.229720
40 db	2.602368e-01	4.624050e-05	1.939774e+01
80 db	2.602368e-01	4.432685e-09	1.939692e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.259429	0.529045	20.117382
40 db	2.594291e-01	4.569454e-05	1.936027e+01
80 db	2.594291e-01	4.532072e-09	1.936093e+01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.422460	0.168344	0.086030
40 db	6.422460e+00	1.631150e-05	1.393912e-05
80 db	6.422460e+00	1.627729e-09	5.95458e-06
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.508980	0.167101	0.085758
40 db	6.508980e+00	1.612839e-05	8.994067e-06
80 db	6.508980e+00	1.658764e-09	7.152723e-07

Kaiser Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.263195	0.047082	0.365068
40 db	2.631947e-01	2.061832e-04	3.121085e-01
80 db	2.631947e-01	2.006732e-04	1.938460e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.258159	0.047304	0.759250
40 db	2.581594e-01	2.784938e-04	7.066263e-01
80 db	2.581594e-01	2.720419e-04	7.066481e-01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.425156	75.605078	0.066798
40 db	6.425156e+00	2.853542e+01	3.688130e-02
80 db	6.425156e+00	6.265124e+01	3.687059e-02
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.463355	15.221117	0.031915
40 db	6.463355e+00	2.629965e+01	1.193115e-05
80 db	6.463355e+00	4.462195e+03	8.952911e-06

Gauss Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.014349	0.016579	0.125133
40 db	1.306500e-04	1.291236e-06	9.529606e-02
80 db	1.289635e-04	1.311091e-10	9.532270e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.013696	0.015765	0.263508
40 db	1.301875e-04	1.279086e-06	1.249372e-01
80 db	1.287318e-04	1.306537e-10	1.248817e-01
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.026425	0.026425	0.013418
40 db	1.098944e-02	1.530451e-06	1.228480e-05
80 db	1.098558e-02	1.565913e-10	1.123055e-05
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.026599	0.015673	0.013339
40 db	1.060153e-02	1.580379e-06	2.771300e-06
80 db	1.059896e-02	1.597583e-10	1.633455e-06

Tukey Window - Amplitude MSE

1st Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012511	0.011925	0.280064
40 db	1.366618e-04	1.142854e-06	7.561294e-02
80 db	1.352705e-04	1.133406e-10	2.732314e-02
2nd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.012102	0.011517	0.145454
40 db	1.359636e-04	1.106696e-06	3.753045e-02
80 db	1.345496e-04	1.148236e-10	9.349646e-02
3rd Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.054670	0.010568	0.009636
40 db	4.366215e-02	1.036039e-06	1.511913e-06
80 db	4.365632e-02	1.044472e-10	5.581006e-07
4th Amplitude	No iteration QHM	2 iterations QHM	FFT
0 db	0.054387	0.010414	0.009491
40 db	4.342570e-02	1.068479e-06	9.771926e-07
80 db	4.342115e-02	1.044184e-10	4.488067e-09

Gauss Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.256720	0.291904	180.610002
40 db	2.567197e-01	2.876941e-05	1.738426e+01
80 db	2.567197e-01	2.777529e-09	1.738281e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.255441	0.295418	32.423938
40 db	2.554411e-01	2.907362e-05	2.148365e+01
80 db	2.554411e-01	2.867623e-09	2.148351e+01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.618723	0.150339	0.062718
40 db	6.618723e+00	1.481595e-05	1.476275e-03
80 db	6.618723e+00	1.540920e-09	1.469108e-03
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.407650	0.152781	0.063961
40 db	6.407650e+00	1.508443e-05	9.692734e-06
80 db	6.407650e+00	1.495498e-09	3.321913e-06

Tukey Window - Frequency MSE

1st Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.257488	0.190454	1861.735832
40 db	2.574882e-01	1.867011e-05	1.954839e+01
80 db	2.574882e-01	1.806925e-09	1.516078e+01
2nd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	0.256030	0.192748	277.687900
40 db	2.560300e-01	1.882587e-05	2.012679e+01
80 db	2.560300e-01	1.798435e-09	2.411420e+01
3rd Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.517178	0.081760	0.060633
40 db	6.517178e+00	8.146874e-06	3.602993e-05
80 db	6.517178e+00	8.499266e-10	2.980023e-05
4th Frequency	No iteration QHM	2 iterations QHM	FFT
0 db	6.481153	0.084525	0.062389
40 db	6.481153e+00	8.264295e-06	1.128411e-05
80 db	6.481153e+00	8.082676e-10	5.211781e-06

Δ. Αναλυτικοί πίνακες με τα SRER για κάθε γλώσσα με εφαρμογή παραθύρων Blackman-Hamming και Blackman-Taylor.

Blackman - Hamming

ARABIC M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.6932	15.0725	12.8301	11.3392	9.9540
SRER	28.0550	21.8450	16.7621	14.5026	12.7792
SEGM. SRER	34.1950(11.4)	26.1103(10.4)	20.7451(10.0)	17.1416(9.1)	14.7688(8.8)
ADAPTATIONS	4	2	2	2	4

ARABIC FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	11.5296	7.7594	5.8481	4.7250	4.0388
SRER	17.2751	10.8153	7.6038	6.8267	5.6629
SEGM. SRER	26.9311(9.5)	19.5769(9.2)	15.9928(8.9)	14.0032(8.6)	12.7024(8.4)
ADAPTATIONS	2	2	3	2	2

CHINESE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	11.1583	5.9478	4.0669	2.9266	2.2163
SRER	15.2559	8.6196	7.0451	5.2838	4.7976
SEGM. SRER	20.1710(9.1)	15.4246(9.1)	12.0985(8.0)	10.5932(7.8)	8.7568(7.0)
ADAPTATIONS	1	3	3	3	1

CHINESE FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	13.7784	10.7256	9.0776	7.9545	7.3055
SRER	17.6940	13.1570	11.3995	10.2314	9.5487
SEGM. SRER	21.1882(9.4)	14.4186(7.9)	11.9325(7.2)	10.9052(7.0)	10.1505(6.5)
ADAPTATIONS	4	3	2	3	3

KOREAN M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.2698	16.0962	13.3686	10.9963	10.0026
SRER	30.1992	20.7602	15.7514	14.1865	12.6278
SEGM. SRER	29.9223(12.4)	22.9751(10.5)	16.5877(8.6)	14.6750(8.3)	13.0547(7.9)
ADAPTATIONS	5	4	1	2	4

KOREAN FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	16.3813	13.6791	11.9463	10.6016	9.8648
SRER	19.6649	15.8901	14.0170	13.1582	12.1848
SEGM. SRER	22.7647(10.1)	18.1248(9.6)	16.0239(9.8)	15.0253(9.8)	13.9730(9.3)
ADAPTATIONS	2	2	2	2	2

RUSSIAN M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.2448	17.0169	14.8653	13.3074	11.9451
SRER	28.6582	23.4310	19.9832	16.6781	15.1515
SEGM. SRER	29.0044(12.2)	24.7828(12.9)	18.9904(11.2)	14.9846(9.7)	13.0160(9.2)
ADAPTATIONS	2	4	6 (!!!)	2	2

RUSSIAN FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	15.8164	12.2771	10.4241	8.9532	7.8970
SRER	21.1500	14.3443	12.5362	11.6556	10.0458
SEGM. SRER	21.5425(11.1)	15.2996(9.8)	12.9986(9.3)	11.6971(8.9)	11.2653(9.0)
ADAPTATIONS	3	3	2	2	2

Blackman - Taylor

ARABIC M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.6932	15.0725	12.8301	11.3392	9.9540
SRER	27.6041(-0.4)	23.5197(-1.7)	18.0177(+1.3)	14.8547(+0.3)	13.6105(+0.9)
SEGM. SRER	30.0819(10.0)	26.9435(10.4)	21.8489(10.2)	16.9836(8.8)	15.0689(8.6)
ADAPTATIONS	2	6	6	2	3

ARABIC FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	11.5296	7.7594	5.8481	4.7250	4.0388
SRER	20.3642(+3.1)	11.6605(+0.8)	8.4664(+0.8)	6.9619(+0.1)	6.1072(+0.5)
SEGM. SRER	28.1294(9.1)	19.9021(9.0)	16.3678(8.8)	14.2144(8.5)	13.0350(8.4)
ADAPTATIONS	7	3	4	2	2

CHINESE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	11.1583	5.9478	4.0669	2.9266	2.2163
SRER	19.0204(+3.8)	8.6197	7.8037(+0.8)	6.2309(+1.0)	5.6902(+0.9)
SEGM. SRER	22.7638(9.9)	14.7474(8.6)	12.4258(7.9)	10.4218(7.5)	9.8042(7.6)
ADAPTATIONS	4	2	5	2	3

CHINESE FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	13.7784	10.7256	9.0776	7.9545	7.3055
SRER	18.9803(+1.3)	13.9838(+0.8)	11.9988(+0.6)	11.0515(+0.8)	10.1735(+0.6)
SEGM. SRER	21.4342(9.1)	14.8347(7.7)	12.4685(7.4)	11.3053(6.8)	10.3718(6.7)
ADAPTATIONS	7	5	5	3	5

KOREAN M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.2698	16.0962	13.3686	10.9963	10.0026
SRER	29.5945(+0.6)	20.9391(+0.2)	16.7196(+1.0)	14.9557(+0.8)	13.3390(+0.7)
SEGM. SRER	27.6001(10.9)	22.0059(9.8)	17.2500(8.8)	15.6141(8.4)	13.4547(7.9)
ADAPTATIONS	6	4	2	5	4

KOREAN FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	16.3813	13.6791	11.9463	10.6016	9.8648
SRER	20.0278(+0.4)	15.8461	14.5484(+0.5)	13.7053(+0.6)	12.8591(+0.7)
SEGM. SRER	22.8083(9.8)	17.7220(9.4)	16.4166(9.7)	15.5440(9.7)	14.4743(9.6)
ADAPTATIONS	4	2	3	5	4

RUSSIAN M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.2448	17.0169	14.8653	13.3074	11.9451
SRER	28.7623(+0.1)	23.8187(+0.4)	20.2982(+0.3)	17.0140(+0.4)	15.5108(+0.4)
SEGM. SRER	27.6600(11.2)	23.4790(11.7)	18.6808(10.4)	15.1947(9.4)	13.4103(9.2)
ADAPTATIONS	3	4	5	3	3

RUSSIAN FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	15.8164	12.2771	10.4241	8.9532	7.8970
SRER	22.2008(+1.1)	14.4225(+0.1)	13.2123(+0.7)	12.3085(+0.9)	10.5794(+0.5)
SEGM. SRER	21.7117(11.4)	14.9928(9.5)	13.1914(9.3)	12.2621(9.0)	11.3432(8.9)
ADAPTATIONS	8 (!!!)	2	2	3	2

Blackman - Hamming

ENGLISH M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	17.6371	14.4593	12.3178	10.6611	9.2207
SRER	24.3472	21.4141	16.4716	13.9853	12.3830
SEGM. SRER	25.3813(11.4)	21.0641(11.2)	12.1145(7.8)	14.9348(8.7)	10.5029(7.3)
ADAPTATIONS	2	4	2	2	2

ENGLISH FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.5417	16.2156	13.7421	11.9524	10.9437
SRER	25.1556	18.8760	17.2020	14.5922	13.3495
SEGM. SRER	26.9084(14.0)	19.2528(11.3)	16.0979(10.7)	14.4114(10.3)	13.5117(9.9)
ADAPTATIONS	4	2	2	1	2

JAPANESE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	23.0165	19.7793	17.6491	15.4760	14.1087
SRER	36.3954	25.7126	22.3081	19.5024	17.6193
SEGM. SRER	38.2036(12.3)	30.3591(10.6)	26.0363(10.1)	22.6738(9.1)	19.8101(9.0)
ADAPTATIONS	4	2	3	4	2

JAPANESE FE.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	21.8327	18.3221	16.2737	15.4312	14.0082
SRER	26.0853	20.5910	19.2443	17.8415	17.0315
SEGM. SRER	24.5925(11.7)	19.2370(10.7)	17.4201(10.6)	16.3617(10.5)	16.1118(10.0)
ADAPTATIONS	6	2	2	2	3

DEUTSCHE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.4074	16.9701	14.8528	12.9604	11.1599
SRER	28.9814	21.6925	17.0728	14.8786	13.5964
SEGM. SRER	30.3858(12.6)	23.7614(12.2)	17.2414(9.8)	15.8438(9.9)	13.9354(8.9)
ADAPTATIONS	2	2	1	2	2

DEUTSCHE FE.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	16.1179	12.2475	10.2301	8.9623	8.1531
SRER	24.1342	17.0403	12.1505	12.0517	10.5015
SEGM. SRER	26.1197(12.9)	18.8115(11.2)	15.2730(10.2)	13.7191(9.6)	12.6577(9.4)
ADAPTATIONS	4	4	2	2	2

Blackman - Hamming

ITALIAN M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	21.1434	18.0650	15.2240	13.5047	12.0281
SRER	28.0727	21.7733	18.0498	15.6969	14.6347
SEGM. SRER	28.4426(11.3)	22.4356(10.6)	18.3996(9.8)	15.6547(9.2)	14.3906(8.8)
ADAPTATIONS	3	2	2	2	2

ITALIAN FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.1347	14.8544	13.2329	11.8289	10.7089
SRER	20.5169	17.1795	15.1843	13.5338	12.2871
SEGM. SRER	24.1010(11.3)	20.1600(11.0)	16.9241(10.2)	15.3206(9.6)	13.8251(9.3)
ADAPTATIONS	1	3	4	2	2

FRANCE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	21.7938	18.7438	16.0626	14.4342	12.8683
SRER	30.198	23.9516	20.7674	18.5173	15.3774
SEGM. SRER	28.0715(10.8)	23.3265(11.0)	19.6064(10.5)	17.2020(9.9)	15.0352(9.0)
ADAPTATIONS	1	1	3	4	2

FRANCE FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.1285	15.9091	13.6914	12.2619	10.9482
SRER	27.2634	19.5699	16.2210	14.4110	13.8338
SEGM. SRER	26.4943(11.6)	19.6770(10.2)	16.8250(9.8)	15.5408(9.7)	14.4489(9.6)
ADAPTATIONS	4	3	2	2	2

GREEK M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.1538	15.9526	13.7194	11.6904	10.2269
SRER	27.5738	24.9261	20.7464	13.9418	13.0772
SEGM. SRER	30.7793(10.9)	26.1807(11.3)	21.2563(10.9)	14.3580(8.0)	13.0730(8.2)
ADAPTATIONS	3	3	4	1	2

GREEK FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.2335	14.9761	12.6855	10.8290	9.9589
SRER	24.0247	17.3261	14.8817	13.3690	11.7377
SEGM. SRER	22.6978(11.3)	16.1146(9.6)	12.9171(8.7)	11.6253(8.3)	10.1116(7.7)
ADAPTATIONS	3	3	2	2	1

Blackman - Taylor

ENGLISH M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	17.6371	14.4593	12.3178	10.6611	9.2207
SRER	25.8613(+1.5)	22.0515(+0.6)	16.3688(-0.1)	14.4926(+0.5)	12.9891(+0.6)
SEGM. SRER	24.5808(11.0)	20.0743(10.2)	15.0743(8.6)	12.8128(7.9)	10.7330(7.2)
ADAPTATIONS	4	4	2	4	2

ENGLISH FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.5417	16.2156	13.7421	11.9524	10.9437
SRER	25.5148(+0.4)	19.5744(+0.7)	17.8764(+0.6)	16.0280(+1.4)	14.0923(+0.6)
SEGM. SRER	23.2841(11.6)	18.8827(11.1)	16.6276(10.7)	15.2677(10.6)	14.6205(10.4)
ADAPTATIONS	2	2	3	2	3

JAPANESE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	23.0165	19.7793	17.6491	15.4760	14.1087
SRER	32.8487(-3.5)	25.6421(-0.1)	22.1461(-0.2)	18.6904(-0.9)	16.5837(-1.1)
SEGM. SRER	33.6567(10.7)	29.1927(10.1)	25.3359(9.7)	21.9273(8.7)	18.3196(8.5)
ADAPTATIONS	4	3	4	3	1

JAPANESE FE.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	21.8327	18.3221	16.2737	15.4312	14.0082
SRER	25.3755(-0.7)	20.8996(+0.3)	19.1109(-0.1)	17.9504(+0.1)	16.7695(+0.3)
SEGM. SRER	23.1760(11.5)	19.2577(10.6)	17.6211(10.8)	16.5201(10.6)	15.8741(10.3)
ADAPTATIONS	5	3	2	3	3

DEUTSCHE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.4074	16.9701	14.8528	12.9604	11.1599
SRER	27.7595(-1.2)	20.9811(-0.6)	18.4107(+1.4)	14.8817	13.7307(+0.2)
SEGM. SRER	28.3759(12.0)	20.5586(9.9)	18.6541(10.3)	15.5602(9.6)	13.7682(8.8)
ADAPTATIONS	3	1	3	2	2

DEUTSCHE FE.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	16.1179	12.2475	10.2301	8.9623	8.1531
SRER	24.0484(-0.1)	17.7794(+0.7)	12.2182(+0.1)	12.5010(+0.5)	11.0235(+0.5)
SEGM. SRER	24.7250(12.2)	18.7329(11.1)	14.2710(9.7)	13.6159(9.5)	12.8183(9.4)
ADAPTATIONS	7	8	1	2	3

Blackman - Taylor

ITALIAN M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	21.1434	18.0650	15.2240	13.5047	12.0281
SRER	27.8143(-0.2)	22.7633(+1.0)	18.5000(+0.5)	16.6396(+1.0)	14.6173
SEGM. SRER	26.8126(10.5)	22.3369(10.5)	18.1670(9.6)	16.0126(9.1)	14.3302(8.6)
ADAPTATIONS	4	5	3	4	2

ITALIAN FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.1347	14.8544	13.2329	11.8289	10.7089
SRER	21.2578(+0.7)	17.1095	15.8337(+0.4)	14.1038(+0.6)	12.5277(+0.3)
SEGM. SRER	25.1221(11.7)	19.5056(10.6)	17.1904(9.9)	15.6419(9.6)	14.1550(9.3)
ADAPTATIONS	4	2	3	3	2

FRANCE M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	21.7938	18.7438	16.0626	14.4342	12.8683
SRER	32.0335(+1.9)	26.3883(+2.4)	20.9994(+0.2)	17.9534(-0.6)	15.7224(+0.4)
SEGM. SRER	28.6718(12.2)	24.6426(11.8)	19.6272(10.1)	16.5752(9.4)	14.5780(8.5)
ADAPTATIONS	3	3	3	2	1

FRANCE FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.1285	15.9091	13.6914	12.2619	10.9482
SRER	25.8490(-1.4)	18.5810(-1.0)	16.5504(+0.3)	14.1740(-0.3)	14.5153(+0.7)
SEGM. SRER	24.8823(11.1)	18.61539(9.6)	16.8281(9.5)	15.7301(9.9)	14.4393(9.5)
ADAPTATIONS	3	1	2	3	2

GREEK M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.1538	15.9526	13.7194	11.6904	10.2269
SRER	26.7930(-0.8)	24.1941(+0.7)	20.7363	15.7687(+1.8)	12.8744(-0.2)
SEGM. SRER	27.6193(10.4)	24.2903(10.2)	19.8183(9.2)	15.7769(8.6)	12.8727(7.8)
ADAPTATIONS	4	4	3	3	2

GREEK FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.2335	14.9761	12.6855	10.8290	9.9589
SRER	24.1198(+0.1)	18.0083(+0.7)	15.3354(+0.5)	13.5619(+0.2)	12.0860(+0.3)
SEGM. SRER	21.9739(10.6)	15.8536(9.4)	13.2654(8.7)	11.8442(8.5)	10.6966(8.0)
ADAPTATIONS	4	3	3	2	2

Blackman - Hamming

SPANISH M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.2017	16.9164	15.3325	12.0728	10.5700
SRER	31.0410	23.7292	20.7868	16.0411	14.0528
SEGM. SRER	30.3647(14.1)	25.3401(12.9)	21.0846(11.5)	16.7530(9.7)	13.9735(8.8)
ADAPTATIONS	3	3	4	3	2

SPANISH FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.7418	16.8328	14.3857	12.5226	11.5642
SRER	29.0291	22.0635	17.8175	15.5807	14.3462
SEGM. SRER	28.5251(12.7)	20.7293(10.2)	17.2342(9.2)	15.642(8.9)	14.4182(8.8)
ADAPTATIONS	6 (!!!)	4	2	2	2

VASCON M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.2234	16.6967	15.0194	12.2278	10.5580
SRER	31.8391	28.0806	20.8460	15.4063	13.8354
SEGM. SRER	28.3898(12.8)	26.1663(14.1)	20.5107(11.4)	14.8246(8.9)	12.6132(8.4)
ADAPTATIONS	2	7 (!!!)	8 (!!!)	2	3

VASCON FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.0975	14.4186	12.4223	11.0454	10.1193
SRER	23.8699	17.2707	14.4151	13.4867	12.3068
SEGM. SRER	25.5387(11.3)	19.4566(10.0)	16.3115(8.7)	14.8907(8.3)	13.6477(7.9)
ADAPTATIONS	2	3	2	2	2

ISLAND M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.3194	16.9531	14.5895	12.4372	11.1755
SRER	30.3885	21.5821	17.7933	15.4614	14.2682
SEGM. SRER	27.2735(11.7)	20.4406(9.9)	16.6993(9.2)	13.9608(8.6)	12.7004(8.4)
ADAPTATIONS	2	2	3	2	2

ISLAND FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.8947	16.6949	14.3770	12.9409	11.9150
SRER	24.8485	19.2627	16.5463	14.9510	13.9449
SEGM. SRER	22.5207(10.6)	16.6011(9.2)	14.2773(9.0)	12.9127(8.8)	12.0486(8.4)
ADAPTATIONS	5	3	3	2	2

Blackman - Taylor

SPANISH M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.2017	16.9164	15.3325	12.0728	10.5700
SRER	30.3030(-1.0)	21.7926(-2.0)	20.7069	15.3340(-0.7)	13.9617(-0.1)
SEGM. SRER	27.2101(12.6)	21.0805(9.8)	20.1053(10.9)	16.1515(9.5)	13.9041(8.5)
ADAPTATIONS	4	1	5	3	2

SPANISH FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.7418	16.8328	14.3858	12.5226	11.5642
SRER	28.2063(-0.7)	22.4367(+0.4)	17.2561(-0.6)	15.6922(+0.1)	14.7738(+0.3)
SEGM. SRER	26.0877(11.3)	20.3027(10.0)	17.0425(9.3)	16.0441(9.0)	14.7930(8.8)
ADAPTATIONS	4	3	2	3	3

VASCON M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.2234	16.6967	15.0194	12.2278	10.5580
SRER	29.5075(-2.3)	25.5077(-2.5)	21.0871(+0.2)	16.0098(+0.6)	14.2420(+0.6)
SEGM. SRER	26.2835(12.6)	23.0271(11.8)	18.9370(10.4)	14.3836(8.7)	12.7946(8.3)
ADAPTATIONS	4	5	4	2	3

VASCON FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.0975	14.4186	12.4223	11.0454	10.1193
SRER	24.7895(+0.9)	17.9730(+0.7)	15.0877(+0.7)	13.8428(+0.4)	12.9223(+0.6)
SEGM. SRER	25.0340(11.6)	19.1289(9.6)	16.4718(8.7)	14.8176(8.5)	13.9446(8.2)
ADAPTATIONS	4	3	3	2	3

ISLAND M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.3194	16.9531	14.5895	12.4372	11.1755
SRER	30.2546(-0.1)	22.1743(+0.6)	18.3684(+0.5)	15.6595(+0.2)	15.0462(+0.8)
SEGM. SRER	26.3849(11.8)	19.9196(9.5)	16.4279(8.8)	13.9211(8.5)	13.3165(8.4)
ADAPTATIONS	4	2	3	2	4

ISLAND FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.8947	16.6949	14.3770	12.9409	11.9150
SRER	24.8747	19.7381(+0.5)	17.1783(+0.4)	15.2640(+0.3)	13.8495(-0.1)
SEGM. SRER	21.8512(10.0)	16.8003(9.0)	14.4918(9.0)	13.0781(8.9)	11.6716(8.3)
ADAPTATIONS	5	5	4	2	2

Blackman - Hamming

INDIA M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	16.5244	12.5640	9.7256	8.0380	6.9337
SRER	24.7411	16.1001	12.1947	10.8246	9.5866
SEGM. SRER	30.9415(10.8)	22.3226(9.1)	17.0208(8.3)	15.1010(7.8)	13.4882(7.5)
ADAPTATIONS	3	3	2	2	3

INDIA FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	13.8345	9.6567	8.6908	7.1796	6.2254
SRER	17.0609	16.2551	13.4401	12.6425	11.4598
SEGM. SRER	23.8759(12.1)	19.9494(11.2)	17.5943(11.2)	16.7918(10.9)	16.2514(10.5)
ADAPTATIONS	3	3	2	2	2

TURKEY M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.2190	14.9769	12.9168	10.9659	9.6519
SRER	26.3872	19.5186	16.6222	14.6013	12.9504
SEGM. SRER	27.3128(13.7)	21.1069(12.7)	15.6668(10.6)	12.9991(9.2)	11.2363(8.5)
ADAPTATIONS	3	4	3	3	2

TURKEY FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.0935	15.4142	13.8603	11.9239	10.8251
SRER	23.9161	19.0114	16.9062	15.2547	14.0011
SEGM. SRER	22.7093(11.3)	17.7096(11.0)	15.8029(10.9)	14.1913(10.3)	13.7746(10.0)
ADAPTATIONS	2	4	3	2	2

FINLAND M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	22.1971	18.2486	15.7174	14.6387	13.3119
SRER	32.4460	29.3633	20.8336	19.6345	16.1444
SEGM. SRER	28.8725(11.8)	28.9946(14.4)	22.9036(12.1)	18.6411(10.9)	14.7225(9.1)
ADAPTATIONS	1	3	2	2	1

FINLAND FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.1392	16.5441	14.1910	12.4466	11.4102
SRER	26.3378	19.8998	16.6450	14.8370	12.9224
SEGM. SRER	25.2423(12.0)	19.6252(11.3)	16.3726(10.5)	14.7769(9.8)	13.4516(9.4)
ADAPTATIONS	2	3	2	2	1

Blackman - Taylor

INDIA M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	16.5244	12.5640	9.7256	8.0380	6.9337
SRER	26.2771(+1.5)	14.8947(-2.1)	13.5009(+1.4)	10.8495	10.4443(+0.9)
SEGM. SRER	30.3320(10.2)	19.9406(8.2)	17.4332(8.2)	15.1618(7.6)	13.9573(7.4)
ADAPTATIONS	6	1	5	2	4

INDIA FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	13.8345	9.6567	8.6908	7.1796	6.2254
SRER	16.8344(+0.2)	15.9830(-0.3)	13.6707(+0.2)	12.8438(+0.2)	11.8948(+0.4)
SEGM. SRER	23.2326(11.7)	19.4564(11.0)	17.3779(10.9)	16.7162(10.7)	16.2189(10.5)
ADAPTATIONS	4	4	2	2	2

TURKEY M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	18.2190	14.9769	12.9168	10.9659	9.6519
SRER	26.4213(+0.1)	20.0102(+0.5)	16.4926(+0.2)	14.6589	13.1110(+0.2)
SEGM. SRER	24.7510(12.4)	19.4382(11.4)	15.3132(10.2)	12.6600(8.9)	11.2860(8.3)
ADAPTATIONS	4	3	3	2	2

TURKEY FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	19.0935	15.4142	13.8603	11.9239	10.8251
SRER	24.3206(+0.4)	18.5882(-0.5)	17.2249(+0.3)	15.8781(+0.6)	14.7368(+0.7)
SEGM. SRER	23.0424(11.2)	17.4411(10.5)	15.6950(10.5)	14.6430(9.5)	14.1481(10.3)
ADAPTATIONS	4	4	3	3	3

FINLAND M.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	22.1971	18.2486	15.7174	14.6387	13.3119
SRER	30.6561(+1.8)	28.0584(-1.3)	23.4511(+2.6)	18.3192(-1.3)	17.5679(+1.4)
SEGM. SRER	26.2558(10.9)	25.4385(12.4)	23.2157(12.7)	16.4085(9.5)	16.2140(9.8)
ADAPTATIONS	1	3	5	1	4

FINLAND FEM.	5 MS	7.5 MS	10 MS	12.5 MS	15 MS
1 ST SRER	20.1392	16.5441	14.1910	12.4466	11.4102
SRER	26.0064(-0.3)	19.8986	16.7903(+0.1)	15.6430(-0.8)	13.4091(+0.5)
SEGM. SRER	24.7194(12.3)	19.2688(11.1)	16.2785(10.3)	14.9213(9.8)	13.8669(9.7)
ADAPTATIONS	4	3	2	2	2

6 Βιβλιογραφία

[1] https://en.wikipedia.org/wiki/Nyquist_frequency

[2] [https://en.wikipedia.org/wiki/Bandwidth_\(signal_processing\)](https://en.wikipedia.org/wiki/Bandwidth_(signal_processing))

[EIK1],[EIK2] https://en.wikipedia.org/wiki/Window_function

[KOS11] George P. Kafentzis, Olivier Rosenc, Yannis Stylianou, Robust full-band adaptive Sinusoidal analysis and synthesis of speech. IEEE Acoustics, Speech and Signal Processing (ICASSP), pages 6260-6264, May 2014

[KPOS12] George P. Kafentzis, Y. Pantazis, Olivier Rosenc, Yannis Stylianou. An extension of the adaptive Quasi Harmonic Model. Orange Labs. Lannion, France, pages 4605-4608, March 2012

[LSM93] J. Laroche Y. Stylianou and E. Moulines. HNM: A Simple, Effecient Harmonic plus Noise Model for Speech. In Workshop on Appl. of Signal Proc. to Audio and Acoustics (WASPAA), pages 169–172, New Paltz, NY, USA, Oct 1993.

[MQ86] R. J. McAulay and T. F. Quatieri. Speech Analysis/Synthesis based on a Sinusoidal Representation. IEEE Trans. on Acoustic Speech and Signal Processing, 34:744–754, 1986.

[PRS08] Y. Pantazis, O. Rosenc, and Y. Stylianou. On the Properties of a Time-Varying Quasi-Harmonic Model of Speech. In Interspeech, Brisbane, September 2008.

[PRS09a] Y. Pantazis, O. Rosenc, and Y. Stylianou. AM-FM estimation for speech based on a time-varying sinusoidal model. In Interspeech, Brighton, September 2009.

[PRS09b] Y. Pantazis, O. Rosenc, and Y. Stylianou. Chirp rate estimation of speech based on a time-varying quasiharmonic model. In Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), pages 3985–3988, April 2009.

[PRS10] Y. Pantazis, O. Rosenc, and Y. Stylianou. On the robustness of the quasi-harmonic model of speech. In Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), pages 4210–4213, March 2010.

[PRS11] Y. Pantazis, O. Rosenc, and Y. Stylianou. Adaptive AM-•FM signal decomposition with application to speech analysis. IEEE Transactions on Audio, Speech, and Language Processing, 19:290–300, 2011.

[PTRS10] Y. Pantazis, G. Tzedakis, O. Rosenc, and Y. Stylianou. Analysis/Synthesis of Speech based on an Adaptive Quasi-Harmonic plus Noise Model. In Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP), Dallas, Texas, USA, March 2010.

[Sty96] Y. Stylianou. Harmonic plus Noise Models for Speech, combined with Statistical Methods, for Speech and Speaker Modification. PhD thesis, E.N.S.T - Paris, 1996.



University of Crete