

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΓΕΝΙΚΟ ΤΜΗΜΑ



ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΕΦΑΡΜΟΣΜΕΝΕΣ ΕΠΙΣΤΗΜΕΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑ

**ΔΙΠΛΩΜΑΤΙΚΗ ΔΙΑΤΡΙΒΗ ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ
ΕΙΔΙΚΕΥΣΗΣ ΚΑΤΕΥΘΥΝΣΗ : «ΕΦΑΡΜΟΣΜΕΝΑ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΑ
ΜΑΘΗΜΑΤΙΚΑ»**

**ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΜΕΓΑΛΩΝ ΑΡΑΙΩΝ
ΣΥΣΤΗΜΑΤΩΝ ΠΟΥ ΠΡΟΚΥΠΤΟΥΝ ΑΠΟ ΤΗΝ
ΔΙΑΚΡΙΤΟΠΟΙΗΣΗ ΠΕΠΕΡΑΣΜΕΝΩΝ ΟΓΚΩΝ
ΕΞΙΣΩΣΕΩΝ ΤΥΠΟΥ BOUSSINESQ**

ΓΑΙΤΑΝΗ Σ. ΜΑΡΙΑ

Επιβλέπων : Επίκουρος Καθηγητής Δελής Ανάργυρος

ΧΑΝΙΑ , 2013

Η διατριβή αυτή εξετάστηκε επιτυχώς στις 12-03-2013 από τη παρακάτω Τριμελή Επιτροπή:

- Επίκ. Καθηγητή Α.Δελή ως επιβλέπων
- Επίκ. Καθηγητή Εμμ.Μαθιουδάκη
- Επίκ Καθηγητή Ι.Νικολό

η οποία ορίστηκε κατά τη 20η/14-2-2013 συνεδρίαση της Γ.Σ.Ε.Σ. του Γενικού Τμήματος του Πολυτεχνείου Κρήτης.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω την Τριμελή Επιτροπή μου και ιδιαίτερα τον Επίκουρο Καθηγητή, κ. Ανάργυρο Ι. Δελή για την ηθική και επιστημονική καθοδήγηση που μου παρείχε. Επίσης, θα ήθελα να ευχαριστήσω όλα τα μέλη του εργαστηρίου Εφαρμοσμένων Μαθηματικών και Η/Υ του Πολυτεχνείου Κρήτης για τη πολύτιμη βοήθεια και τη συμπαράσταση.

Θερμές ευχαριστίες οφείλω στην Υποψήφια Διδάκτορα κ. Μαρία Καζολέα για τις συμβουλές και την βοήθεια που μου παρείχε καθ' όλη τη διάρκεια της υλοποίησης της εργασίας. Τέλος, αισθάνομαι την ανάγκη να ευχαριστήσω την οικογένεια μου και τον Υποψήφιο Διδάκτορα κ. Ιωάννη Αθανασάκη για τα ηθικά και πνευματικά αγαθά που μου παρείχαν.

Περίληψη

Στην παρούσα διατριβή παρουσιάζονται τα αποτελέσματα της αριθμητικής επίλυσης μεγάλων και αραιών γραμμικών συστημάτων που προκύπτουν από την εφαρμογή των αριθμητικών μεθόδων πεπερασμένων όγκων κατά τη διακριτοποίηση εκτεταμένων διαφορικών εξισώσεων τύπου Boussinesq σε μη δομημένα υπολογιστικά πλέγματα. Γίνεται σύγκριση μεταξύ διαφόρων επαναληπτικών μεθόδων σε συνδυασμό με προρρύθμιση και αναδιάταξη των εξισώσεων με σκοπό τη καλύτερη επιλογή για μείωση του υπολογιστικού κόστους και την εύρεση της ταχύτερης αριθμητικής λύσης. Παρουσιάζεται μελέτη της συμπεριφοράς επίλυσης σε εκτεταμένα αριθμητικά πειράματα για διαφόρους τύπους και μεγέθη υπολογιστικών πλεγμάτων μεταβάλλοντας φυσικές παραμέτρους, όπως είναι η μέση στάθμη του νερού. Προτείνεται ο καταλληλότερος και αποδοτικότερος συνδυασμός για την επίλυση του παραπάνω προβλήματος, όσο αφορά την ακρίβεια και το υπολογιστικό κόστος της μεθόδου.

Περιεχόμενα

Ευχαριστίες	ii
Περίληψη	iii
1 Εισαγωγή	1
2 Φυσικό Πρόβλημα	5
2.1 Μαθηματικό Μοντέλο	7
2.2 Αριθμητικό Σχήμα Πεπερασμένων Όγκων	10
2.3 Γραμμικό Σύστημα	17
2.4 Υπολογιστικά Πλέγματα και Δείκτης κατάστασης	21
3 Αραιοί Πίνακες	31
3.1 Αραιές μορφές αποθήκευσης	32
3.1.1 COO (Coordinate Format)	32
3.1.2 CSR (Compressed Sparse Row)	33
3.1.3 CSC (Compressed Sparse Column)	34
3.1.4 MSR (Modified Sparse Row)	34
3.2 SPARSKIT	35
4 Επαναληπτικές Μέθοδοι Επίλυσης	38
4.1 Εισαγωγή	38
4.1.1 Μέθοδοι Jacobi και Gauss-Seidel	39

4.1.2 Successive Overrelaxation (SOR) και Symmetric Successive Over-relaxation (SSOR)	41
4.2 Krylov Subspace Methods	42
4.2.1 GMRES	43
4.2.2 BICGSTAB	45
4.2.3 FGMRES	46
4.2.4 TFQMR	47
4.2.5 CG	48
4.3 Προρρυθμιστές (Preconditioners)	49
4.3.1 ILU	51
4.3.2 $ILU0$	52
4.3.3 $ILU(k)$	54
4.3.4 $ILUT$	55
4.4 Αναδιάταξη Αγνώστων/Εξισώσεων	56
4.4.1 CMK	57
4.4.2 RCM	57
5 Αριθμητικά Αποτελέσματα	60
5.1 Μελέτη πινάκων για μέση στάθμη $h = 1m$	60
5.2 Μελέτη πινάκων για μέση στάθμη $h = 10m$	66
5.3 Μελέτη πινάκων για βάθος $h = 100m$	70
5.4 Μελέτη πινάκων με τοπογραφία.	74
6 Συμπεράσματα	83

Κατάλογος Σχημάτων

2.1 Ροή με ελεύθερο σύνορο υπό την επίδραση της βαρύτητας.	7
2.2 Όγκος ελέγχου για ένα εσωτερικό (αριστερά) και ένα συνοριακό κόμβο (δεξιά). 1 1	
2.3 Ορισμός του χωρίου όπου γίνεται αριθμητικός υπολογισμός της κλίσης και της απόκλισης.	13
2.4 Υπολογιστικά κελιά για την κλίση της απόκλισης : εσωτερικό κελί (αριστερά) και συνοριακό κελί (δεξιά).	16
2.5 Τα μη μηδενικά στοιχεία (αριστερά) και ο CPU χρόνος (δεξιά) σε συνάρτηση με το N [33].	21
2.6 Αντιπροσωπευτικοί τύποι πλεγμάτων.	22
2.7 Δομή πίνακα που προκύπτει από Distorted πλέγμα με $N_x=15$, $N = 297$. .	23
2.8 Δομή πίνακα που προκύπτει από Distorted πλέγμα με $N_x = 30$, $N = 1102$. 23	
2.9 Δομή πίνακα που προκύπτει από Distorted πλέγμα με $N_x = 60$, $N = 4305$. 23	
2.10 Δομή πίνακα που προκύπτει από Equilateral πλέγμα με $N_x=15$, $N = 297$. 24	
2.11 Δομή πίνακα που προκύπτει από Equilateral πλέγμα με $N_x = 30$, $N = 1102$. 24	
2.12 Δομή πίνακα που προκύπτει από Equilateral πλέγμα με $N_x = 60$, $N = 4305$. 24	
2.13 Δομή πίνακα που προκύπτει από Orthogonal (type I) πλέγμα με $N_x = 15$, $N = 481$	25
2.14 Δομή πίνακα που προκύπτει από Orthogonal (type I) πλέγμα με $N_x = 30$, $N = 1861$	25
2.15 Δομή πίνακα που προκύπτει από Orthogonal (type I) πλέγμα με $N_x = 60$, $N = 7321$	25

2.16 Δομή πίνακα που προκύπτει από Orthogonal (type II) $N_x = 15$, $N = 256$.	26
2.17 Δομή πίνακα που προκύπτει από Orthogonal (type II) $N_x = 30$, $N = 961$.	26
2.18 Δομή πίνακα που προκύπτει από Orthogonal (type II) πλέγμα με $N_x = 15$, $N = 37114$.	26
2.19 Ιδιοτιμές του Distorted πίνακα $N = 297$.	28
2.20 Ιδιοτιμές του Distorted πίνακα $N = 1102$.	28
2.21 Ιδιοτιμές του Orthogonal (type I) πίνακα $N = 481$.	29
2.22 Ιδιοτιμές του Orthogonal (type I) πίνακα $N = 1861$.	29
2.23 Ιδιοτιμές του Orthogonal (type II) πίνακα $N = 256$.	30
2.24 Ιδιοτιμές του Orthogonal (type II) πίνακα $N = 961$.	30
4.1 Μορφή πίνακα τύπου I μετά την χρήση της CMK.	58
4.2 Μορφή πίνακα τύπου II μετά την χρήση της CMK.	58
5.1 Συγκριτικό ραβδόγραμμα των προρρυθμιστών ILUT και ILUk για $N_x = 15$ και μέση στάθμη $h = 1m$.	65
5.2 Συγκριτικό ραβδόγραμμα για το χρόνο εκτέλεσης των προρρυθμιστών ILUT και ILUk για $N_x = 15$ και μέση στάθμη $h = 1m$.	66
5.3 Συγκριτικό ραβδόγραμμα της GMRES και των προρρυθμιστών ILUT και ILUk για $N_x = 15$ και βάθος $h = 10m$.	75
5.4 Συγκριτικό ραβδόγραμμα της BiCGSTAB και των προρρυθμιστών ILUT και ILUk για $N_x = 15$ και βάθος $h = 10m$.	76
5.5 Συγκριτικό ραβδόγραμμα των μεθόδων GMRES και BicGSTAB μέσης στάθ- μης $h = 10m$.	77
5.6 Μορφή τοπογραφίας.	80

Κεφάλαιο 1

Εισαγωγή

Στη θαλάσσια και παράκτια μηχανική η μοντελοποίηση των μη γραμμικών και διασπειρόμενων κυματισμών σε ρεαλιστικά περιβάλλοντα είναι σημαντική και έχει αντικαταστήσει σε μεγάλο βαθμό τα εργαστηριακά πειράματα για την μελέτη και κατασκευή παράκτιων κατασκευών αλλά και την πρόβλεψη πλημμυρικών φαινομένων. Για αυτό το λόγο στην ερευνητική περιοχή της μαθηματικής και αριθμητικής μοντελοποίησης αναπτύχθηκαν μέθοδοι τις τελευταίες δεκαετίες, έτσι ώστε να παρέχουν τα κατάλληλα μέσα για ακριβή πρόβλεψη των κυματικών φαινομένων κοντά στην ακτή όπως το φαινόμενο της ρήξης, της διάθλασης, της περίθλασης, των αρμονικών αλληλεπιδράσεων καθώς και της αναρρίχησης των κυμάτων στην ακτή.

Ιδιαίτερη επιστημονική προσπάθεια έχει αναπτυχθεί για την επίλυση σημαντικών θεμάτων προσομοίωσης που περιλαμβάνουν, την εγκυρότητα των μαθηματικών μοντέλων στις παράκτιες ζώνες, όπως και σε βαθύτερα νερά, την θραύση των κυματισμών, την μετάβαση από υπο-κρίσιμη σε υπερ-κρίσιμη ροή, την διασπορά και τη σωστή αριθμητική προσέγγιση της φυσικής τοπογραφίας καθώς και τις διαδικασίες υπολογισμού υγρών/στεγανών μετώπων στις ροές. Για το σκοπό αυτό και την αποφυγή της πολυπλοκότητας και του υπολογιστικού κόστους για την επίλυση των εξισώσεων Navier Stokes [21] χρησιμοποιείται η ολοκλήρωση κατά βάθος ώστε να απλοποιηθούν οι εξισώσεις και τα μοντέλα που προκύπτουν να είναι αριθμητικά εύκολα επιλύσιμα. Υποθέτοντας ότι η κάθετη συνιστώσα της ταχύτητας είναι αμελητέα και χρησιμοποιώντας την εξίσωση

διατήρησης της μάζας και την εξίσωση διατήρησης της ορμής παράγονται τα ευρέως χρησιμοποιούμενα μοντέλα μέσου βάθους. Οι πολύ διαδεδομένες μη γραμμικές εξισώσεις ρηχών υδάτων ανήκουν σε αυτή τη κατηγορία. Παρόλο που η χρήση των μη γραμμικών εξισώσεων ρηχών υδάτων είναι ικανή να μοντελοποιήσει βασικές πτυχές της διάδοσης κυματισμών, τα γενικά χαρακτηριστικά της αναρρίχησης αυτών στην ακτή, δεν είναι κατάλληλα για βαθύτερα νερά όπου τα φαινόμενα διασποράς είναι πιο σημαντικά από τα μη γραμμικά φαινόμενα. Από την άλλη πλευρά οι εξισώσεις τύπου Boussinesq εισάγουν όρους διασποράς και είναι καταλληλότερες σε νερά όπου η διασπορά ξεκινάει να επηρεάζει την ελεύθερη επιφάνεια του υγρού. Οι εξισώσεις τύπου Boussinesq είναι κυρίως διδιάστατα μοντέλα τα οποία δεν εξαρτώνται από το πρόβλημα του ελεύθερου συνόρου λόγω του ότι εξαλείφουν την κάθετη συνιστώσα της ταχύτητας.

Κατά την αριθμητική επίλυση των εξισώσεων τύπου Boussinesq είτε χρησιμοποιώντας μεθόδους πεπερασμένων διαφορών, είτε πεπερασμένων όγκων, είτε πεπερασμένων στοιχείων πολλές φορές καταλήγουμε σε ένα γραμμικό σύστημα, το οποίο πρέπει να λυθεί σε κάθε διακριτό χρονικό βήμα για να πάρουμε τη λύση του πεδίου ροής, ή απαιτείται η επίλυση του σε ένα ενδιάμεσο βήμα, ώστε να λάβουμε τις ταχύτητες του πεδίου ροής [14, 35, 33, 34]. Ακολουθώντας την εργασία [33] και χρησιμοποιώντας πεπερασμένους όγκους, η επίλυση ενός αραιού γραμμικού συστήματος είναι απαραίτητη σε κάθε χρονικό βήμα για τη λήψη του διανύσματος της ταχύτητας του πεδίου ροής. Το θέμα της παρούσας εργασίας είναι η μελέτη της συμπεριφοράς επίλυσης του παραπάνω αραιού γραμμικού συστήματος, για να προταθεί η καταλληλότερη μέθοδος επίλυσης του χρησιμοποιώντας τις πιο διαδεδομένες επαναληπτικές μεθόδους σε συνδυασμό με διαδικασίες αναδιάταξης και προρρύθμισης.

Η ιστορική αναδρομή των επαναληπτικών μεθόδων που ανήκουν στη κατηγορία των υπόχωρων Krylov ξεκινάει το 1952, με τους Lanczos [20], Hestenes και Stiefel [17], οι οποίοι πρότειναν τη δημιουργία της CG (Conjugate Gradient) για την επίλυση γραμμικών συστημάτων με συμμετρικούς και θετικά ορισμένους πίνακες συντελεστών. Για

διάφορους λόγους, η χρήση της CG ήταν περιορισμένη για σχεδόν 20 χρόνια, με την SOR (Successive Overrelaxation) και την μέθοδο Chebushev να χρησιμοποιούνται περισσότερο για λύση αραιών συστημάτων.

Το 1970 με την εργασία του Reid [26], έγινε γνωστό ότι για μεγάλα αραιά συστήματα με καλό δείκτη κατάστασης η CG είναι η αποδοτικότερη δίνοντας ικανοποιητικές προσεγγίσεις σε λιγότερα από n βήματα, όπου n η τάξη του πίνακα. Η εξέλιξη της CG για μόνο συμμετρικούς πίνακες έδωσε τις MINRES (Minimal Residual Method) και SYMMLQ (Symmetric LQ Method) μεθόδους [24]. Το 1980 οι Saad και Schultz [32] εξέλιξαν την GMRES για μη συμμετρικές περιπτώσεις, οι Freund και Nachtigal [13] βελτίωσαν την μέθοδο QMR (Quasi-Minimal Residual Method) και τέλος πιο πρόσφατα εμφανίστηκε η μέθοδος BiCGTAB (Biconjugate Gradient Stabilized Method) από τον van der Vorst [16]. Για την επίλυση του αραιού συστήματος θα μελετήσουμε τις επαναληπτικές μεθόδους που ανήκουν στην κατηγορία των υπόχωρων Krylov, αφού είναι δημοφιλέστερες για αραιούς πίνακες και υποστηρίζονται από την βιβλιοθήκη υποπρογραμμάτων γραμμικής άλγεβρας Sparskit.

Σε πολλές περιπτώσεις, και ανάλογα το πρόβλημα, η επαναληπτική μέθοδος μπορεί να μην είναι αποδοτική, όσο αφορά τη σύγκλιση. Η χρήση προρρυθμιστή σε αυτή τη περίπτωση είναι αναγκαία για την βελτίωση-μετατροπή του αρχικού γραμμικού συστήματος σε ένα άλλο καταλληλότερο. Η ιδέα της προρύθμισης είναι, αντί του αρχικού γραμμικού συστήματος να λυθεί ένα προρρυθμισμένο με καλύτερες ιδιότητες από το αρχικό, ώστε να υπάρχει ταχύτερη σύγκλιση. Στην παρούσα εργασία θα μελετηθούν οι προρρυθμιστές ILU0 (Incomplete LU), ILUK (Incomplete LU με level of fill) και ILUT (Incomplete LU με Threshold). Μία ακόμα τεχνική διόρθωσης του μεγάλου δείκτη κατάστασης του πίνακα είναι η αναδιάταξη των στοιχείων του. Η επιρροή της αναδιάταξης στην σύγκλιση των προρρυθμισμένων Krylov μεθόδων έχει απασχολήσει το επιστημονικό περιβάλλον, γιατί οδηγεί σε μείωση του επιπέδου γεμίσματος, εισάγει την εφαρμογή παραλληλοποίησης και επηρεάζει τον ρυθμό σύγκλισης [7]. Κάποιες από τις κλασσικές στρατηγικές αναδιά-

ταξης είναι η Cuthill-McKee (CMK), Reverse Cuthill-McKee (RCM) και Sloan's ordering [7].

Η παρούσα εργασία δομείται ως εξής: Στο δεύτερο κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο του μαθηματικού μοντέλου που χρησιμοποιείται καθώς και η αριθμητική διακριτοποίηση αυτού, τόσο στο χώρο όσο και στο χρόνο. Μετά την εφαρμογή του πολυβηματικού χρονικού σχήματος, προκύπτει ένα γραμμικό σύστημα το οποίο πρέπει να επιλυθεί σε κάθε βήμα, του οποίου ο πίνακας είναι αραιός με σχεδόν συμμετρική μορφή. Επιπρόσθετα, παρουσιάζονται τα υπολογιστικά πλέγματα, δομημένα και μή που προέρχονται από τριγωνισμούς, από τα οποία προκύπτουν οι αραιοί πίνακες. Η γεωμετρία του πλέγματος επηρεάζει σημαντικά τη δομή του πίνακα. Στο τρίτο κεφάλαιο παρουσιάζονται οι δυνατές αραιές μορφές αποθήκευσης, καθώς και η βιβλιοθήκη γραμμικής άλγεβρας Sparskit. Στο τέταρτο κεφάλαιο, αναλύονται οι στατικές επαναληπτικές μέθοδοι καθώς και αυτές που ανήκουν στον υπόχωρο Krylov, παρουσιάζονται οι αλγόριθμοι τους και περιγράφονται οι διάφοροι παράμετροι τους. Στη συνέχεια, μελετώνται πιθανοί προορυθμιστές που μπορούν να χρησιμοποιηθούν, καθώς και δύο δυνατότητες αναδιάταξης. Στο επόμενο κεφάλαιο παρουσιάζονται και συγκρίνονται εκτεταμένα αριθμητικά πειράματα που προκύπτουν μεταβάλλοντας κάποιες από τις παραμέτρους τους. Τέλος, προτείνεται ο καταλληλότερος και αποδοτικότερος τρόπος επίλυσης του υπό μελέτη γραμμικού συστήματος.

Κεφάλαιο 2

Φυσικό Πρόβλημα

Οι ροές νερού με ελεύθερη επιφάνεια κάτω από την επίδραση της βαρύτητας αποτελούν μία μεγάλη κλάση προβλημάτων επιστημονικού ενδιαφέροντος. Φυσικά φαινόμενα όπως πλημμύρες, κατάρρευση φραγμάτων, ροή ποταμών και διάδοση μακρών κυματισμών απασχολούν την επιστημονική κοινότητα λόγω του μεγάλου πρακτικού ενδιαφέροντος που παρουσιάζουν. Σε πολλές περιπτώσεις τέτοιες ροές μοντελοποιούνται ικανοποιητικά από ένα σύνολο μη γραμμικών διαφορικών εξισώσεων, οι οποίες προέρχονται από τις εξισώσεις Navier Stokes [21].

Λόγω του μεγάλου υπολογιστικού κόστους και της πολυπλοκότητας για την επίλυση των εξισώσεων Navier Stokes χρησιμοποιούνται πιο απλοποιημένες εξισώσεις όπως είναι οι εξισώσεις ρηχών υδάτων. Οι τελευταίες δεν είναι κατάλληλες για βαθιά νερά λόγω των φαινομένων διασποράς που υπερισχύουν αυτών της μη γραμμικότητας, οπότε για αυτό το λόγο γίνεται χρήση των εξισώσεων Boussinesq. Υπό την υπόθεση ότι η μη γραμμικότητα και η διασπορά είναι ασθενής και διατηρώντας την ίδια τάξη προσέγγισης ο Peregrine [25] διατύπωσε τις πρότυπες Boussinesq εξισώσεις για μεταβαλλόμενη μέση στάθμη χρησιμοποιώντας την μετατόπιση της ελεύθερης επιφάνειας και τη μέση τιμή της ταχύτητας ως εξαρτώμενες μεταβλητές. Οι πρότυπες Boussinesq εξισώσεις δεν ισχύουν όταν η μέση στάθμη είναι μεγαλύτερη από το ένα πέμπτο του ισοδύναμου μήκους κύματος σε βαθιά νερά και έτσι η χρήση τους περιορίζεται σε σχετικά ρηχά νερά.

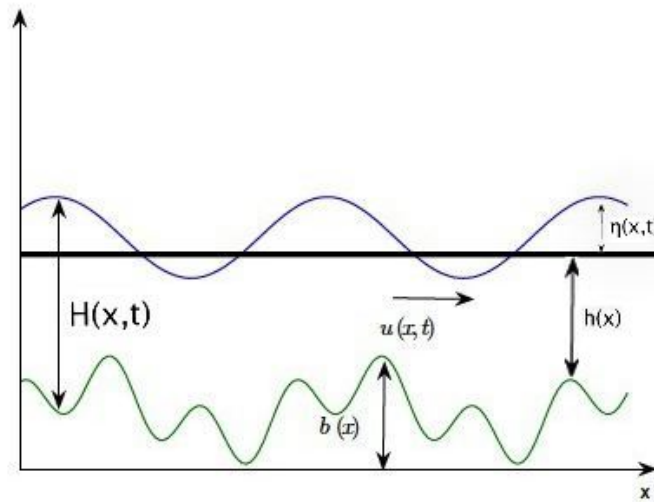
Αξιοσημείωτες προσπάθειες έχουν γίνει τις τελευταίες δύο δεκαετίες για την επέκταση

της εγκυρότητας και της εφαρμογής των πρότυπων Boussinesq εξισώσεων σε βαθύτερα νερά. Ως εκ τούτου έχουν παρουσιαστεί αρκετές εναλλακτικές διατυπώσεις με βελτιωμένες ιδιότητες στη περιγραφή φαινομένων διασποράς και μη γραμμικότητας. Ιδιαίτερης προσοχής έχουν λάβει οι εκτεταμένες Boussinesq εξισώσεις του Nwogu [23], οι οποίες μοντελοποιούν την διάδοση κυματισμών από βαθιά σε ρηχά νερά. Τις τελευταίες δεκαετίες έχουν αναπτυχθεί διάφορα μοντέλα τύπου Boussinesq, όπως είναι οι εξισώσεις των Madsen και Sørensen [22] και οι εξισώσεις του Nwogu [23]. Οι εξισώσεις του Nwogu είναι διατυπωμένες χρησιμοποιώντας την ελεύθερη επιφάνεια του νερού και την οριζόντια ταχύτητα σε ένα ορισμένο βάθος το οποίο έχει διαλεχθεί έτσι ώστε να γίνεται ελαχιστοποίηση του σφάλματος διάδοσης κυματισμών με βάση τη γραμμική θεωρία.

Για την αριθμητική επίλυση των εξισώσεων τύπου Boussinesq μέχρι πρόσφατα οι πεπερασμένες διαφορές ήταν η κύρια μέθοδος επίλυσης [14, 22, 23]. Η ευρεία χρήση αυτής της μεθόδου μπορεί να δικαιολογηθεί χάριν της ευκολίας με την οποία οι παράγωγοι υψηλότερης τάξης μπορούν να προσεγγιστούν και λόγω των καλά δομημένων γραμμικών συστημάτων τα οποία καταλήγουν. Τα συστήματα αυτά μπορούν πολύ εύκολα να επιλυθούν. Όμως το βασικό μειονέκτημα στις διδιάστατες προσομοιώσεις είναι ότι μόνο δομημένα πλέγματα χρησιμοποιούνται ακόμα και για ακανόνιστα φυσικά χωρία. Για να χρησιμοποιήσουμε μη δομημένα πλέγματα μπορούμε να επιλέξουμε ανάμεσα στις μεθόδους πεπερασμένων στοιχείων [35, 11, 9] και στις μεθόδους πεπερασμένων όγκων [33, 10, 19, 27, 28]. Το μεγάλο μειονέκτημα της μεθοδολογίας των πεπερασμένων στοιχείων είναι ότι συνήθως οι υψηλότερης τάξης παράγωγοι που είναι παρούσες και σε χαμηλότερης τάξης Boussinesq εξισώσεις πρέπει να εξαλειφθούν. Αυτό συνήθως γίνεται λύνοντας μία επιπρόσθετη εξίσωση. Οι μέθοδοι πεπερασμένων όγκων συνήθως απαιτούν λιγότερο υπολογιστικό κόπο σε σχέση με την μέθοδο πεπερασμένων στοιχείων, ενώ οι μη γραμμικοί όροι μεταφοράς και οι όροι τοπογραφίας μπορούν να αντιμετωπιστούν πιο εύκολα σε σχέση με την μέθοδο πεπερασμένων διαφορών.

Για να οδηγηθούμε σε ένα μοντέλο που είναι σχετικά εύκολα επιλύσιμο αριθμητικά

κάνουμε την παραδοχή ότι η μέση στάθμη του νερού είναι μικρή σε σχέση με την καμπυλότητα της ελεύθερης επιφάνειας. Έτσι λαμβάνουμε τις πιο διαδεδομένες εξισώσεις υπερβολικού τύπου, που περιγράφουν ροές ελεύθερης επιφάνειας, τις εξισώσεις ρηχών υδάτων. Όμως ο βασικός περιορισμός αυτών είναι ότι δεν εφαρμόζονται σε ενδιάμεσου βάθους και σε βαθύτερα νερά διότι η διασπορά επιδρά στην ροή του υγρού. Για αυτό το λόγο χρησιμοποιούνται εξισώσεις τύπου Boussinesq που περιέχουν και όρους διασποράς. Παρακάτω παρουσιάζονται οι εξισώσεις τύπου Boussinesq του Nwogu οι οποίες εφαρμόζονται για διάδοση κυμάτων σε αργά μεταβαλλόμενη βαθυμετρία.



Σχήμα 2.1: Ροή με ελεύθερο σύνορο υπό την επίδραση της βαρύτητας.

2.1 Μαθηματικό Μοντέλο

Ο Nwogu [23] παρήγαγε ένα σύστημα εξισώσεων χρησιμοποιώντας την μέση τιμή των ταχυτήτων $\mathbf{u} \equiv \mathbf{u}_a$ σε μία αυθαίρετη απόσταση z_a από ένα σταθερό βάθος h (μέση στάθμη) με μεταβλητή ταχύτητα αντί να χρησιμοποιηθεί η μέση τιμή κατά το βάθος. Ως βέλτιστη τιμή χρησιμοποιήθηκε η $z_a = -0.531h$, έτσι ώστε οι ιδιότητες της διασποράς του συστήματος να προσεγγίζουν την γραμμική θεωρία κύματος κάνοντας τις εξισώσεις να εφαρμόζονται σε μεγαλύτερο φάσμα βάθους σε σύγκριση με τις κλασσικές εξισώσεις

Boussinesq, [25]. Οι εξισώσεις του Nwogu δίνονται ως εξής:

$$\begin{aligned} \eta_t + (Hu)_x + (Hv)_y + \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h(u_x + v_y)_x + \left(z_a + \frac{h}{2} \right) h((hu)_x + (hv)_y)_x \right]_x \\ + \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h(u_x + v_y)_y + \left(z_a + \frac{h}{2} \right) h((hu)_x + (hv)_y)_y \right]_y = 0 \end{aligned} \quad (2.1)$$

$$u_t + g\eta_x + uu_x + vu_y + \left[\frac{z_a^2}{2}(u_{xx} + v_{yx}) + z_a((hu)_{xx} + (hv)_{yx}) \right]_t = 0 \quad (2.2)$$

$$v_t + g\eta_y + uv_x + vv_y + \left[\frac{z_a^2}{2}(u_{xy} + v_{yy}) + z_a((hu)_{xy} + (hv)_{yy}) \right]_t = 0 \quad (2.3)$$

Οι δείκτες $\mathbf{x} = [x, y] \in \mathbb{R}^2$ και $t \in \mathbb{R}^+$ ορίζουν μερικές παραγώγους του χώρου και του χρόνου, το $H(\mathbf{x}, t) = \eta(\mathbf{x}, t) + h(\mathbf{x})$ είναι το συνολικό βάθος του υγρού (με $\eta(\mathbf{x}, t)$ να είναι η ελεύθερη επιφάνεια ανύψωσης και $h(\mathbf{x})$ τη μέση σταθερή στάθμη), $\mathbf{u}(\mathbf{x}, t) = [u, v]^T$ είναι η μέση οριζόντια ταχύτητα και ως g ορίζουμε την επιτάχυνση της βαρύτητας.

Αναδιατάσσοντας τις παραπάνω εξισώσεις και κάνοντας υπολογισμούς προσπαθούμε να δημιουργήσουμε την συντηρητική μορφή ενός νόμου διατήρησης. Ακολουθώντας τις εργασίες των [19], [33] και [27] λαμβάνουμε:

$$\mathbf{U}_t + \nabla \cdot \mathcal{H}(\mathbf{U}) = \mathbf{S}(\mathbf{U}) \quad \text{στο} \quad \Omega \times [0, t] \subset \mathbb{R}^2 \times \mathbb{R}^+, \quad (2.4)$$

όπου $\Omega \times [0, t]$ είναι το χωρίο του Καρτεσιανού χρόνου-χώρου μέσα στο οποίο περιλαμβάνονται οι λύσεις. $\mathbf{U}$ είναι το διάνυσμα των νέων μεταβλητών, με $\mathcal{H} = [\mathbf{F}, \mathbf{G}]$ να είναι τα μη γραμμικά διανύσματα ροής (fluxes) και $\mathbf{S} = \mathbf{S}_b + \mathbf{S}_d$ οι πηγαίοι όροι που περιλαμβάνουν την τοπογραφία $\mathbf{S}_b$ και τους όρους διασποράς $\mathbf{S}_d$, που μας δίνουν

$$\mathbf{U} = \begin{bmatrix} H \\ P_1 \\ P_2 \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} Hu \\ Hu^2 + \frac{1}{2}gH^2 \\ Huv \end{bmatrix},$$

$$\mathbf{G}(\mathbf{U}) = \begin{bmatrix} Hv \\ Huv \\ Hv^2 + \frac{1}{2}gH^2 \end{bmatrix}, \quad \mathbf{S}_b = \begin{bmatrix} 0 \\ -gHb_x \\ -gHb_y \end{bmatrix}, \quad \mathbf{S}_d = \begin{bmatrix} -\psi_c \\ -u\psi_c + \psi_{M_x} \\ -v\psi_c + \psi_{M_y} \end{bmatrix},$$

όπου

$$P_1 = Hu + H\mathcal{A},$$

$$P_2 = Hv + H\mathcal{B},$$

ή σε διανυσματική μορφή

$$\mathbf{P} = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = H \begin{bmatrix} \frac{z_a^2}{2} \nabla(\nabla \cdot \mathbf{u}) + z_a \nabla(\nabla \cdot h\mathbf{u}) + \mathbf{u} \end{bmatrix},$$

$$\psi_c = \mathcal{C} + \mathcal{D} = \nabla \cdot \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h \nabla(\nabla \cdot \mathbf{u}) + \left(z_a + \frac{h}{2} \right) h \nabla(\nabla \cdot h\mathbf{u}) \right]$$

και

$$\psi_{\mathbf{M}} = \begin{bmatrix} \psi_{M_x} \\ \psi_{M_y} \end{bmatrix} = \begin{bmatrix} H_t \mathcal{A} \\ H_t \mathcal{B} \end{bmatrix} = H_t \frac{z_a^2}{2} \nabla(\nabla \cdot \mathbf{u}) + H_t z_a \nabla(\nabla \cdot h\mathbf{u}),$$

με

$$\mathcal{A} = \left(\frac{z_a^2}{2} (u_{xx} + v_{yx}) + z_a ((hu)_{xx} + (hv)_{yx}) \right),$$

$$\mathcal{B} = \left(\frac{z_a^2}{2} (u_{xy} + v_{yy}) + z_a ((hu)_{xy} + (hv)_{yy}) \right),$$

$$\mathcal{C} = \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h (u_x + v_y)_x + \left(z_a + \frac{h}{2} \right) h ((hu)_x + (hv)_y)_x \right]_x$$

και

$$\mathcal{D} = \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h (u_x + v_y)_y + \left(z_a + \frac{h}{2} \right) h ((hu)_x + (hv)_y)_y \right]_y.$$

Η εξίσωση (2.4) περιέχει τους ίδιους όρους ροής, όπως και οι μη γραμμικές εξισώσεις ριγών υδάτων. Το διάνυσμα $\mathbf{P}$ περιέχει όλους τους όρους διασποράς με χρονικές παραγώγους, ενώ το διάνυσμα διασποράς $\mathbf{S_d}$ περιλαμβάνει μονάχα χωρικές παραγώγους. Το σχήμα πεπερασμένων όγκων που αναπτύχθηκε στις εργασίες [33] βασίζεται στην ολοκληρωτική μορφή της (2.4) σε ένα γενικό υπολογιστικό χωρίο Ω :

$$\frac{\partial}{\partial t} \iint_{\Omega} \mathbf{U} d\Omega + \iint_{\Omega} (\nabla \cdot \mathcal{H}) d\Omega = \iint_{\Omega} \mathbf{S} d\Omega. \quad (2.5)$$

Εφαρμόζοντας το θεώρημα απόκλισης του Gauss στο ολοκλήρωμα ροής έχουμε:

$$\frac{\partial}{\partial t} \iint_{\Omega} \mathbf{U} d\Omega + \oint_{\Gamma} (\mathcal{H} \cdot \tilde{\mathbf{n}}) d\Gamma = \iint_{\Omega} \mathbf{S} d\Omega, \quad (2.6)$$

όπου ορίζουμε Γ να είναι το σύνορο του Ω και $\tilde{\mathbf{n}} = [\tilde{n}_x, \tilde{n}_y]^T$ είναι το μοναδιαίο κάθετο διάνυσμα. Μία προσέγγιση της εξίσωσης (2.6) εφαρμόζεται σε ένα όγκο ελέγχου (υπολογιστικό κελί) $C_P \subset \Omega$ έτσι ώστε το ολοκλήρωμα του όγκου στην περιοχή $|C_P|$ να αναπαριστά ολοκλήρωση του κέντρου ελέγχου και το επιφανειακό ολοκλήρωμα να αναπαριστά την συνολική ροή (μή γραμμικούς και πηγαίους όρους) διαμέσου των συνόρων του C_P . Ορίζοντας

$$\mathbf{U}_P = \frac{1}{|C_P|} \iint_{C_P} \mathbf{U} d\Omega.$$

την μέση τιμή των συντηρητικών ποσοτήτων του κέντρου ελέγχου σε μία δεδομένη χρονική στιγμή, λαμβάνουμε για κάθε υπολογιστικό κελί

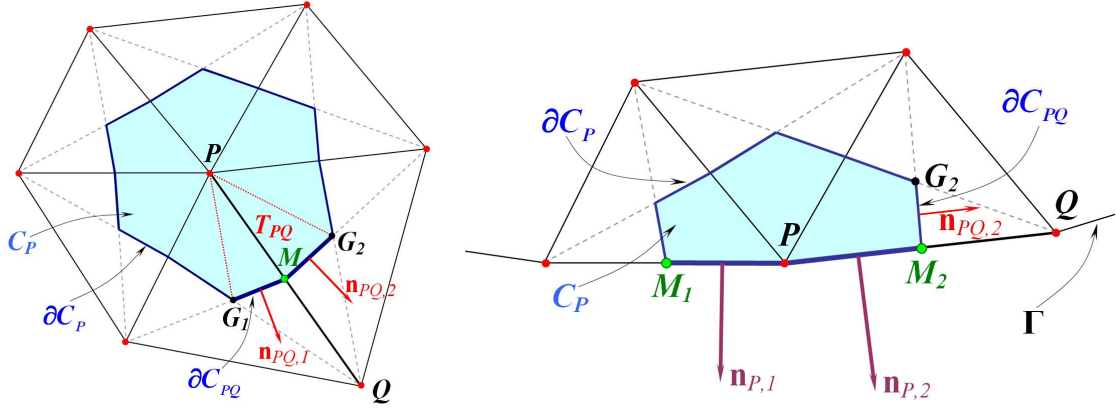
$$\frac{\partial \mathbf{U}_P}{\partial t} = -\frac{1}{|C_P|} \oint_{\Gamma} (\mathbf{F}\tilde{n}_x + \mathbf{G}\tilde{n}_y) d\Gamma + \frac{1}{|C_P|} \iint_{\Omega} \mathbf{S} d\Omega. \quad (2.7)$$

2.2 Αριθμητικό Σχήμα Πεπερασμένων Όγκων

Η γενική προσέγγιση των πεπερασμένων όγκων απαιτεί τον διαχωρισμό του υπολογιστικού χωρίου $\Omega \subset \mathbb{R}^2$ σε ένα σύνολο μη επικαλυπτόμενων όγκων ελέγχου και την αριθμητική διακριτοποίηση της ολοκληρωτικής μορφής της εξίσωσης (2.4) σε κάθε υπολογιστικό κελί. Τα διακριτά σημεία του χώρου, ονομάζονται σημεία δεδομένων (data points) (P) και είναι οι κορυφές των τριγώνων ενώ τα σύνορα των όγκων ελέγχου λέγονται μέτωπα. Ο όρος ακμή (edge) αναφέρεται στην γραμμή που ενώνει τα γειτονικά data points. Ως $\mathcal{T}^{h_N}$ του Ω θέτουμε να είναι ένα σύνολο πεπερασμένων τριγωνικών υποσυνόλων $T_p \subset \Omega$, $p = 1, 2, \dots$, με h_N να είναι το χαρακτηριστικό μήκος τριγωνισμού που χρησιμοποιούμε για κάθε πλέγμα.

Τα κεντροθετημένα ανά κόμβο (node-centered) σχήματα, σύμφωνα με τις εργασίες [18] και [8] έχουν όγκους ελέγχου που σχηματίζονται από το πλέγμα που είναι δυϊκό στο αρχικό $\mathcal{T}^{h_N}$. Όπως φαίνεται στο Σχήμα 2.2, τα σύνορα ∂C_P των όγκων ελέγχου C_P γύρω από τον εσωτερικό κόμβο P ορίζονται αν συνδέσουμε τα βαρύκεντρα των τριγώνων που έχουν το P σαν κοινή κορυφή με τα μέσα των αντίστοιχων ακμών που έχουν ως κοινή

κορυφή τον κόμβο P .



Σχήμα 2.2: Όγκος ελέγχου για ένα εσωτερικό (αριστερά) και ένα συνοριακό κόμβο (δεξιά).

Ορίζουμε $\partial C_{PQ} = \partial C_P \cap \partial C_Q$ και M ως το μέσον της ακμής PQ . Το εξωτερικό κάθετο διάνυσμα στο ∂C_{PQ} είναι $\mathbf{n}_{PQ} = [n_{PQx}, n_{PQy}]^T$, ενώ $\tilde{\mathbf{n}}_{PQ} = [\tilde{n}_{PQx}, \tilde{n}_{PQy}]^T$ είναι το αντίστοιχο μοναδιαίο. Εάν $\mathbf{n}_{PQ,1}$ είναι κάθετο στο $\overline{G_1M}$ με νόρμα ίση με το μήκος $\overline{G_1M}$, και $\mathbf{n}_{PQ,2}$ κάθετο στο $\overline{MG_2}$, τότε:

$$\mathbf{n}_{PQ} = \int_{\partial C_{PQ}} \tilde{\mathbf{n}} dl = \mathbf{n}_{PQ,1} + \mathbf{n}_{PQ,2},$$

όπου dl είναι ένα μέτρο ∂C_{PQ} .

Για ένα κόμβο P στο σύνορο του πλέγματος ο ορισμός για τον όγκο ελέγχου περιγράφεται στο σχήμα 2.2. Το εξωτερικό κάθετο διάνυσμα $\overline{M_1PM_2}$ είναι $\mathbf{n}_P = [n_{Px}, n_{Py}]^T$, ενώ έχουμε $\tilde{\mathbf{n}}_P = [\tilde{n}_{Px}, \tilde{n}_{Py}]^T$ σύμφωνα με το μοναδιαίο κάθετο διάνυσμα. Εάν, $\mathbf{n}_{P,1}$ είναι κάθετο στο $\overline{M_1P}$ (έχοντας μία νόρμα ίση με το μήκος του $\overline{M_1P}$), ενώ $\mathbf{n}_{P,2}$ είναι κάθετο στο $\overline{PM_2}$ όταν:

$$\mathbf{n}_P = \int_{\overline{M_1PM_2}} \tilde{\mathbf{n}} dl = \mathbf{n}_{P,1} + \mathbf{n}_{P,2},$$

όπου dl υπολογίζεται μέσω $\overline{M_1 P M_2}$. Επίσης, στην περίπτωση του συνοριακού κόμβου P , $\partial C_{PQ} = \overline{M_2 G_2}$ και εάν $\mathbf{n}_{PQ,2}$ είναι κάθετο στο $\overline{M G_2}$ τότε :

$$\mathbf{n}_{PQ} = \int_{\partial C_{PQ}} \tilde{\mathbf{n}} dl = \mathbf{n}_{PQ,2},$$

όπου dl υπολογίζεται μέσω ∂C_{PQ} .

Από την εξίσωση (2.7), εάν έχουμε ότι το Γ είναι το κυρίως σύνορο και

$$K_P := \{Q \in \mathbb{N} \mid \partial C_P \cap \partial C_Q \neq \emptyset\}$$

το σύνολο των γειτονικών κόμβων P , τότε ∂C_P για ένα συνοριακό κόμβο περιγράφεται ως εξής :

$$\partial C_P = \bigcup_{Q \in K_P} \partial C_{PQ} + (\partial C_P \cap \Gamma).$$

Η ολοκληρωτική μορφή των εξισώσεων δίνει :

$$\begin{aligned} \frac{\partial \mathbf{U}_P}{\partial t} = & -\frac{1}{|C_P|} \sum_{Q \in K_P} \left\{ \int_{\partial C_{PQ}} (\mathbf{F} \tilde{n}_x + \mathbf{G} \tilde{n}_y) dl \right\} - \frac{1}{|C_P|} \int_{\partial C_P \cap \Gamma} (\mathbf{F} \tilde{n}_x + \mathbf{G} \tilde{n}_y) dl \quad (2.8) \\ & + \frac{1}{|C_P|} \sum_{Q \in K_P} \left\{ \iint_{C_P} \mathbf{S}_b dx dy \right\}, \end{aligned}$$

με διανύσματα ροής :

$$\Phi_{PQ} = \int_{\partial C_{PQ}} (\mathbf{F} \tilde{n}_x + \mathbf{G} \tilde{n}_y) dl \quad \text{και} \quad \Phi_{P,\Gamma} = \int_{\partial C_P \cap \Gamma} (\mathbf{F} \tilde{n}_x + \mathbf{G} \tilde{n}_y) dl.$$

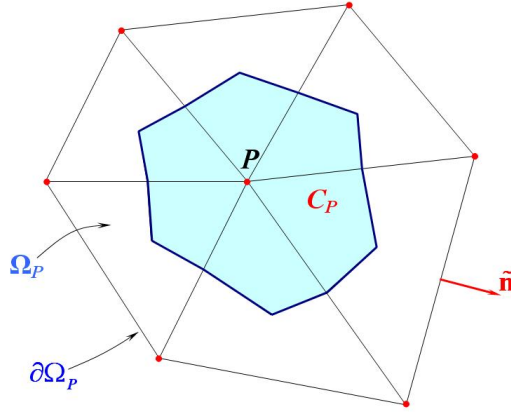
Οπότε η εξίσωση (2.8) γίνεται :

$$\frac{\partial \mathbf{U}_P}{\partial t} = -\frac{1}{|C_P|} \sum_{Q \in K_P} \Phi_{PQ} - \frac{1}{|C_P|} \Phi_{P,\Gamma} + \frac{1}{|C_P|} \sum_{Q \in K_P} \left\{ \iint_{C_P} \mathbf{S}_b dx dy \right\}, \quad (2.9)$$

Για κάθε ακμή του μη δομημένου πλέγματος τα μη γραμμικά διανύσματα της ροής Φ_{PQ} πρέπει να υπολογίζονται και να προστίθενται στο άθροισμα των ροών των γειτονικών

κελιών C_P και C_Q αντίστοιχα. Μία θεμελιώδης ιδέα για τη μέθοδο των πεπερασμένων όγκων είναι αυτή της αντικατάστασης της ροής στα μέτωπα των όγκων ελέγχου με τη συνάρτηση αριθμητικής ροής. Γενικά, αυτή η συνάρτηση υπολογίζεται σαν μία ακριβής ή προσεγγιστική λύση ενός προβλήματος Riemann. Στην έρευνα των [33] χρησιμοποιήθηκε ο προσεγγιστικός επιλύτης Riemann του Roe [29].

Για την επίτευξη χωρικής ακρίβειας μεγαλύτερης του ένα στο αριθμητικό σχήμα χρειάζεται ο υπολογισμός των κλίσεων των μεταβλητών. Η μέση τιμή των κλίσεων μίας μεταβλητής w (*gradient*) $(\nabla w)_P$ σε ένα κελί μπορεί να υπολογιστεί σε κάθε κόμβο πλέγματος P εφαρμόζοντας τα θεωρήματα των Green-Gauss στο χωρίο Ω_P , όπως φαίνεται στο σχήμα 2.3, όπου περιγράφονται από το σύνολο όλων των τριγώνων που μοιράζονται την κορυφή P . Ακολουθώντας τους [3, 4, 5], υπολογίζουμε το ολοκλήρωμα και υποθέτουμε ότι η διακριτή λύση μιας ποσότητας του w μεταβάλλεται γραμμικά, δηλαδή η κλίση είναι σταθερή στο Ω_P . Έστω,



Σχήμα 2.3: Ορισμός του χωρίου όπου γίνεται αριθμητικός υπολογισμός της κλίσης και της απόκλισης.

$$\iint_{\Omega_P} \nabla w dA = \oint_{\partial\Omega_P} w \tilde{\mathbf{n}} dl,$$

από το οποίο λαμβάνουμε :

$$(\nabla w)_P = \frac{1}{|C_P|} \sum_{Q \in K_P} \frac{1}{2} (w_P + w_Q) \mathbf{n}_{PQ}. \quad (2.10)$$

Στην περίπτωση συνοριακού κόμβου P (σχήμα 2.2) η προηγούμενη εξίσωση παίρνει την μορφή :

$$(\nabla w)_P = \frac{1}{|C_P|} \left\{ \sum_{Q \in K_P} \frac{1}{2} (w_P + w_Q) \mathbf{n}_{PQ} + w_P (\mathbf{n}_{P,1} + \mathbf{n}_{P,2}) \right\}. \quad (2.11)$$

Στη συνέχεια παρουσιάζεται ο υπολογισμός της μέσης τιμής της απόκλισης για το διάνυσμα της ταχύτητας για το χωρίο Ω_P . Ο υπολογισμός αυτός θα χρησιμοποιηθεί αργότερα για τη διακριτοποίηση των όρων διασποράς. Για ακόμα μια φορά χρησιμοποιώντας το θεώρημα απόκλισης και προσεγγίζοντας τα επικαμπύλια ολοκληρώματα με το κανόνα του τραπεζίου παίρνουμε :

$$\int_{\Omega_P} \nabla \cdot \mathbf{u} d\Omega = \oint_{\partial\Omega_P} \mathbf{u} \cdot \tilde{\mathbf{n}} dl = \sum_{Q \in K_P} \frac{3}{2} (\mathbf{u}_P + \mathbf{u}_Q) \cdot \mathbf{n}_{PQ}. \quad (2.12)$$

Άρα η μέση τιμή της απόκλισης σε ένα κόμβο P δίδεται από :

$$(\nabla \cdot \mathbf{u})_P = \frac{3}{|\Omega_P|} \sum_{Q \in K_P} \frac{1}{2} (\mathbf{u}_P + \mathbf{u}_Q) \cdot \mathbf{n}_{PQ} = \frac{1}{|C_P|} \sum_{Q \in K_P} \frac{1}{2} (\mathbf{u}_P + \mathbf{u}_Q) \cdot \mathbf{n}_{PQ}, \quad (2.13)$$

λαμβάνοντας υπόψιν ότι $|C_P| = \frac{1}{3} |\Omega_P|$.

Στη περίπτωση όπου ο κόμβος είναι συνοριακός η εξίσωση (2.12) παίρνει τη μορφή :

$$(\nabla \cdot \mathbf{u})_P = \frac{1}{|C_P|} \left\{ \sum_{Q \in K_P} \frac{1}{2} (\mathbf{u}_P + \mathbf{u}_Q) \cdot \mathbf{n}_{PQ} + \mathbf{u}_P \cdot (\mathbf{n}_{P,1} + \mathbf{n}_{P,2}) \right\}. \quad (2.14)$$

Η εξίσωση (2.8) περιέχει το ολοκλήρωμα με τον όρο διασποράς ψ_C στον πηγαίο όρο $\mathbf{S}_d$. Για την διακριτοποίηση του όρου χρησιμοποιούμε το θεώρημα απόκλισης, με το

οποίο έχουμε :

$$(\psi_c)_P = \frac{1}{|C_P|} \iint_{C_P} \psi_c d\Omega \quad (2.15)$$

$$= \frac{1}{|C_P|} \iint_{C_P} \nabla \cdot \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h \nabla (\nabla \cdot \mathbf{u}) + \left(z_a + \frac{h}{2} \right) h \nabla (\nabla \cdot h \mathbf{u}) \right] \quad (2.16)$$

$$= \frac{1}{|C_P|} \oint_{\partial C_P} \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h \nabla (\nabla \cdot \mathbf{u}) + \left(z_a + \frac{h}{2} \right) h \nabla (\nabla \cdot h \mathbf{u}) \right] \cdot \tilde{\mathbf{n}} dl \quad (2.17)$$

$$= \frac{1}{|C_P|} \sum_{Q \in K_P} \left\{ \int_{\partial C_{PQ}} \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h \nabla (\nabla \cdot \mathbf{u}) \right] \cdot \tilde{\mathbf{n}} dl + \int_{\partial C_{PQ}} \left[\left(z_a + \frac{h}{2} \right) h \nabla (\nabla \cdot h \mathbf{u}) \right] \cdot \tilde{\mathbf{n}} dl \right\}$$

Όπως και στην αριθμητική ροή Φ_{PQ} , και για όλες τις ακμές, τα παραπάνω ολοκληρώματα πρέπει να υπολογιστούν και έπειτα να προστεθούν στο κάθε υπολογιστικό κελί C_P , C_Q με το αντίστοιχο πρόσημο.

Για τον υπολογισμό του $(\psi_c)_P$ απαιτείται η προσέγγιση των δύο επικαμπύλιων ολοκληρωμάτων τα οποία υπολογίζονται χρησιμοποιώντας τον κανόνα του μέσου. Πιο συγκεκριμένα :

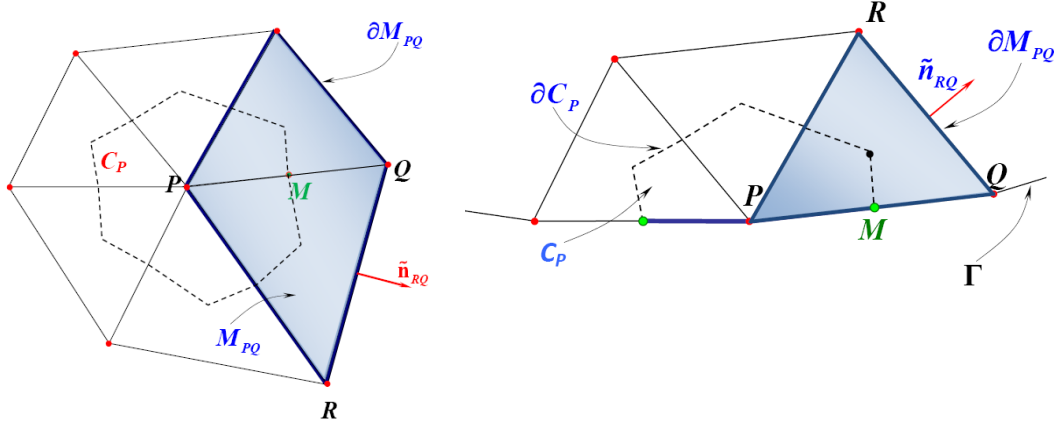
$$\int_{\partial C_{PQ}} \left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h \nabla (\nabla \cdot \mathbf{u}) \cdot \tilde{\mathbf{n}} dl \approx \left[\left(\frac{z_a^2}{2} - \frac{h^2}{6} \right) h \right]_M [\nabla (\nabla \cdot \mathbf{u}) \cdot \mathbf{n}_{PQ}]_M, \quad (2.18)$$

$$\int_{\partial C_{PQ}} \left(z_a + \frac{h}{2} \right) h \nabla (\nabla \cdot h \mathbf{u}) \cdot \tilde{\mathbf{n}} dl \approx \left[\left(z_a + \frac{h}{2} \right) h \right]_M [\nabla (\nabla \cdot h \mathbf{u}) \cdot \mathbf{n}_{PQ}]_M. \quad (2.19)$$

Οι παραπάνω όροι απαιτούν τον υπολογισμό της κλίσης της απόκλισης του διανύσματος της ταχύτητας $\mathbf{u}$ και της ποσότητας $h\mathbf{u}$ κατά μήκος των ενδιάμεσων σημείων M . Γενικά, ο υπολογισμός της κλίσης μίας ποσότητας w στο M απαιτεί τον ορισμό ενός νέου υπολογιστικού κελιού που δομείται από την ένωση των δύο τριγώνων τα οποία έχουν κοινή την ακμή PQ (βλέπε σχήμα. 2.4).

Για τον υπολογισμό της κλίσης μιας ποσότητας w στο νέο υπολογιστικό κελί, ορίζουμε :

$$K_{PQ} := \{R \in \mathbb{N} \mid R \text{ είναι μία κορυφή } M_{PQ} \text{ και } RQ \in \partial M_{PQ}\},$$



Σχήμα 2.4: Υπολογιστικά κελιά για την κλίση της απόκλισης : εσωτερικό κελί (αριστερά) και συνοριακό κελί (δεξιά).

και χρησιμοποιώντας το θεώρημα Green-Gauss στο M_{PQ} λαμβάνουμε :

$$\iint_{M_{PQ}} \nabla w d\Omega = \oint_{\partial M_{PQ}} w \tilde{\mathbf{n}}_{RQ} dl = \sum_{\substack{R, Q \in K_{PQ} \\ R \neq Q}} \frac{1}{2} (w_R + w_Q) \mathbf{n}_{RQ} \quad (2.20)$$

που οδηγεί στην εξίσωση

$$(\nabla w)_M = \frac{1}{|M_{PQ}|} \sum_{\substack{R, Q \in K_{PQ} \\ R \neq Q}} \frac{1}{2} (w_R + w_Q) \mathbf{n}_{RQ}, \quad (2.21)$$

όπου $\mathbf{n}_{RQ}$ το κάθετο διάνυσμα στην ακμή RQ . Στην περίπτωση των συνοριακών κόμβων το M_{PQ} μειώνεται στο τρίγωνο PQR (βλέπε Σχήμα 3) και η υπολογιστική τιμή από (2.21) αναθέτεται στο M . Εν κατακλείδι, για να υπολογίσουμε τους όρους $[\nabla(\nabla \cdot \mathbf{u})]_M$ και $[\nabla(\nabla \cdot h\mathbf{u})]_M$ στις εξισώσεις (2.18) και (2.19) χρησιμοποιούμε την (2.20) για την κλίση και για τον υπολογισμό των $(\nabla \cdot \mathbf{u})_R$ και $(\nabla \cdot h\mathbf{u})_R$ στους κόμβους $R \in K_{PQ}$ χρησιμοποιούμε τη φόρμουλα (2.13).

Στη συνέχεια, για τους πηγαίους όρους διασποράς στην εξίσωση ορμής έχουμε :

$$\frac{1}{|C_P|} \iint_{C_P} -\mathbf{u} \psi_c + \psi_M d\Omega = -\mathbf{u}_P \iint_{C_P} \psi_c d\Omega + \frac{1}{|C_P|} \iint_{C_P} \psi_M d\Omega. \quad (2.22)$$

Ο πρώτος όρος στο δεξί μέλος της εξίσωσης διακριτοποιείται όπως παρουσιάστηκε προη-

γουμενώς και ο δεύτερος όρος παίρνει την διακριτή μορφή:

$$(\psi_{\mathbf{M}})_P = \quad (2.23)$$

$$= \frac{1}{|C_P|} \iint_{C_P} \psi_{\mathbf{M}} d\Omega = \frac{1}{|C_P|} \iint_{C_P} H_t \frac{z_a^2}{2} \nabla(\nabla \cdot \mathbf{u}) + H_t z_a \nabla(\nabla \cdot h\mathbf{u}) d\Omega \quad (2.24)$$

$$= \frac{1}{|C_P|} \iint_{C_P} H_t \frac{z_a^2}{2} \nabla(\nabla \cdot \mathbf{u}) d\Omega + \frac{1}{|C_P|} \iint_{C_P} H_t z_a \nabla(\nabla \cdot h\mathbf{u}) d\Omega \quad (2.25)$$

$$\approx \left[H_t \frac{z_a^2}{2} \right]_P |C_P| [\nabla(\nabla \cdot \mathbf{u})]_P + [H_t z_a]_P |C_P| [\nabla(\nabla \cdot h\mathbf{u})]_P, \quad (2.26)$$

όπου οι διακριτές αποκλίσεις $(\nabla \cdot \mathbf{u})_P$ και $(\nabla \cdot h\mathbf{u})_P$ υπολογίζονται ξανά χρησιμοποιώντας την φόρμουλα (2.13). Με τη χωρική διακριτοποίηση πεπερασμένων όγκων των μη γραμμικών όρων της τοπογραφίας και των πηγαίων όρων διασποράς παίρνουμε την μορφή της εξίσωσης

$$\frac{\partial \mathbf{U}_P}{\partial t} = \mathcal{L}(\mathbf{U}).$$

όπου $\mathcal{L}(\mathbf{U})$ ο διακριτός χωρικός τελεστής. Για το συνολικό διακριτό πλέγμα, έχουμε να διακριτοποιήσουμε τον τελεστή που εξελίσσεται στο χρόνο. Στην εργασία των [33], η βέλτιστη υπό την έννοια της CFL συνθήκης, η τρίτης τάξης ρητή μέθοδος Strong Stability Preserving Runge-Kutta (SSP-RK) υιοθετήθηκε. Αυτή η μέθοδος είναι γνωστή ως τρίτης τάξης TVD Runge-Kutta και περιγράφεται ως εξής:

$$\begin{aligned} \mathbf{U}_P^{(1)} &= \mathbf{U}_P^{(n)} + \Delta t^n \mathcal{L}(\mathbf{U}^{(n)}); \\ \mathbf{U}_P^{(2)} &= \frac{3}{4} \mathbf{U}_P^{(n)} + \frac{3}{4} \mathbf{U}_P^{(1)} + \Delta t^n \frac{1}{4} \mathcal{L}(\mathbf{U}^{(1)}); \\ \mathbf{U}_P^{(n+1)} &= \frac{1}{3} \mathbf{U}_P^{(n)} + \frac{2}{3} \mathbf{U}_P^{(2)} + \Delta t^n \frac{2}{3} \mathcal{L}(\mathbf{U}^{(2)}); \end{aligned}$$

όπου $\Delta t^n = t^{n+1} - t^n$ είναι το χρονικό βήμα.

2.3 Γραμμικό Σύστημα

Σε κάθε βήμα του χρονικού σχήματος ένα γραμμικό σύστημα $\mathbf{M}\mathbf{u} = \mathbf{P}$, $\mathbf{M} \in \mathbb{R}^{2N \times 2N}$, πρέπει να λυθεί για να λάβουμε τις ταχύτητες u, v όπου $\mathbf{P} = [P_1 \ P_2]^T$, το διάνυσμα των

συντελεστών των μεταβλητών των εξισώσεων ορμής. Ο πίνακας $\mathbf{M}$ προκύπτει από την διακριτοποίηση του $\mathbf{P}$. Είναι ένας αραιός πίνακας και η δομή του εξαρτάται από το πλέγμα. Πιο συγκεκριμένα από τον ορισμό του διανύσματος $\mathbf{P}$ για κάθε κόμβο $\mathbf{P}$ και χρονικό βήμα i στη Runge-Kutta έχουμε:

$$H_P^{(i)} \left[\frac{z_a^2}{2} \nabla(\nabla \cdot \mathbf{u}) + z_a \nabla(\nabla \cdot h\mathbf{u}) + \mathbf{u} \right]_P^{(i)} = \mathbf{P}_P^{(i)}, \quad i = 1, 2, n+1. \quad (2.27)$$

Τώρα, η διακριτή κλίση $\nabla(\nabla \cdot \mathbf{u})_P$ και $\nabla(\nabla \cdot h\mathbf{u})_P$ χρειάζεται να υπολογιστεί βάση των διακριτοποιήσεων:

$$H_P \left[\frac{(z_a^2)_P}{2} \frac{1}{|C_P|} \sum_{Q \in K_P} (\nabla \cdot \mathbf{u})_M \mathbf{n}_{PQ} + \frac{(z_a)_P}{|C_P|} \sum_{Q \in K_P} (\nabla \cdot h\mathbf{u})_M \mathbf{n}_{PQ} + \mathbf{u}_P \right] = \mathbf{P}_P. \quad (2.28)$$

Υπολογίζουμε ποσότητες $\nabla \cdot \mathbf{u}$ και $\nabla \cdot h\mathbf{u}$ στο σημείο M , σύμφωνα με το σχήμα 2.4, τα νέα υπολογιστικά κελιά μπορούν να χρησιμοποιηθούν ξανά. Έναν όμοιο υπολογισμό κάναμε στην $(\nabla \cdot h\mathbf{u})$. Η σχέση του αριστερού και του δεξιού μέλους της εξίσωσης (2.28) μας δίνει:

$$\begin{aligned} \frac{(z_a^2)_P}{2} \frac{1}{|C_P|} \sum_{Q \in K_P} \left[\frac{1}{|M_{PQ}|} \sum_{\substack{R, Q \in K_{PQ} \\ R \neq Q}} \frac{1}{2} (\mathbf{u}_R + \mathbf{u}_Q) \cdot \mathbf{n}_{RQ} \right] \mathbf{n}_{PQ} = \\ \frac{(z_a^2)_P}{2|C_P|} \sum_{Q \in K_P} \left[\frac{1}{2|M_{PQ}|} \sum_{\substack{R, Q \in K_{PQ} \\ R \neq Q}} \mathbf{u}_R \cdot (\mathbf{n}_{PR} + \mathbf{n}_{RQ}) \right] \mathbf{n}_{PQ}, \end{aligned} \quad (2.29)$$

που μπορεί να γραφεί ως:

$$\frac{(z_a^2)_P}{2|C_P|} \sum_{Q \in K_P} \left[\frac{1}{2|M_{PQ}|} \left(\sum_{\substack{R, Q \in K_P \cap K_{PQ} \\ R \neq Q}} \mathbf{u}_R \cdot (\mathbf{n}_{PR} + \mathbf{n}_{RQ}) + \mathbf{u}_P \cdot (\mathbf{n}_{SP} + \mathbf{n}_{PR}) \right) \right] \mathbf{n}_{PQ},$$

ή

$$\frac{(z_a^2)_P}{2} \frac{1}{|C_P|} \sum_{Q \in K_P} (\nabla \cdot \mathbf{u})_M \mathbf{n}_{PQ} = \frac{(z_a^2)_P}{2|C_P|} \sum_{Q \in K_P} (\mathbf{A}_{Qx} u_Q + \mathbf{A}_{Qy} v_Q + \mathbf{A}_{Px} u_P + \mathbf{A}_{Py} v_P) \quad (2.30)$$

όπου $\mathbf{A}_{Qx}, \mathbf{A}_{Qy}, \mathbf{A}_{Px}, \mathbf{A}_{Py} \in \mathbb{R}^2$ και εξαρτώνται από τα γεωμετρικά χαρακτηριστικά των ποσοτήτων $\mathbf{n}_{PQ}$, το κάθετο διάνυσμα στις ακμές του ∂M_{PQ} και στο εμβαδό $|M_{PQ}|$.

Οπότε το αραιό γραμμικό σύστημα που προκύπτει έχει την μορφή:

$$\begin{aligned} & \frac{(z_a^2)_P}{2|C_P|} \sum_{Q \in K_P} ([\mathbf{A}_{Qx} \mathbf{A}_{Qy}] \mathbf{u}_Q + [\mathbf{A}_{Px} \mathbf{A}_{Py}] \mathbf{u}_P) \\ & + \frac{(z_a)_P}{|C_P|} \sum_{Q \in K_P} ([\mathbf{B}_{Qx} \mathbf{B}_{Qy}] \mathbf{u}_Q + [\mathbf{B}_{Px} \mathbf{B}_{Py}] \mathbf{u}_P) + \mathbf{I} \mathbf{u}_P \\ & = \frac{1}{H_P} \mathbf{P}_P, \quad P = 1 \dots N, \end{aligned} \quad (2.31)$$

Όπου κάθε δύο συνεχόμενες γραμμές του συστήματος του πίνακα $\mathbf{M}$, που αντιστοιχούν στον κόμβο P , έχουν μη μηδενικά στοιχεία τα οποία αντιστοιχούν στους συντελεστές των αγνώστων των ταχυτήτων στον κόμβο P και στους γειτονικούς τους $Q \in K_P$. Πιο συγκεκριμένα οι στήλες των 2×2 πινάκων (2.31) δίνονται από:

$$\mathbf{A}_{Qx} = \frac{1}{2|M_{PQ}|} \sum_{R, Q \in K_P \cap K_{PQ}} (n_{PRx} + n_{RQx}) \mathbf{n}_{PQ}, \quad (2.32)$$

$$\mathbf{A}_{Qy} = \frac{1}{2|M_{PQ}|} \sum_{R, Q \in K_P \cap K_{PQ}} (n_{PRy} + n_{RQy}) \mathbf{n}_{PQ}, \quad (2.33)$$

$$\mathbf{B}_{Qx} = \frac{1}{2|M_{PQ}|} \sum_{R, Q \in K_P \cap K_{PQ}} h_R (n_{PRx} + n_{RQx}) \mathbf{n}_{PQ}, \quad \mathbf{B}_{Qy} = \frac{1}{2|M_{PQ}|} \sum_{R, Q \in K_P \cap K_{PQ}} h_R (n_{PRy} + n_{RQy}) \mathbf{n}_{PQ},$$

$$\mathbf{A}_{Px} = \frac{1}{2|M_{PQ}|} (n_{SPx} + n_{PRx}) \mathbf{n}_{PQ}, \quad \mathbf{A}_{Py} = \frac{1}{2|M_{PQ}|} (n_{SPy} + n_{PRy}) \mathbf{n}_{PQ},$$

$$\mathbf{B}_{Px} = \frac{1}{2|M_{PQ}|} h_P (n_{SPx} + n_{PRx}) \mathbf{n}_{PQ}, \quad \mathbf{B}_{Py} = \frac{1}{2|M_{PQ}|} h_P (n_{SPy} + n_{PRy}) \mathbf{n}_{PQ}.$$

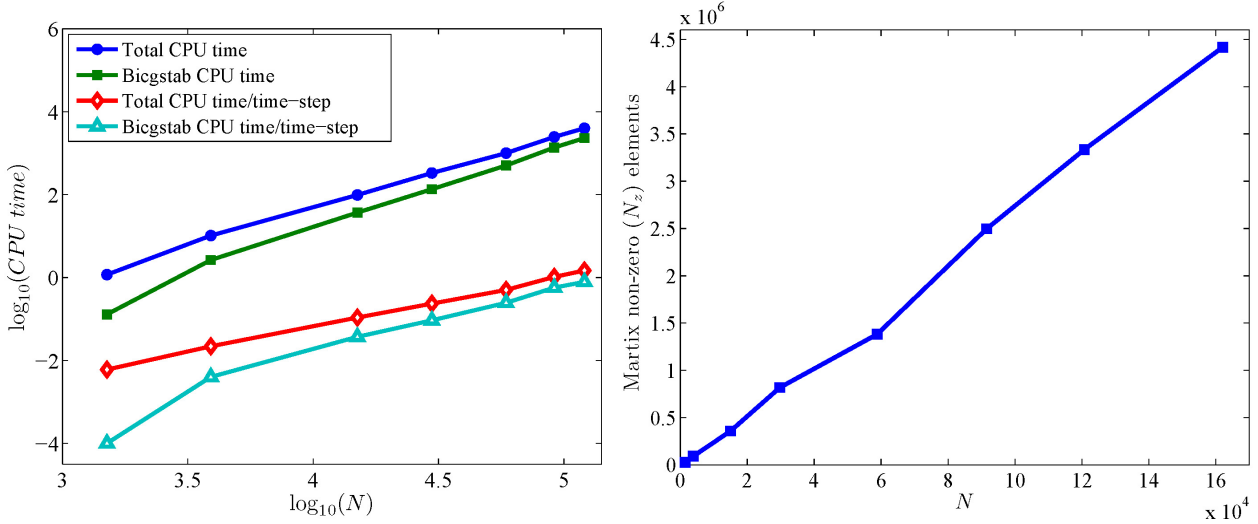
Τα μέγιστα μη μηδενικά στοιχεία του πίνακα $\mathbf{M}$ σε κάθε γραμμή P στην (2.31) είναι διπλάσια του αριθμού των γειτονικών κόμβων του P συν ένα.

Ο πίνακας διάστασης $2N \times 2N$ που προκύπτει από το γραμμικό σύστημα (2.31) είναι αραιός και έχει σχεδόν συμμετρική δομή (structural symmetric) αλλά δεν είναι συμμετρικός. Η δομή του εξαρτάται από το πλέγμα, την φυσική κατάσταση του προβλήματος και τον αριθμό των κόμβων που χρησιμοποιούμε.

Στην εργασία των [33] έγινε μία προσέγγιση στην επίλυση του γραμμικού συστήματος και ήταν ως εξής: Για την αποθήκευση του πίνακα χρησιμοποιήθηκε η compressed sparse row (CSR) μορφή αποθήκευσης (Saad) [30] αραιών πινάκων. Η CSR μορφή αποθήκευσης καθώς και οι διάφορες μορφές αποθήκευσης που υπάρχουν αναφέρονται και περιγράφονται αναλυτικά στο Κεφάλαιο 2. Για την επίλυση του γραμμικού συστήματος χρησιμοποιήθηκε η Bi-Conjugate Gradient Stabilized Method (BiCGStab) επαναληπτική μέθοδος για τη λύση των μη συμμετρικών συστημάτων και σαν αρχική προσέγγιση x^0 χρησιμοποιείται η αριθμητική λύση των ταχυτήτων στο προηγούμενο χρονικό βήμα και ως ανοχή του υπολοίπου (residual error tolerance) το $5 \cdot 10^{-6}$ με χρήση του προρρυθμιστή ILUT. Στο σχήμα 2.5 παρουσιάζεται μία σχετική σύγκριση των αποτελεσμάτων που προκύπτουν με την αύξηση του αριθμού των κόμβων N και των μη μηδενικών στοιχείων N_z του αραιού γραμμικού συστήματος. Όπως ήταν αναμενόμενο, ο αριθμός των μη μηδενικών στοιχείων αυξάνεται γραμμικά σε σχέση με το N . Ο χρόνος που χρειάζεται η Bi-Conjugate Gradient Stabilized Method ανά χρονικό βήμα είναι της τάξεως $O(n^{3/2})$, ενώ ο συνολικός χρόνος ανά χρονικό βήμα αυξάνεται με τάξη $O(n^{5/4})$. Όπως φαίνεται από το σχήμα λοιπόν, η επίλυση του γραμμικού συστήματος με την επαναληπτική μέθοδο BiCGStab είναι η πιο χρονοβόρα διαδικασία του αλγορίθμου. Στο κεφάλαιο 3 γίνεται ανάλυση των διαφόρων επαναληπτικών μεθόδων και προρρυθμιστών όπως η ILUT καθώς και των παραμέτρων που τις επηρεάζουν.

Στην παρούσα μεταπτυχιακή εργασία θα γίνει εφαρμογή, σύγκριση και μελέτη μεταξύ των διαφόρων επαναληπτικών μεθόδων επίλυσης του πίνακα που προκύπτει με σκοπό τη βελτίωση του χρόνου επίλυσης του, καθώς μεταβάλλονται διάφορα μεγέθη όπως το μέγεθος του πλέγματος (ο αριθμός των κόμβων) και κάποια φυσικά μεγέθη (πχ. μέση

στάθμη). Ταυτόχρονα, θα χρησιμοποιηθούν διάφοροι προρρυθμιστές (preconditioners) και αναδιάταξη (reorderings). Αυτό θα γίνει έτσι ώστε να βρούμε τον καλύτερο συνδυασμό σε χρόνο και ακρίβεια.



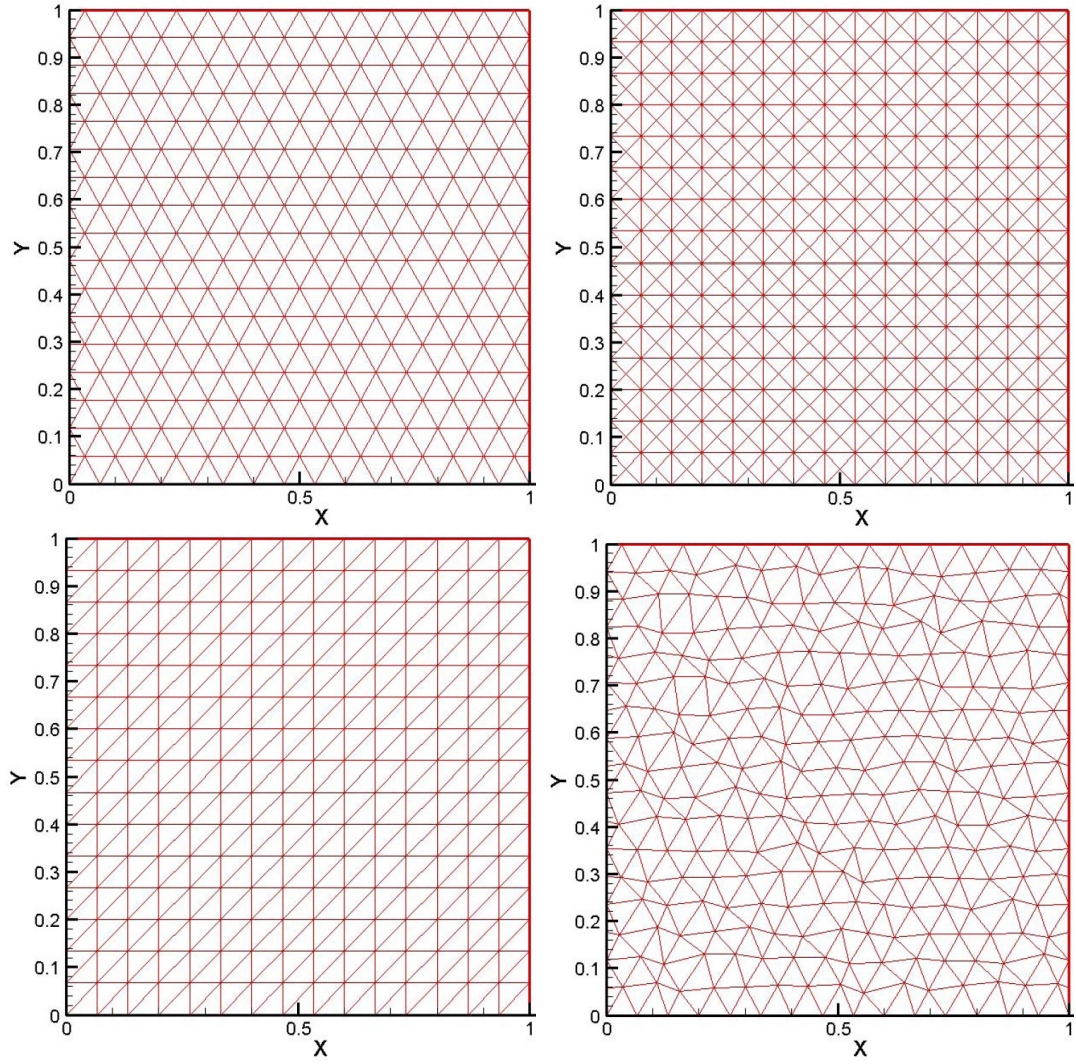
Σχήμα 2.5: Τα μη μηδενικά στοιχεία (αριστερά) και ο CPU χρόνος (δεξιά) σε συνάρτηση με το N [33].

2.4 Υπολογιστικά Πλέγματα και Δείκτης κατάστασης

Τα πλέγματα (grids) που μελετήθηκαν στη συγκεκριμένη μεταπτυχιακή εργασία, είναι αυτά που χρησιμοποιούν στην έρευνα τους οι [18]. Μπορούμε να τα διακρίνουμε σε δομημένα (regular) και μη δομημένα (irregular) όπως φαίνονται στο Σχήμα 2.6 και είναι τα εξής:

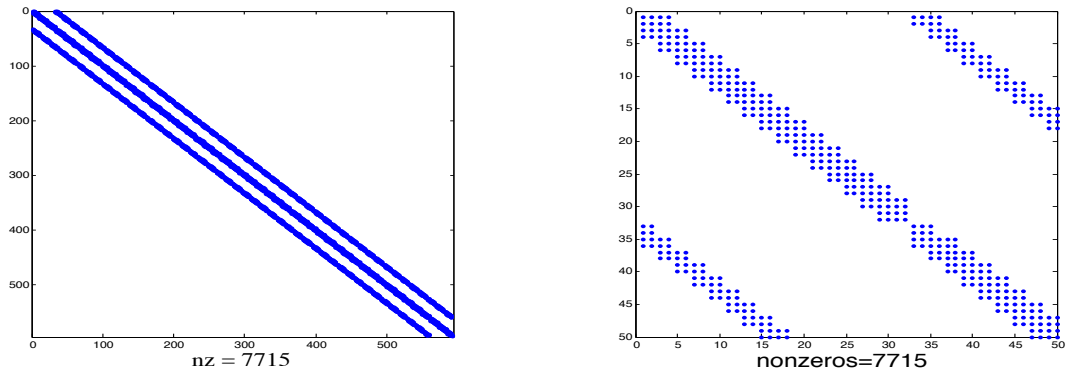
- (1) Equilateral triangular (τύπου I)
- (2) Orthogonal (τύπου II),
- (3) Orthogonal (τύπου III),
- (4) Distorted (τύπου IV).

Το equilateral triangular (ισόπλευρο τύπου I) είναι γνωστό και ως isometric. Τα πλέγματα Distorted (τύπου IV) προέρχονται από τυχαίες μετακινήσεις των κόμβων κατά $0.4r\Delta x$ με $r \in [-1/2, 1/2]$ και Δx είναι το μήκος της πλευράς του τριγώνου στον άξονα x .

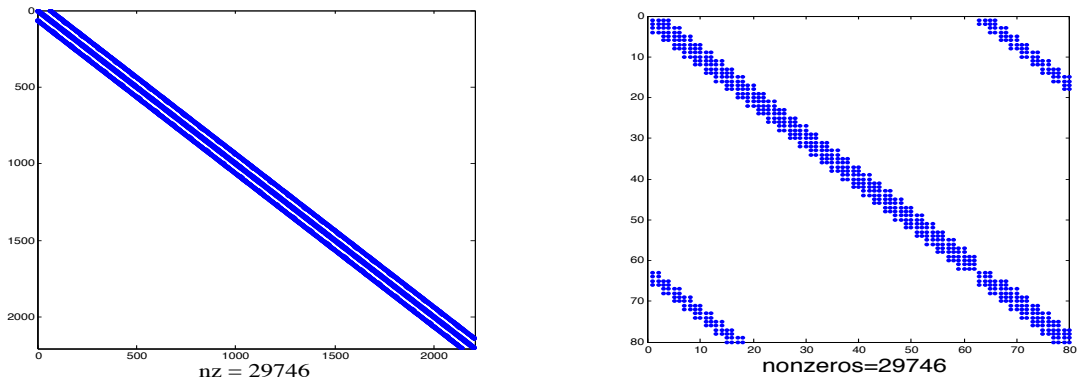


Σχήμα 2.6: Αντιπροσωπευτικοί τύποι πλεγμάτων.

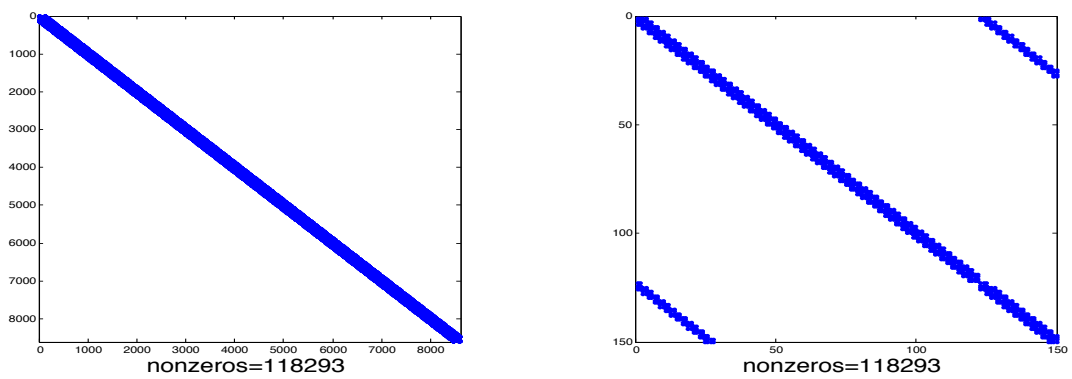
Στα Σχήματα 2.7-2.9 παρουσιάζεται η δομή των πινάκων που αναφέραμε στη παράγραφο 2.3 που προκύπτει από ένα Distorted πλέγμα (τύπου IV). Στην αριστερή στήλη φαίνονται οι μορφές των πινάκων για κόμβους $N=297$ ($N_x = 15$) με N_x το σύνολο των κόμβων στον άξονα τον x , $N=1102$ ($N_x = 30$) και $N=4305$ ($N_x = 60$). Στη δεξιά στήλη έχει γίνει μεγένθυση ώστε να φανεί καλύτερα η κατανομή των στοιχείων του πίνακα που προκύπτει για κάθε περίπτωση. Στην συνέχεια, έχει γίνει η ίδια παρουσίαση για ένα equilateral πλέγμα στα Σχήματα 2.10-2.12 (τύπου I), ένα Orthogonal (τύπου II) στα Σχήματα 2.13-2.15 και Orthogonal (τύπου III) στα Σχήματα 2.16-2.18.



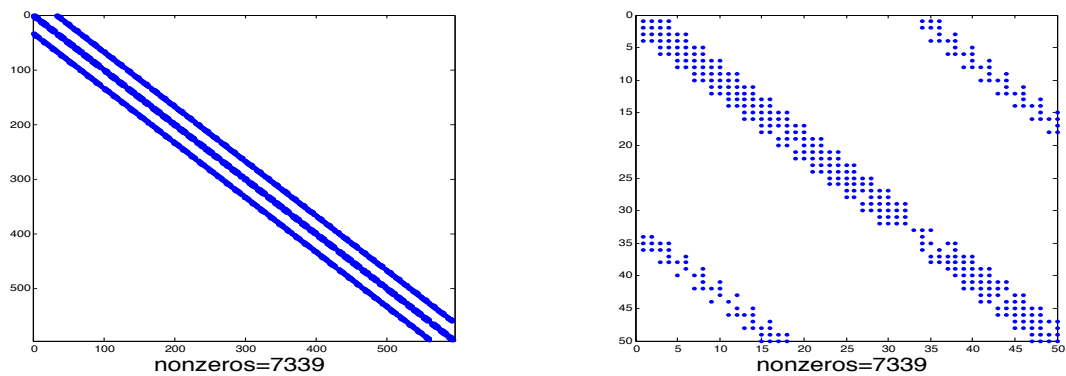
Σχήμα 2.7: Δομή πίνακα που προκύπτει από Distorted πλέγμα με $N_x=15$, $N = 297$.



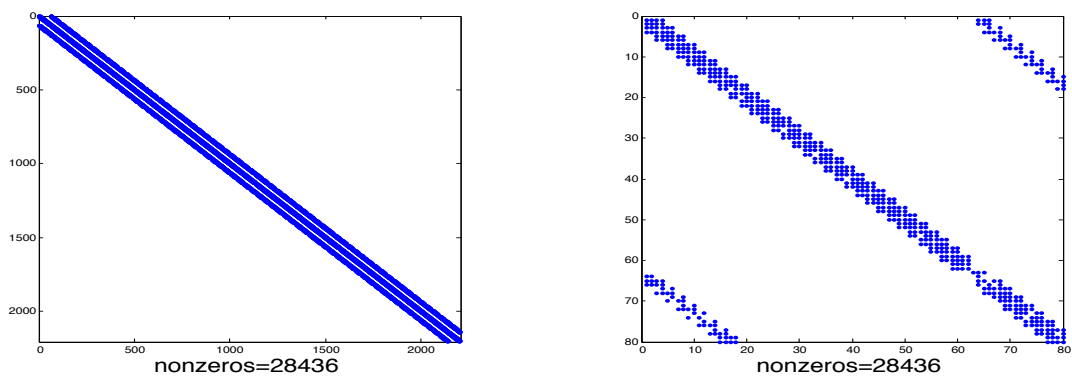
Σχήμα 2.8: Δομή πίνακα που προκύπτει από Distorted πλέγμα με $N_x = 30$, $N = 1102$.



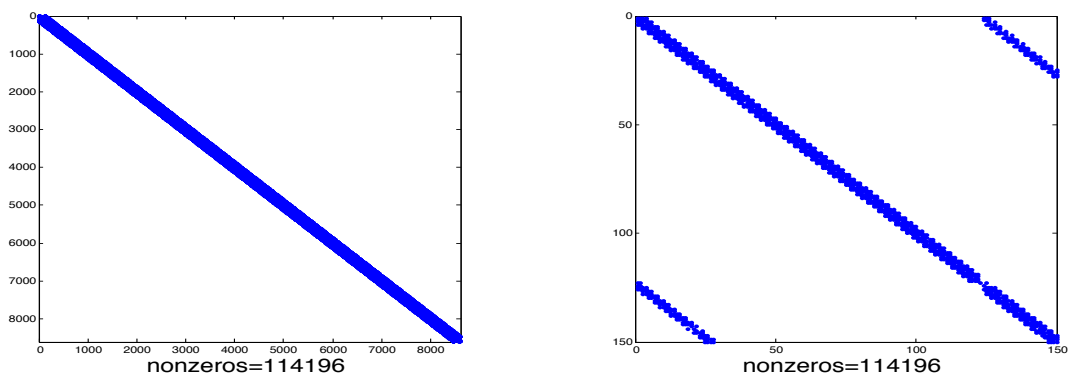
Σχήμα 2.9: Δομή πίνακα που προκύπτει από Distorted πλέγμα με $N_x = 60$, $N = 4305$.



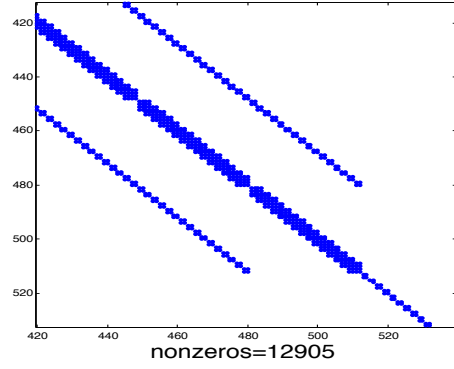
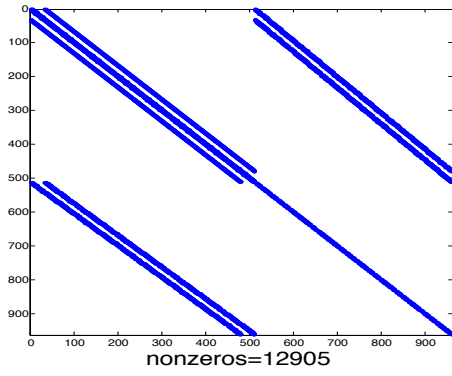
Σχήμα 2.10: Δομή πίνακα που προκύπτει από Equilateral πλέγμα με $N_x=15$, $N = 297$.



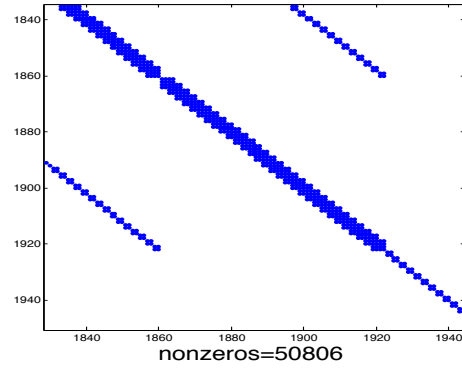
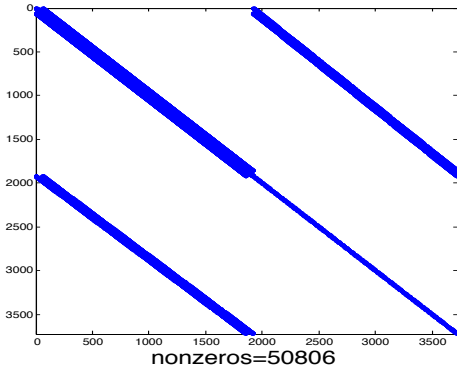
Σχήμα 2.11: Δομή πίνακα που προκύπτει από Equilateral πλέγμα με $N_x = 30$, $N = 1102$.



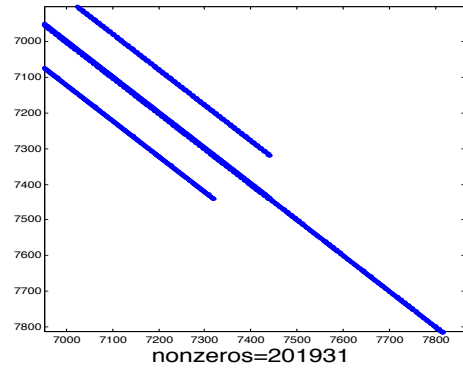
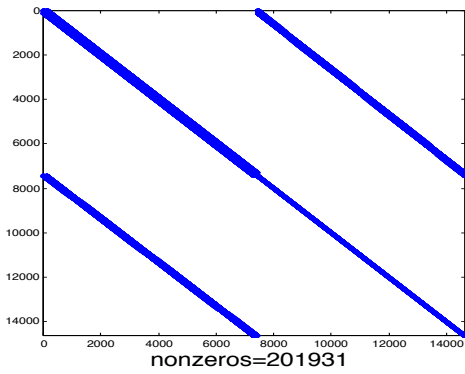
Σχήμα 2.12: Δομή πίνακα που προκύπτει από Equilateral πλέγμα με $N_x = 60$, $N = 4305$.



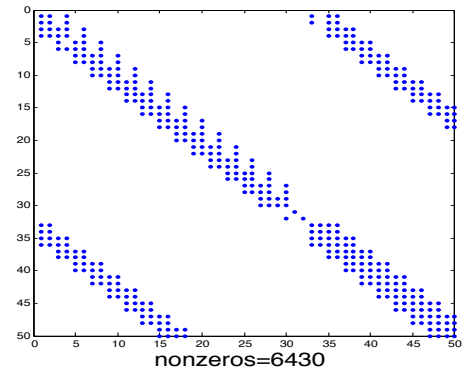
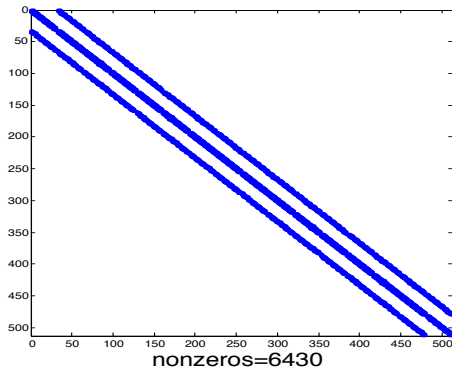
Σχήμα 2.13: Δομή πίνακα που προκύπτει από Orthogonal (type I) πλέγμα με $N_x = 15$, $N = 481$.



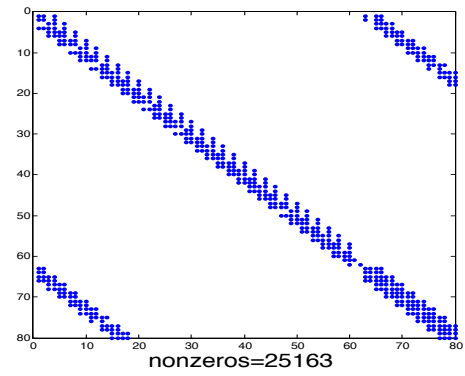
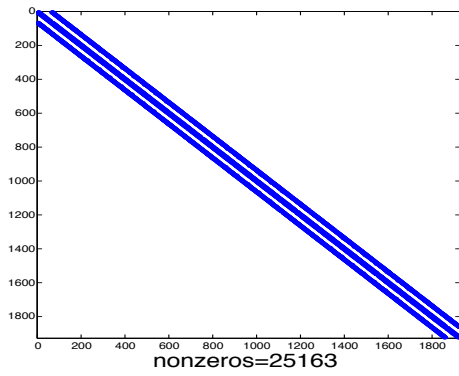
Σχήμα 2.14: Δομή πίνακα που προκύπτει από Orthogonal (type I) πλέγμα με $N_x = 30$, $N = 1861$.



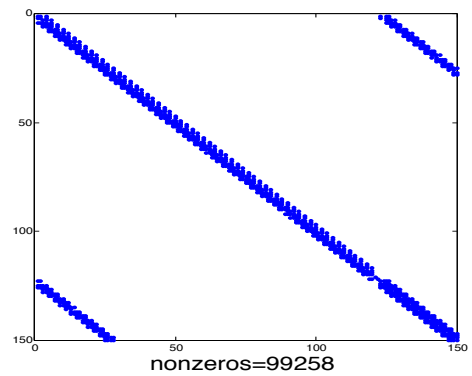
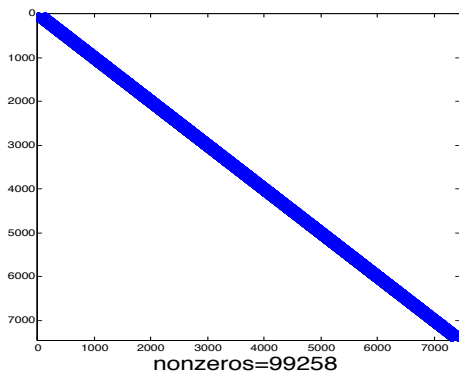
Σχήμα 2.15: Δομή πίνακα που προκύπτει από Orthogonal (type I) πλέγμα με $N_x = 60$, $N = 7321$.



Σχήμα 2.16: Δομή πίνακα που προκύπτει από Orthogonal (type II) $N_x = 15$, $N = 256$.



Σχήμα 2.17: Δομή πίνακα που προκύπτει από Orthogonal (type II) $N_x = 30$, $N = 961$.



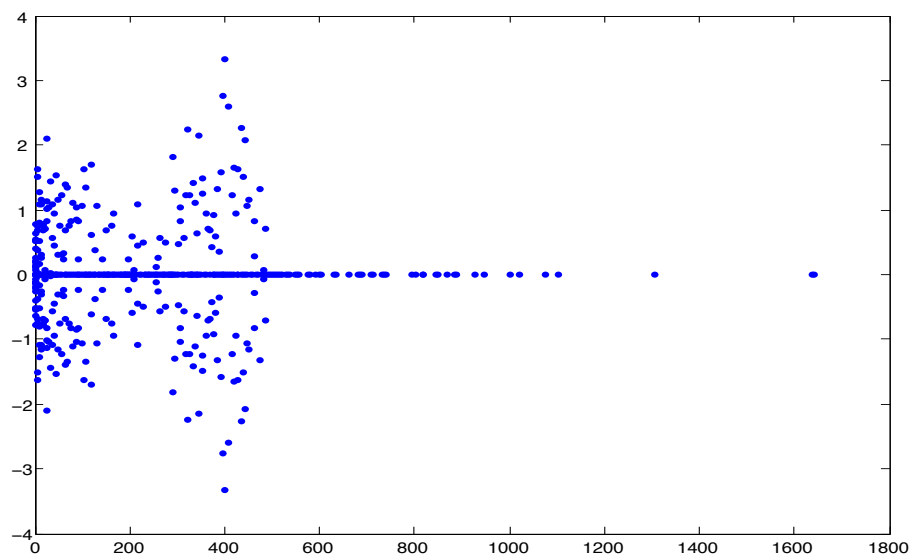
Σχήμα 2.18: Δομή πίνακα που προκύπτει από Orthogonal (type II) πλέγμα με $N_x = 15$, $N = 37114$.

Επίσης, στα Σχήματα 2.19-2.24 παρουσιάζονται οι ιδιοτιμές που προκύπτουν από τα 4 διαφορετικά είδη πλέγματος που χρησιμοποιούμε, όπου παρατηρούμε το φάσμα, δηλαδή το σύνολο των ιδιοτιμών και το πως κατανέμονται αυτές στο μιγαδικό επίπεδο. Στα σχήματα των ιδιοτιμών του πίνακα Orthogonal (τύπου II) παρατηρούμε ότι το φάσμα των ιδιοτιμών είναι όλο πραγματικό (βλέπε Σχήματα 2.22 και 2.23). Ενώ, στα υπόλοιπα σχήματα οι ιδιοτιμές ανήκουν στο μιγαδικό επίπεδο. Μελετάμε τις ιδιοτιμές του πίνακα έτσι ώστε από την μορφή τους να ελέγξουμε εάν χρειάζεται προρρύθμιση και εάν όχι. Ένα άλλο κριτήριο για την προρρύθμιση που παρουσιάζεται είναι ο δείκτης κατάστασης.

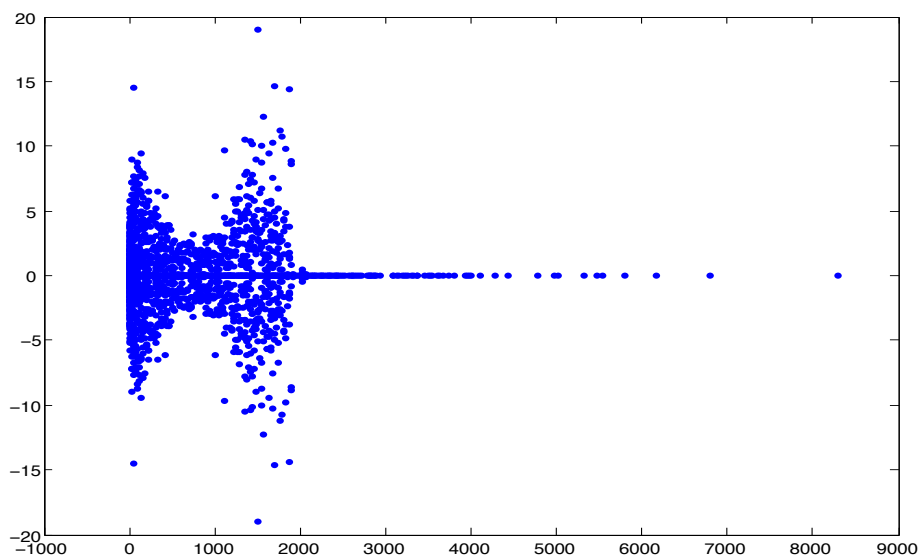
Ορίζουμε ως δείκτη κατάστασης $k(A)$ condition number του πίνακα A ως προς τη νόρμα $\|\cdot\|$, τον αριθμό που είναι πάντοτε μεγαλύτερος ή ίσος με την μονάδα γιατί

$$1 = \|I\| = \|A^{-1}A\| \leq \|A^{-1}\| \|A\| = k(A).$$

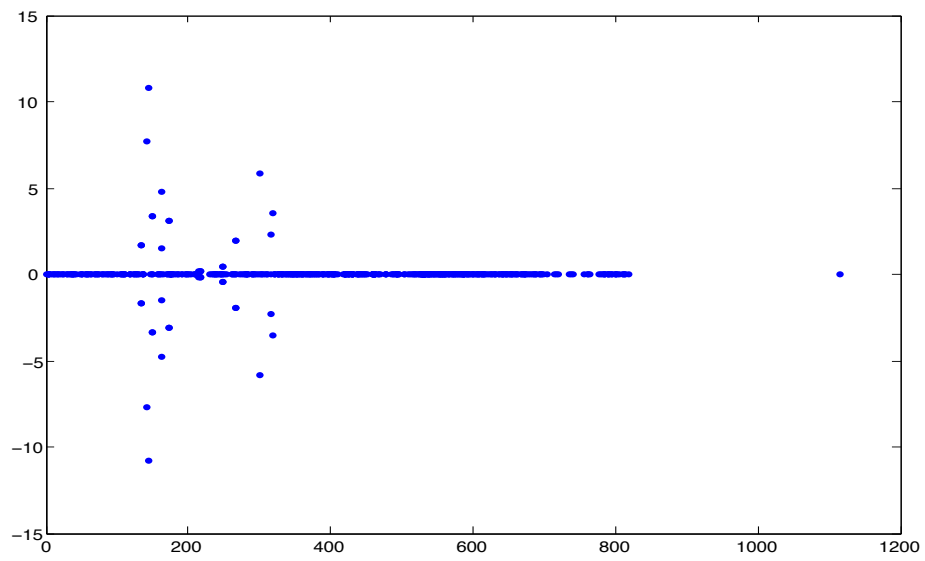
Για να υπολογίσουμε τον δείκτη κατάστασης στην πράξη καλό θα ήταν να αποφύγουμε τον υπολογισμό του A^{-1} , όπου είναι δαπανηρός για την επίλυση του συστήματος. Στην πράξη λοιπόν δεν υπολογίζουμε ακριβώς τον δείκτη κατάστασης αλλά τον εκτιμούμε κατά προσέγγιση, αφού μονάχα την τάξη μεγέθους του συστήματος χρειαζόμαστε για να συμπεράνουμε αν έχει καλή ή κακή κατάσταση. Γενικά, λέμε ότι ο πίνακας (ή το σύστημα $Ax = b$) θα έχει κακή κατάσταση, δηλαδή κακό δείκτη κατάστασης αν $k(A) \gg 1$. Στην περίπτωση που έχουμε κακό δείκτη κατάστασης μία μικρή μεταβολή του διανύσματος b , μπορεί να προκαλέσει μεταβολή στο σύστημα. Το πρόβλημα που παρουσιάστηκε στην παρούσα εργασία είναι ότι οι πίνακες που προκύπτουν από την 2.31 έχουν πολύ κακό δείκτη κατάστασης. Για παράδειγμα για τον πίνακα που φαίνεται στο σχήμα 2.6 που προκύπτει από ένα distorted πλέγμα έχει δείκτη κατάστασης 7,7732E+05, ενώ για τον πίνακα του σχήματος 2.9 του equilateral πλέγματος ο δείκτης που προκύπτει είναι 5,5348E+03.



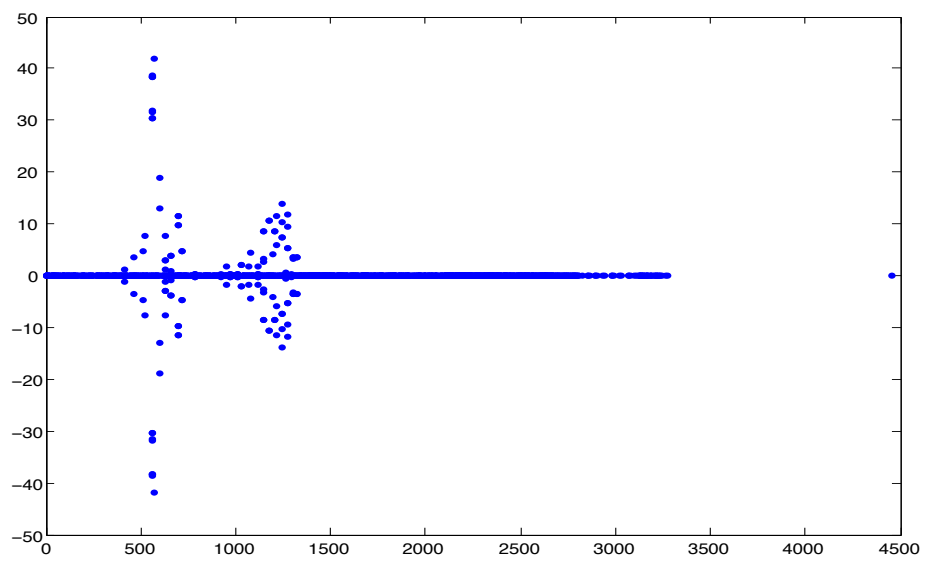
Σχήμα 2.19: Ιδιοτιμές του Distorted πίνακα $N = 297$



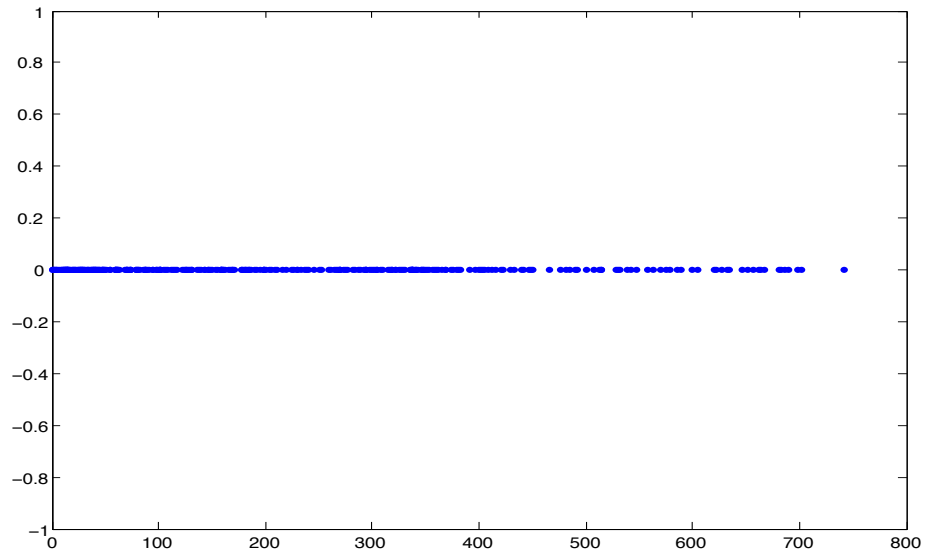
Σχήμα 2.20: Ιδιοτιμές του Distorted πίνακα $N = 1102$



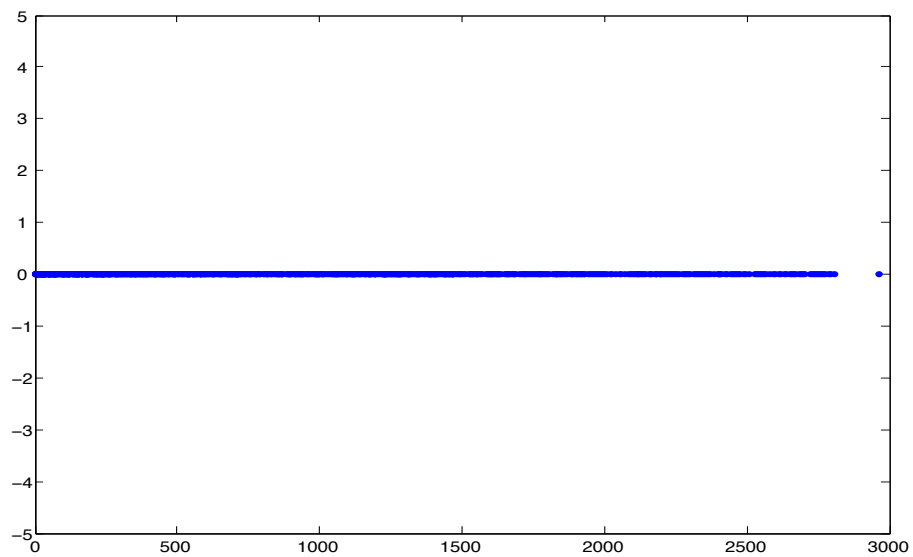
Σχήμα 2.21: Ιδιοτιμές του Orthogonal (type I) πίνακα $N = 481$



Σχήμα 2.22: Ιδιοτιμές του Orthogonal (type I) πίνακα $N = 1861$



Σχήμα 2.23: Ιδιοτιμές του Orthogonal (type II) πίνακα $N = 256$



Σχήμα 2.24: Ιδιοτιμές του Orthogonal (type II) πίνακα $N = 961$

Κεφάλαιο 3

Αραιοί Πίνακες

Αραιός θεωρείται ένας πίνακας όταν το πλήθος των μηδενικών στοιχείων είναι μεγαλύτερο από το πλήθος των μη μηδενικών. Τέτοιοι πίνακες προκύπτουν κατά την αριθμητική επίλυση Συνήθων και Μερικών Διαφορικών Εξισώσεων. Για την ελαχιστοποίηση του χώρου αποθήκευσης τους με αποφυγή αποθήκευσης των μηδενικών στοιχείων, δημιουργήθηκαν ειδικές μορφές αποθήκευσης όπως η COO (Coordinate format), η CSR (Compressed Sparse Row) η CSC (Compressed Sparse Column) και η MSR (Modified Sparse Row) μορφή [30]. Οι αραιές μορφές αποθήκευσης τους εξοικονομούν υπολογιστική μνήμη καθώς και χρόνο εκτέλεσης πράξεων. Οι παραπάνω μορφές αποθήκευσης αραιών πινάκων χρησιμοποιούνται σε αρκετές βιβλιοθήκες και λογισμικά, όπως στο Sparskit και στο ITSOL του Yousef Saad [30]. Οι βιβλιοθήκες λογισμικού (software) που θα μπορούσαν να χρησιμοποιηθούν είναι το ITSOL [1], η οποία διαθέτει υλοποίηση σε γλώσσα προγραμματισμού C και παρέχει πρόσθετους προορυθμιστές (ILUK, ILUT, ILUC, VBI-LUK, VBILUT) για την επίλυση γενικών αραιών γραμμικών συστημάτων και το Sparskit [2] που περιλαμβάνει υλοποιήσεις αριθμητικών διαδικασιών της γλώσσας προγραμματισμού Fortran. Το τελευταίο, περιέχει πολλά χρήσιμα εργαλεία για την επίλυση αραιών συστημάτων, καθώς και βασικές διαδικασίες αλγεβρικών πράξεων μεταξύ αραιών πινάκων. Παρακάτω παρουσιάζονται αναλυτικά οι διάφορες μορφές για την αποθήκευση αραιών πινάκων καθώς και παραδείγματα για την κάθε μία περίπτωση.

3.1 Αραιές μορφές αποθήκευσης

3.1.1 COO (Coordinate Format)

Ο κύριος σκοπός της coordinate format αποθήκευσης είναι να αποθηκευτούν μόνο τα μη μηδενικά στοιχεία ενός πίνακα εξετάζοντας τον πίνακα κατά γραμμές. Έστω ότι συμβολίζουμε το πλήθος των μη μηδενικών στοιχείων με Nz για ένα πραγματικό πίνακα διάστασης $N \times N$.

Σύμφωνα με τη μορφή αραιής αποθήκευσης coordinate format (COO) αποθηκεύουμε κάθε μη μηδενικό στοιχείο του πίνακα σε ένα διάνυσμα AA , Nz θέσεων, διατρέχοντας όλες τις γραμμές του πίνακα και αποθηκεύοντας ταυτόχρονα τη θέση του στοιχείου σε δύο διανύσματα jr, jc (γραμμή και στήλη αντίστοιχα), μεγέθους Nz .

Στο πρώτο διάνυσμα αποθηκεύονται τα μη μηδενικά στοιχεία του πίνακα, το δεύτερο είναι το διάνυσμα των στηλών και το τρίτο είναι το διάνυσμα που κρατάει τις γραμμές των μη μηδενικών στοιχείων. Χαρακτηριστικά της COO δομής είναι η εύκολη χρήση και κατανόηση της καθώς και το γεγονός ότι οι γραμμές και οι στήλες της δεν χρειάζεται να βρίσκονται σε συγκεκριμένη σειρά [30].

Για παράδειγμα στο πίνακα

$$A = \begin{pmatrix} 1. & 0. & 0. & 2. & 0. \\ 3. & 8. & 0. & 5. & 0. \\ 2. & 0. & 7. & 8. & 9. \\ 0. & 0. & 10. & 15. & 0. \\ 0. & 0. & 0. & 0. & 19. \end{pmatrix}$$

τα διανύσματα AA , jr , και jc θα είναι αντίστοιχα:

$$\begin{aligned} AA &= (1. \ 2. \ 3. \ 8. \ 5. \ 2. \ 7. \ 8. \ 9. \ 10. \ 15. \ 19.), \\ jr &= (1 \ 1 \ 2 \ 2 \ 2 \ 3 \ 3 \ 3 \ 3 \ 4 \ 4 \ 5), \\ jc &= (1 \ 4 \ 1 \ 2 \ 4 \ 1 \ 3 \ 4 \ 5 \ 3 \ 4 \ 5). \end{aligned}$$

Γενικά, η COO είναι μία ευέλικτη δομή που χρησιμοποιείται ευρέως όμως έχει δύο μειονεκτήματα. Πρώτον, χρειάζεται κάποιες πράξεις μεταξύ ακεραίων για να εξαγάγει τους

δείκτες των στηλών και των γραμμών κάθε στοιχείου όταν χρειάζεται. Δεύτερον, για τεράστιους πίνακες ίσως οδηγηθούμε σε πρόβλημα (overflow) λόγω των ακεραίων που μπορεί να είναι της τάξεως N^2 .

3.1.2 CSR (Compressed Sparse Row)

Η πιο διαδεδομένη και εύχρηστη μορφή αποθήκευσης είναι η CSR η οποία δεν βασίζεται σε κάποια ιδιότητα του πίνακα για αυτό και είναι εφικτή η αποθήκευση οποιουδήποτε αραιού πίνακα. Είναι μία παραλλαγή της μεθόδου αποθήκευσης COO. Σε αυτή την μορφή υπάρχουν τρία διανύσματα. Το πρώτο διάνυσμα AA αποθηκεύει τις τιμές των a_{ij} όπως υπάρχουν ανά γραμμές από το ένα έως το N και το μήκος του AA είναι Nz . Το δεύτερο διάνυσμα ja (ακέραιο διάνυσμα) περιέχει την πληροφορία για τις στήλες των αντίστοιχων μη μηδενικών στοιχείων του AA , το μήκος του είναι Nz . Τέλος, υπάρχει το ακέραιο διάνυσμα ia που περιέχει τους δείκτες στην αρχή κάθε γραμμής των διανυσμάτων AA και ja με μήκος $N + 1$ όπου $ia[N + 1]$ περιέχει το άθροισμα των $ia[1] + Nz$ [30].

Παρακάτω εμφανίζεται ένα απλό παράδειγμα. Έστω πίνακας:

$$A = \begin{pmatrix} 1. & 0. & 0. & 2. & 0. \\ 3. & 8. & 0. & 5. & 0. \\ 2. & 0. & 7. & 8. & 9. \\ 0. & 0. & 10. & 15. & 0. \\ 0. & 0. & 0. & 0. & 19. \end{pmatrix}$$

τότε τα διανύσματα AA , ja και ia θα είναι αντίστοιχα:

$$\begin{aligned} AA &= (1. \ 2. \ 3. \ 8. \ 5. \ 2. \ 7. \ 8. \ 9. \ 10. \ 15. \ 19.), \\ ja &= (1 \ 4 \ 1 \ 2 \ 4 \ 1 \ 3 \ 4 \ 5 \ 3 \ 4 \ 5), \\ ia &= (1 \ 3 \ 6 \ 10 \ 12 \ 13). \end{aligned}$$

3.1.3 CSC (Compressed Sparse Column)

Υπάρχουν αρκετές παραλλαγές της CSR μορφής όπως η μέθοδος CSC (compressed sparse column) που αποτελεί μία δομή αποθήκευσης αραιών πινάκων που ανήκει στην οικογένεια των προηγούμενων δύο. Η βασική διαφοροποίηση από την CSR δομή είναι ότι αποθηκεύονται στήλες αντί για τις γραμμές του πίνακα. Δηλαδή, η CSC μορφή είναι πανομοιότυπη με την CSR για τον πίνακα A^T [30].

Παρακάτω εμφανίζεται στο ίδιο παράδειγμα η τροποποίηση αυτής της δομής. Έστω

$$A = \begin{pmatrix} 1. & 0. & 0. & 2. & 0. \\ 3. & 4. & 0. & 5. & 0. \\ 6. & 0. & 7. & 8. & 9. \\ 0. & 0. & 10. & 11. & 0. \\ 0. & 0. & 0. & 0. & 12. \end{pmatrix}.$$

Τα διανύσματα AA , ja , και ia θα είναι αντίστοιχα:

$$\begin{aligned} AA &= (1. \ 3. \ 6. \ 4. \ 7. \ 10. \ 2. \ 5. \ 8. \ 11. \ 9. \ 12.), \\ ja &= (1 \ 2 \ 3 \ 2 \ 3 \ 4 \ 1 \ 2 \ 3 \ 4 \ 3 \ 5), \\ ia &= (1 \ 4 \ 5 \ 7 \ 11 \ 13). \end{aligned}$$

3.1.4 MSR (Modified Sparse Row)

Η τεχνική της δομής αποθήκευσης MSR (Modified Sparse Row) αποτελεί μια παραλλαγή της CSR μορφής. Χρησιμοποιείται σε πίνακες που προκύπτουν από προβλήματα Μερικών Διαφορικών εξισώσεων και εκμεταλλεύεται το γεγονός ότι στους πίνακες συντελεστών των αγνώστων που προκύπτουν από τις διακριτοποιήσεις των εξισώσεων, τα διαγώνια στοιχεία τους είναι συνήθως μη μηδενικά. Σε αυτήν την δομή υπάρχουν μόνο δύο διανύσματα, το AA που αποθηκεύει τα στοιχεία του πίνακα και το διάνυσμα ja που είναι ένα διάνυσμα που περιέχει τους δείκτες.

Η MSR μορφή είναι ακόμα πιο διαδεδομένη σε σχέση με την CSR μορφή, διότι αποθηκεύει τα διαγώνια στοιχεία του πίνακα A ξεχωριστά. Αυτό είναι το μεγάλο της πλεονέκτημα γιατί αρκετοί πίνακες έχουν γεμάτη με στοιχεία τη διαγώνιο και έτσι τα κρατάμε

σε ένα διάνυσμα μήκους N . Αυτή η μορφή είναι επίσης χρήσιμη σε τριγωνικούς πίνακες με μη μηδενικά στοιχεία στην διαγώνιο.

Στις πρώτες N θέσεις του AA θα αποθηκευτούν τα διαγώνια στοιχεία του πίνακα, ενώ η $N + 1$ θέση μένει κενή (για κάποια μελλοντική πληροφορία του πίνακα). Έπειτα ξεκινώντας από την θέση $N + 2$ αποθηκεύονται όλα τα μη μηδενικά στοιχεία του πίνακα (ανά γραμμή) εκτός της διαγώνιου. Για κάθε μη διαγώνιο στοιχείο στην θέση $AA(k)$ αποθηκεύεται στο $ja(k)$ πληροφορία για την στήλη του στοιχείου. Στις $N + 1$ θέσεις του διανύσματος ja έχουν αποθηκευτεί δείκτες που δείχνουν σε ποια θέση του διανύσματος AA υπάρχει στοιχείο που αλλάζει γραμμή στην δομή του πίνακα. Τελικά επιλέγουμε την CSR διότι αποθηκεύονται δεδομένα σε τρία διανύσματα και είναι ευκολότερη στην χρήση της σε αλγόριθμο. Παρακάτω υπάρχει ένα παράδειγμα για την αποθήκευση τύπου MSR. Για

$$A = \begin{pmatrix} 1. & 0. & 0. & 2. & 0. \\ 3. & 8. & 0. & 5. & 0. \\ 2. & 0. & 7. & 8. & 9. \\ 0. & 0. & 10. & 15. & 0. \\ 0. & 0. & 0. & 0. & 19. \end{pmatrix},$$

τα διανύσματα AA και ja θα είναι αντίστοιχα:

$$\begin{aligned} AA &= (1. \ 8. \ 7. \ 15. \ 19. \ * \ 2. \ 3. \ 5. \ 2. \ 8. \ 9. \ 10), \\ ja &= (7 \ 8 \ 10 \ 13 \ 14 \ 14 \ 4 \ 1 \ 4 \ 1 \ 4 \ 5 \ 3). \end{aligned}$$

3.2 SPARSKIT

Στην παράγραφο 2.3 παρουσιάστηκε ο τρόπος με τον οποίο έγινε η εξαγωγή των γραμμικών συστημάτων με τους αντίστοιχους πίνακες που μελετώνται στη συγκεκριμένη διατριβή. Ο πίνακας που προκύπτει είναι αραιός και έχει μορφή *strucural symmetric*. Λόγω της αναγκαιότητας η αποθήκευση του πίνακα να γίνει σε αραιή μορφή στον κώδικα χρησιμοποιήθηκε η βιβλιοθήκη γραμμικής άλγεβρας για αραιούς πίνακες SPARSKIT. Το λογισμικό αυτό υποστηρίζει όλες τις μορφές αποθήκευσης που αναφέρθηκαν παραπάνω. Ταυτόχρονα, είναι σχετικά εύκολο στη χρήση του και ευρέως διαδεδομένο. Σε αυτό

το πακέτο περιέχονται βασικές πράξεις γραμμικής άλγεβρας, καθώς και επαναληπτικές μέθοδοι επίλυσης γραμμικών συστημάτων: General Minimal Residual (GMRES), Flexible General Minimal Residual (FGMR), Transpose Free quasi minimal residual (TFQMR), Biconjugate Gradient Stabilized (BiCGSTAB), Biconjugate Gradient (BCG), Conjugate Gradient (CG), Conjugate Gradient Method (CGNR), Full Orthogonalization (FOM), BCG με μερική οδήγηση (partial pivoting) (DBCG), Direct versions of Quasi Minimum Residual Method (DQGMRES). Ως προορυθμιστής του πίνακα μπορεί να επιλεγεί μεταξύ των Incomplete LU factorization with dual threshold strategy (ILUT), ILUT με οδήγηση ανά στήλη (ILUTP), ILU with single dropping +diagonal compensation (ILUD), ILUD με οδήγηση ανά στήλη (ILUDP), level-k-ILU (ILUK), απλή ILU(0) preconditioner (ILU0) και τέλος δίνονται οι επιλογές για reorderings, ανάμεσα σε Cuthill McKee (CMK), Reverse Cuthill McKee (RCM), MULTI-COLOR.

Το SPARSKIT είναι γραμμένο από τον Yousef Saad (Version 2) το 1994 σε Fortran 90. Οι παραπάνω μεθοδολογίες αναλύονται και περιγράφονται συνοπτικά στο κεφάλαιο 3. Παρουσιάζονται παρακάτω οι υποκατάλογοι που περιέχονται σε αυτή τη βιβλιοθήκη λογισμικού:

BLASSM: Περιέχει κάποιες από τις βασικές πράξεις γραμμικής άλγεβρας για αραιούς πίνακες καθώς και τα υποπρογράμματα `blasm.f` και `matvec.f`.

DOC: Περιέχει την βασική τεκμηρίωση του πακέτου.

INFO: Περιέχει τις απαραίτητες πληροφορίες για τις υπορουτίνες.

FORMATS: Περιέχει τις υπορουτίνες για την μετατροπή στις διάφορες μορφές αποθήκευσης αραιών πινάκων στο αρχείο `formats.f` και οι υπορουτίνες για τις πράξεις που βρίσκονται στο `unary.f`.

INOUT: Περιέχει τις υπορουτίνες εισόδου και εξόδου.

ITSOL: Περιέχει το πακέτο των επαναληπτικών μεθόδων και των preconditioners.

MATGEN: Περιέχει υπορουτίνες για την δημιουργία πινάκων στους εξής υποκαταλόγους.

- **FDIF**: Περιέχει υπορουτίνες για την κατασκευή πινάκων που προκύπτουν από τη διακριτοποίηση Μερικών Διαφορικών Εξισώσεων με χρήση της μεθόδου Πεπερασμένων Διαφορών.
- **FEM**: Περιέχει υπορουτίνες για την κατασκευή πινάκων που προκύπτουν από τη διακριτοποίηση Μερικών Διαφορικών Εξισώσεων με χρήση της μεθόδου Πεπερασμένων Στοιχείων.
- **MISC**: Περιέχει υπορουτίνες για την κατασκευή Markov Chain πινάκων που πηγάζουν από προβλήματα τυχαίων περιπάτων.

ORDERINGS: Περιέχει τις υπορουτίνες reorderings.

UNSUPP: Περιέχει διάφορες υπορουτίνες και κάποια προγράμματα οδήγησης (drivers) τα οποία ταξινομούνται στους εξής υποκαταλόγους.

- **BLAS1.f**: Περιέχει κάποιες από τις υπορουτίνες της βιβλιοθήκης BLAS
- **MATEXP**: Περιέχει υπορουτίνες σχετικές με εκθετικά πινάκων.
- **PLOTS**: Περιέχει υπορουτίνες σχετικές με τη γραφική απεικόνιση των πλεγμάτων και των προβλημάτων.

Κεφάλαιο 4

Επαναληπτικές Μέθοδοι Επίλυσης

4.1 Εισαγωγή

Έστω το γραμικό σύστημα

$$Ax = b, \quad (3.1)$$

με $A \in \mathbb{R}^{n \times n}$ και $b \in \mathbb{R}^n$ και τη λύση του x να ανήκει στον $\mathbb{R}^n$. Για το παραπάνω πρόβλημα υπάρχουν δύο βασικοί τρόποι αριθμητικής επίλυσης. Η χρήση κάποιας άμεσης μεθόδου (direct method) όπως η απαλοιφή Gauss και με τη χρήση μιας επαναληπτικής μεθόδου (iterative method), όπως είναι η Generalized Minimal Residual (GMRES), Biconjugate Gradient Stabilized (BICGSTAB), Quasi Minimum Residual Method (QMR), Transpose Free QMR (TFQMR), Jacobi, Gauss-Seidel, Successive Overrelaxation (SOR). Τόσο ο πρώτος, όσο και ο δεύτερος τρόπος επίλυσης διαθέτουν πλεονεκτήματα και μειονεκτήματα αντίστοιχα.

Οι άμεσες μέθοδοι χρησιμοποιούνται σε επιμέρους διαδικασίες επίλυσης προβλημάτων CFD (Computational fluid dynamics) και στην περιοχή της structural analysis [7], ενώ, οι επαναληπτικές μέθοδοι έγιναν ιδιαίτερα γνωστοί τον τελευταίο αιώνα.

Στις άμεσες μεθόδους βασιζόμαστε στην παραγοντοποίηση του πίνακα των συντελεστών A , τα αποτελέσματα της επίλυσης είναι αξιόπιστα και χρησιμοποιούνται ευρέως. Όμως, για μεγάλα γραμμικά συστήματα οι επαναληπτικές μέθοδοι απαιτούν λιγότερο χώρο

αποθήκευσης και λιγότερες πράξεις, αφού ο χρήστης είναι σε θέση να καθορίσει τη προσέγγιση της αριθμητικής στη πραγματική λύση του γραμμικού συστήματος. Στις επαναληπτικές μεθόδους δεν τροποποιείται ο πίνακας A , όπως στην απαλοιφή Gauss, αλλά ξεκινώντας από μία αρχική προσέγγιση x^0 της λύσης x (συνήθως αυθαίρετη) κατασκευάζεται μία ακολουθία διαδοχικών προσεγγίσεων $x^m, m = 1, 2, \dots, n$ η οποία συγκλίνει στην x όταν $m \rightarrow \infty$. Σε ορισμένες περιπτώσεις οι επαναληπτικές μέθοδοι αποτυγχάνουν στην επίλυση του γραμμικού συστήματος και η χρήση προρρύθμισης είναι αναγκαία. Οι επαναληπτικές μέθοδοι μπορούν να εκφραστούν με την απλή μορφή $x^{(k)} = Bx^{(k+1)} + c$ όπου όταν ο πίνακας B και το διάνυσμα c δεν εξαρτώνται από τον αριθμό των επαναλήψεων k , ονομάζονται στατικές (stationary) μέθοδοι, διαφορετικά ονομάζονται μη στατικές (nonstationary) μέθοδοι. Οι βασικές παράμετροι τους υπολογίζονται από τα εσωτερικά γινόμενα των υπολοίπων (residual) $r = b - Ax$ ή από άλλα επιμέρους διανύσματα της επαναληπτικής διαδικασίας. Στις στατικές μεθόδους ανήκουν οι Gauss-Seidel, Jacobi, Successive Overrelaxation (SOR), Symmetric Successive Overrelaxation (SSOR), ενώ στην άλλη κατηγορία ανήκουν οι μέθοδοι των υπόχωρων Krylov όπως θα δούμε αναλυτικότερα στο κεφάλαιο 3.2.

4.1.1 Μέθοδοι Jacobi και Gauss-Seidel

Οι δύο πιο γνωστές επαναληπτικές μέθοδοι είναι οι μέθοδοι του Jacobi και του Gauss-Seidel, οι οποίες προϋποθέτουν ότι τα διαγώνια στοιχεία a_{ii} του A είναι μη μηδενικά. Γράφουμε το σύστημα $Ax = b$ στη μορφή

$$x_i = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j), \quad i = 1, \dots, n. \quad (4.1)$$

Για κάποια δεδομένη αρχική προσέγγιση της λύσης $x^{(0)} \in \mathbb{R}^n$. Κάθε βήμα $x^{(m)} \rightarrow x^{(m+1)}$ της μεθόδου του Jacobi είναι

$$x_i^{(m+1)} := \left\{ \frac{1}{a_{ii}}b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(m)} - \sum_{j=i+1}^n a_{ij}x_j^{(m)} \right\}, \quad i = 1, \dots, n. \quad (4.2)$$

Αν για τον υπολογισμό των $x_i^{(m+1)}$ ($i > 1$) αντί των $x_1^{(m)}, \dots, x_{i-1}^{(m)}$ χρησιμοποιήσουμε τις υπολογισμένες τιμές $x_1^{(m+1)}, \dots, x_{i-1}^{(m+1)}$, τότε προκύπτει η μέθοδος Gauss-Seidel:

$$x_i^{(m+1)} := \left\{ \frac{1}{a_{ii}} b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)} \right\}, \quad i = 1, \dots, n. \quad (4.3)$$

Η μέθοδος Gauss-Seidel είναι "καλύτερη" της μεθόδου του Jacobi, αφού είναι λογικό να αναμένει κανείς ότι τα $x_j^{m+1}, j = 1, \dots, i-1$, αποτελούν καλύτερες προσεγγίσεις των x_j απ' ό,τι τα $x_j^{(m)}$. Έτσι μπορούν να υπάρξουν κατηγορίες συστημάτων για τα οποία η μέθοδος των Gauss-Seidel συγκλίνει ταχύτερα από τη μέθοδο του Jacobi. Γενικά όμως υπάρχουν συστήματα για τα οποία η μία από τις αυτές τις μεθόδους να συγκλίνει ενώ η άλλη μπορεί να μην συγκλίνει.

Άλλη μία ενδιαφέρουσα κατηγορία πινάκων, από την άποψη των πρακτικών εφαρμογών, είναι οι συμμετρικοί, θετικά ορισμένοι πίνακες $A \in \mathbb{R}^{n,n}$. Για αυτούς, μπορεί να αποδειχθεί ότι συγκλίνει η Gauss-Seidel, αλλά ότι η Jacobi ίσως να μην συγκλίνει.

Θέτουμε $D := \text{diag}(a_{11}, \dots, a_{nn})$.

$$L := \begin{pmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{pmatrix}, U := \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ & 0 & a_{23} & \dots & a_{2n} \\ & & 0 & & \vdots \\ & & & \ddots & a_{n-1,n} \\ & 0 & & & 0 \end{pmatrix}$$

Έχουμε ότι $A = L + D + U$ και οι 4.2 και οι 4.3 γράγονται στη μορφή

$$Dx^{(m+1)} = -(L + U)x^{(m)} + b,$$

$$(L + D)x^{(m+1)} = -Ux^{(m)} + b.$$

Παρατηρούμε λοιπόν ότι οι δύο μέθοδοι είναι ειδικές περιπτώσεις της γενικής μεθόδου

$$Mx^{(m+1)} = Nx^{(m)} + b, \quad A = M - N.$$

Η μέθοδος Gauss-Seidel είναι παραπλήσια με την Jacobi με την διαφορά ότι αναβαθμίζει τις τιμές όταν χρειάζεται. Η μέθοδος Jacobi είναι αρκετά απλή στην χρήση της εκτός της περίπτωσης που ο πίνακας είναι αυστηρά κυρίαρχα διαγώνιος και γενικά είναι μέθοδος η οποία διαθέτει αρκετές ιδιότητες παραλληλισμού. Σύμφωνα με το βιβλίο [36] η Jacobi έχει αυξημένη την ιδιότητα της παραλληλοποίησης, ενώ η Gauss-Seidel όχι. Οι λεγόμενες 'παράλληλες' μέθοδοι αποκτούν όλο και μεγαλύτερη σημασία από τη ραγδαία ανάπτυξη των παράλληλων επεξεργαστών.

4.1.2 Successive Overrelaxation (SOR) και Symmetric Successive Overrelaxation (SSOR)

Η SOR (μέθοδος διαδοχικής υπερχαλάρωσης) μπορεί να συγκλίνει γρηγορότερα από την Gauss-Seidel κατά μία τάξη μεγέθους. Έχει παρατηρηθεί ότι κατά την επίλυση μεγάλων αραιών συστημάτων που προκύπτουν από τη διακριτοποίηση ελλειπτικών μερικών Διαφορικών Εξισώσεων με μεθόδους πεπερασμένων διαφορών ισχύει ότι:

Έστω η $x^{(m)}$ μία γνωστή προσέγγιση της x και έστω $\bar{x}^{(m+1)}$ η προσέγγιση Gauss-Seidel. Ο γραμμικός συνδυασμός τους θα είναι:

$$x^{(m+1)} = w\bar{x}^{(m+1)} + (1 - w)x^{(m)},$$

που έχει μικρότερο σφάλμα απ' ότι η $\bar{x}^{(m+1)}$, υπό ορισμένες προϋποθέσεις και τιμές της παραμέτρου $w \neq 0$ [36].

Για το σύστημα $Ax = b$ με $a_{ii} \neq 0$ το βήμα $m \rightarrow m + 1$ της μεθόδου SOR θα έχει τη μορφή:

$$x_i^{(m+1)} = \frac{w}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i}^{i-1} a_{ij} x_j^{(m)} \right] + (1 - w)x_i^m,$$

$1 \leq i \leq n$ που σε μορφή πινάκων γράφεται:

$$(D + wL)x^{(m+1)} = [(1 - w)D - U]x^{(m)} + wb.$$

- Φυσικά, ο πίνακας $(D + wL)$ είναι αντιστρέψιμος αν $a_{ii} \neq 0 \forall i$.
- Για $w = 1$ η SOR συμπίπτει με την μέθοδο Gauss-Seidel.
- Για πίνακα συμμετρικό και θετικά ορισμένο και $0 < w < 2$ τότε γνωρίζουμε ότι η SOR συγκλίνει.

Το παραπάνω επαναληπτικό σχήμα ονομάζεται σχήμα χαλάρωσης και το w λέγεται παράμετρος χαλάρωσης. Αν το $w < 1$, το σχήμα λέγεται successive underrelaxation (διαδοχική υποχαλάρωση), ενώ για $w > 1$ το σχήμα λέγεται successive overrelaxation. Το βέλτιστο w εξαρτάται από τον πίνακα και είναι δύσκολο να υπολογισθεί. Η μέθοδος SSOR, χρησιμοποιείται σαν προρρυθμιστής (preconditioner) για μη στατικά (nonstationary) προβλήματα.

4.2 Krylov Subspace Methods

Στον υπόχωρο Krylov ανήκουν όλοι οι αλγόριθμοι στους οποίους διατηρούμε όλες τις προσεγγίσεις x_i και τις συνδυάζουμε ώστε να προσεγγίσουν καλύτερη λύση. Ο υπόχωρος Krylov είναι ο χώρος που παράγεται $K_n(A, r_0) = \text{span}(r_0, Ar_0, A^2r_0, \dots, A^{n-1}r_0)$ με $r_0 = b - Ax$. Πήρε το όνομα του από τον Ρώσο μαθηματικό Alexei Krylov που τα εφάρμοσε, ο οποίος δημοσίευσε για πρώτη φορά εργασία με αυτό το θέμα το 1931.

Οι πιο γνωστές Krylov Subspace μέθοδοι είναι οι Lanczos, Arnoldi, Conjugate Gradient (CG), Generalized Minimal Residual Method (GMRES), Flexible version of Generalized Minimal Residual Method (FGMRES), Biconjugate Gradient Stabilized (BICGSTAB), Quasi Minimum Residual Method (QMR), Transpose Free QMR (TFQMR) [30]. Παρακάτω περιγράφονται συνοπτικά οι αλγόριθμοι των μεθόδων που υλοποιήθηκαν και μελετήθηκαν στην συγκεκριμένη εργασία.

4.2.1 GMRES

Η Generalized Minimal Residual (GMRES) είναι μία επαναληπτική μέθοδος που εξελίχθηκε το 1986. Αρχικά ήταν ισοδύναμη με τις υπόλοιπες τεχνικές, όμως αργότερα έγινε περισσότερο προσιτή λόγω της καλής αριθμητικής συμπεριφοράς της και του χαμηλού κόστους της σε θέμα μνήμης. Είναι από τις πιο κατάλληλες για την εύρεση της αριθμητικής λύσης ενός μη συμμετρικού γραμμικού συστήματος. Η μέθοδος προσεγγίζει τη λύση με ένα διάνυσμα από τον υπόχωρο Krylov ελαχιστοποιώντας το υπόλοιπο, ($r_0 = b - Ax$ με A τον πίνακα συντελεστών και b το δεξί μέρος του γραμμικού συστήματος).

Είναι κατάλληλη για επίλυση γραμμικών συστημάτων με μη συμμετρικούς πίνακες και για ένα καθορισμένο αριθμό επαναληπτικών βημάτων η GMRES τείνει να μειώσει το υπόλοιπο, κατασκευάζοντας μία ορθογώνια βάση για τον χώρο των λύσεων. Δυστυχώς, αυτό μπορεί να οδηγήσει σε μεγαλύτερη απαίτηση χώρου μνήμης όπου αυξάνει γραμμικά σε σχέση με τον αριθμό των επαναληπτικών βημάτων.

Το πρόβλημα αυτό αντιμετωπίστηκε αργότερα, κάνοντας επανεκκίνηση της επαναληπτικής διαδικασίας κάθε m βήματα με την GMRES(m). Ένα πρακτικό μειονέκτημα της μεθόδου GMRES είναι ότι πρέπει να αποθηκεύονται όλα τα διαδοχικά υπόλοιπα σε διανύσματα και έτσι η λύση του συστήματος γίνεται όλο και πιο περίπλοκη καθώς η σύγκλιση δεν είναι εγγυημένη [31]. Ο προφανής τρόπος αποφυγής του παραπάνω προβλήματος είναι η βελτιωμένη έκδοση της GMRES(m). Η επιλογή του m δεν είναι τυχαία αφού είναι ένας ακέραιος που αλλάζει ανάλογα με το σύστημα που λύνουμε σε κάθε περίπτωση. Η επιλογή μικρού m μπορεί να οδηγήσει, είτε σε χαμηλής τάξεως σύγκλιση, είτε ακόμα και στη μη σύγκλιση της μεθόδου. Για να περιορίσουμε την αύξηση του απαιτούμενου χώρου αποθήκευσης η επανεκκίνηση είναι απαραίτητη. Πότε θα γίνεται αυτό απαιτεί καλή γνώση, επιδεξιότητα και εμπειρία και εξαρτάται από τον πίνακα A και το δεξί μέλος δηλαδή το b .

Αλγόριθμος GMRES

Υπολόγισε $r_0 = b - Ax_o$ $\beta := \|r_0\|_2$, and $v_1 := r_0/\beta$

Για $j = 1, 2, \dots, maxit$

Υπολόγισε $w_j := Av_j$

Για $i = 1, \dots, j$

$$h_{ij} := (w_j, v_i)$$

$$w_j := w_j - h_{ij}v_i$$

Τέλος Για

$$h_{j+1,j} = \|w_j\|_2.$$

Εάν $h_{j+1,j} = 0$, βάλε $m := j$ και *break*

$$v_{j+1} = w_j/h_{j+1,j}$$

Τέλος Για

Ορίζουμε $(m+1) \times m$ Hessenberg matrix $\bar{H}_m = \{h_{ij}\}_{1 \leq i \leq m+1, 1 \leq j \leq m}$

Υπολογίζουμε το y_m που ελαχιστοποιεί $\|\beta e_1 - \bar{H}_m y\|$ και $x_m = x_o + V_m y_m$

$$r = b - Ax_m$$

Εάν $\|r\|$ αρκετά μικρό τότε, *break*

Αλγόριθμος GMRES(m)

Υπολόγισε $r_0 = b - Ax_o$, $\beta := \|r_0\|_2$, and $v_1 := r_0/\beta$

Για $l = 1, 2, \dots, maxout$

Για $j = 1, 2, \dots, m$

Υπολόγισε $w_j := Av_j$

Για $i = 1, \dots, j$

$$h_{ij} := (w_j, v_i)$$

$$w_j := w_j - h_{ij}v_i$$

Τέλος Για

$$h_{j+1,j} = \|w_j\|_2.$$

$$v_{j+1} = w_j/h_{j+1,j}$$

Τέλος Για

Ορίζουμε $(m + 1) \times m$ Hessenberg matrix $\bar{H} = \{h_{ij}\}_{1 \leq i \leq m+1, 1 \leq j \leq m}$

Υπολογίζουμε το y_m που ελαχιστοποιεί $\|\beta e_1 - \bar{H}_m y\|_2$, και $x_0 = x_m + V_m y_m$

$$r = b - Ax_m$$

Εάν $\|r\|$ αρκετά μικρό τότε, *break*

Αλλιώς $x_o = x_m$ και $v_1 = r/\beta$

Τέλος Για

Σύμφωνα με τον παραπάνω αλγόριθμο ορίζονται δύο είδη επαναλήψεων για την μέθοδο GMRES(m). Κάθε επανάληψη μεταξύ δύο επανεκκινήσεων του αλγορίθμου λέγεται εσωτερική επανάληψη. Οι m εσωτερικές επαναλήψεις δημιουργούν τη βάση του υπόχωρου Krylon. Ως εξωτερική επανάληψη ορίζεται το σύνολο των m εσωτερικών επαναλήψεων μεταξύ δύο επανεκκινήσεων του αλγορίθμου.

4.2.2 BICGSTAB

Η μέθοδος Biconjugate Gradient Stabilized Method, γνωστή και ως BiCGSTAB, είναι μία επαναληπτική μέθοδος που αναπτύχθηκε από τον H.A. van der Vorst για την επίλυση μη συμμετρικών γραμμικών συστημάτων. Διαφέρει από την biconjugate gradient method (BiCG) και έχει ταχύτερη και ομαλέστερη σύγκλιση σε σχέση με την αρχική έκδοση της BiCG.

Η παραπάνω μέθοδος ανήκει στις μεθόδους των υπόχωρων Krylon και για την λύση του συστήματος $Ax = b$ ξεκινάει από μία αρχική προσέγγιση x^0 . Αξίζει να σημειωθεί ότι η BiCGSTAB έχει δύο κριτήρια τερματισμού. Έτσι ελέγχεται και η μισή επανάληψη, ώστε να μην γίνονται άσκοπα πράξεις. Το ίδιο συμβαίνει και στην TFQMR που αναλύεται λίγο παρακάτω. Επίσης, απαιτεί δύο πολλαπλασιασμούς μεταξύ πίνακα και διανύσματος και τέσσερα εσωτερικά γινόμενα. Χρησιμοποιεί δύο εσωτερικά γινόμενα περισσότερα από την BiCG και την CGS (Conjugate Gradient Square). Η μέθοδος αυτή μπορεί να ερμηνευτεί ως το προϊόν της BiCG και της επαναλαμβανόμενης εφαρμογής της GMRES.

Αλγόριθμος BICGSTAB

Υπολόγισε $r_o = b - Ax_o$, r_o^* τυχαίο

$p_o := r_o$

Για $j = 0, 1, \dots, maxit$

$$a_j := (r_j, r_o^*) / (Ap_j, r_o^*)$$

$$s_j := r_j - a_j Ap_j$$

$$w_j := (As_j, s_j) / (As_j, As_j)$$

$$x_{j+1} := x_j + a_j p_j + w_j s_j$$

Εάν $\|b - Ax_{j+1}\|$ αρκετά μικρό τότε *break*, αλλιώς,

$$r_{j+1} = s_j - w_j As_j$$

$$\beta_j := \frac{(r_{j+1}, r_o^*)}{(r_j, r_o^*)} \times \frac{a_j}{w_j}$$

$$p_{j+1} := r_{j+1} + \beta_j (p_j - w_j Ap_j)$$

Τέλος Για

4.2.3 FGMRES

Στο σημείο αυτό θα αναφερθεί περιληπτικά η προορυθμισμένη έκδοση των μεθόδων του υπόχωρου Krylov. Υπάρχουν η προορυθμισμένη CG, η προορυθμισμένη GMRES και διάφορες άλλες. Υπάρχουν όμως διάφορες επαναληπτικές διαδικασίες που επιδέχονται πολλές διαφοροποιήσεις, όπως για παράδειγμα να αλλάζει ο preconditioner από βήμα σε βήμα. Τέτοιες διαδικασίες ονομάζονται ευέλικτες διαδικασίες (flexible iterations). Μία ευέλικτη επαναληπτική μέθοδος είναι η FGMRES (Flexible- GMRES), στην οποία γίνεται χρήση δεξιού προορυθμιστή.

Αλγόριθμος FGMRES

1. Υπολόγισε $r_o = b - Ax_o$, $\beta = \|r_o\|_2$, και $v_1 = r_o/\beta$ (1)

Για $j = 1, \dots, m$

$$\text{Υπολόγισε } z_j := M_j^{-1} v_j$$

Υπολόγισε $w := Az_j$

Για $i = 1, \dots, j$

$$h_{i,j} := (w, v_i)$$

$$w := w - h_{i,j}v_i$$

Τέλος Για

Υπολόγισε $h_{j+1,j} = \|w\|_2$ και $v_{j+1} = w/h_{j+1,j}$

Ορίζουμε $Z_m := [z_1, \dots, z_m]$, $\bar{H}_m = \{h_{ij}\}_{1 \leq i \leq m+1: 1 \leq j \leq m}$

Τέλος Για

Υπολογίζουμε το y_m που ελαχιστοποιεί $\|\beta e_1 - \bar{H}_m y\|_2$ και $x_m = x_o + Z_m y_m$

Εάν $\|b - Ax_m\|$ αρκετά μικρό τότε, *break*

Αλλιώς, βάλε $x_o \leftarrow x_m$ και πήγαινε στο 1.

4.2.4 TFQMR

Η μέθοδος Transpose -free QMR (TFQMR) προτάθηκε από τον Freund [12] και προέρχεται από τον αλγόριθμο της CGs Όλοι οι αλγόριθμοι αυτής της οικογένειας προσπαθούν να ελαχιστοποιήσουν την νόρμα της ποσότητας $r_o = b - Ax_o$. Το quasi-residual σχετίζεται με το πραγματικό υπόλοιπο.

Αλγόριθμος *TFQMR*

Υπολόγισε $w_o = u_o = r_o = b - Ax_o$, $v_o = Au_o$, $d_o = 0$

$$t_o = \|r_o\|_2, \vartheta = h_o = 0$$

Διάλεξε r_o^* ώστε $\rho_o \equiv (r_o^*, r_o) \neq 0$

Για $m = 0, 1, 2, \dots, \maxit$

Εάν m άρτιος τότε

$$a_{m+1} = a_m = \rho_m / (v_m, r_o^*)$$

$$u_{m+1} = u_m - a_m v_m$$

Τέλος Για

$$w_{m+1} = w_m - a_m A u_m$$

$$d_{m+1} = u_m + (\vartheta_m^2 / a_m) h_m d_m$$

$$\vartheta_{m+1} = \|w_{m+1}\|_2 / t_m, \quad c_{m+1} = (1 + \vartheta_{m+1}^2)^{-1/2}$$

$$t_{m+1} = t_m \vartheta_{m+1} c_{m+1}, \quad h_{m+1} = c_{m+1}^2 a_m$$

$$x_{m+1} = x_m + h_{m+1} d_{m+1}$$

Εάν $\|b - Ax_m\|$ αρκετά μικρό τότε *break*, αλλιώς

Εάν m περιπτώς τότε

$$\rho_{m+1} = (w_{m+1}, r_o^*), \quad \beta_{m-1} = \rho_{m+1} / \rho_{m-1}$$

$$u_{m+1} = w_{m+1} + \beta_{m-1} u_m$$

$$v_{m+1} = A u_{m+1} + \beta_{m-1} (A u_m + \beta_{m-1} v_{m-1})$$

Τέλος Εάν

Τέλος Για

4.2.5 CG

Η CG μέθοδος είναι κατάλληλη για συμμετρικούς θετικά ορισμένους πίνακες SPD. Σε αυτή την περίπτωση η ταχύτητα σύγκλισης εξαρτάται από το δείκτη κατάστασης του πίνακα. Εάν οι ιδιοτιμές είναι καλά διασκορπισμένες τότε το αποτέλεσμα της συμπεριφοράς της σύγκλισης μπορεί να είναι γραμμικό (superlinear).

Αλγόριθμος CG

Υπολόγισε $r_0 = b - Ax_o$, $p_o := r_o$

Για $j = 0, 1, \dots, maxit$

$$a_j := (r_j, r_j) / (Ap_j, p_j)$$

$$x_{j+1} := x_j + a_j p_j$$

Εάν $\|b - Ax_{j+1}\|$ αρκετά μικρό τότε *break*, αλλιώς,

$$r_{j+1} := r_j - a_j A p_j$$

$$\beta_j := (r_{j+1}, r_{j+1})$$

$$p_{j+1} := r_{j+1} + \beta_j p_j$$

Τέλος Για

Η συμπεριφορά της σύγκλισης μίας επαναληπτικής μεθόδου εξαρτάται από πολλούς παραμέτρους, όπως είναι η διανομή των ιδιοτιμών και οι γωνίες που σχηματίζονται μεταξύ των ιδιοδιανυσμάτων. Παράμετροι που για τους περισσότερους που μελετούν και χρησιμοποιούν τις επαναληπτικές μεθόδους είναι άγνωστοι. Από όλα τα παραπάνω μπορούμε να διαπιστώσουμε ότι η επιλογή της κατάλληλης μεθόδου για ένα πρόβλημα είναι δύσκολη.

Οι επαναληπτικές μέθοδοι δε συγκλίνουν πάντοτε, δηλαδή ένα συγκεκριμένο κριτήριο μπορεί να μην ικανοποιείται. Υπάρχουν περιπτώσεις όπως θα δούμε στα κεφάλαια που ακολουθούν που η σύγκλιση είναι ταχύτερη σε σχέση με άλλες, έτσι απαιτείται σαφώς λιγότερος χρόνος σε σχέση με την γνωστή απαλοιφή *Gauss*, όπου ο πίνακας $A = LU$. Ο L πίνακας είναι κάτω τριγωνικός με μονάδες στη διαγώνιο. Κάτω από τη διαγώνιο περιέχει τους πολλαπλασιαστές της απαλοιφής, ενώ ο πίνακας U είναι άνω τριγωνικός και προκύπτει μετά την τριγωνοποίηση του αρχικού πίνακα.

4.3 Προρρυθμιστές (Preconditioners)

Με τον όρο προρρύθμιση (preconditioning) αναφερόμαστε στην μετατροπή του συστήματος $Ax = b$, σε ένα άλλο το οποίο θα έχει την ίδια λύση αλλά με καλύτερες ιδιότητες για την επαναληπτική λύση του. Η σύγκλιση μίας επαναληπτικής μεθόδου εξαρτάται από συγκεκριμένες ιδιότητες των ιδιοτιμών του πίνακα συντελεστών του γραμμικού συστήματος και για να βελτιωθούν αυτές, συχνά μετατρέπουμε το γραμμικό σύστημα σε ένα άλλο καταλληλότερο. Αυτή η διαδικασία είναι γνωστή ως προρρύθμιση (preconditioning). Ο προρρυθμιστής λοιπόν, είναι ένας πίνακας που επιφέρει αυτή τη μετατροπή στο γραμμικό σύστημα.

Με την προρρύθμιση προσπαθούμε να βελτιώσουμε τις φασματικές ιδιότητες του πίνακα συντελεστών. Για ένα συμμετρικό και θετικά ορισμένο πίνακα για παράδειγμα, η τάξη σύγκλισης της *conjugate gradient* μεθόδου εξαρτάται από τον τρόπο με τον οποίο είναι διασκορπισμένες οι ιδιοτιμές του πίνακα A . Φυσικά ο μετασχηματισμένος πίνακας θα έχει μικρότερο δείκτη κατάστασης και οι ιδιοτιμές του θα βρίσκονται γύρω στο 1.

Για μη συμμετρικούς πίνακες, όπως αυτούς που μελετάμε στην παρούσα εργασία η κατάσταση είναι περισσότερο πολύπλοκη και οι ιδιοτιμές μπορεί να μην περιγράφουν την σύγκλιση του μη συμμετρικού πίνακα, όπως συμβαίνει με την GMRES [7]. Παρόλα αυτά, ένα φάσμα απομακρυσμένο από το 0, συχνά έχει ως αποτέλεσμα γρήγορη σύγκλιση, ειδικά στις περιπτώσεις που ο μετασχηματισμένος πίνακας είναι πολύ κοντά στον κανονικό.

Εάν M είναι ένας αντιστρέψιμος πίνακας που προσεγγίζει τον A , τότε το γραμμικό σύστημα $M^{-1}Ax = M^{-1}b$ έχει την ίδια λύση με το σύστημα $Ax = b$ και μπορεί να λυθεί εύκολα. Ο πίνακας M λέγεται πίνακας προρρύθμισης και παραπάνω εφαρμόστηκε από την αριστερή πλευρά (left). Το σύστημα μπορεί να προρρυθμιστεί και από δεξιά (right), $AM^{-1}y = b, x = M^{-1}y$, είτε μπορεί να υπάρξει ο συνδυασμός τους *split preconditioner*, δηλαδή $M_1^{-1}AM_2^{-1}y = M_1^{-1}b, x = M_2^{-1}y$, όπου $M = M_1M_2$. Το είδος προρρύθμισης που θα χρησιμοποιηθεί (αριστερό, δεξί, split) εξαρτάται από την επαναληπτική μέθοδο που θα εφαρμοστεί. Στις επαναληπτικές μεθόδους με ελαχιστοποίηση υπολοίπου όπως είναι η Gmres, χρησιμοποιείται προρρύθμιση από δεξιά συνήθως. Σε πραγματικούς υπολογισμούς τα υπόλοιπα του συστήματος που έχει εφαρμοστεί προρρύθμιση από δεξιά είναι ίδια με το υπόλοιπο $r_k = b - Ax_k$.

Αξιοσημείωτο είναι ότι οι πίνακες $M^{-1}A, AM^{-1}$ και $M_1^{-1}AM_2^{-1}$ είναι όμοιοι για αυτό και έχουν ίδιες ιδιοτιμές. Στην περίπτωση που ο A και ο M είναι συμμετρικοί και θετικά ορισμένοι (*SPD*), η σύγκλιση της *CG* μεθόδου θα είναι ίδια για όλες τις περιπτώσεις. Από την άλλη πλευρά όμως, σε μη κανονικές περιπτώσεις όπως για την GMRES η συμπεριφορά της εξαρτάται από το είδος του προρρυθμιστή. Ένας προρρυθμιστής θεωρείται

καλός, όταν το γραμμικό σύστημα που προκύπτει διαθέτει τα παρακάτω:

- μπορεί να λυθεί εύκολα και γρήγορα,
- επιλύεται αποδοτικά χωρίς την απαίτηση μεγάλου υπολογιστικού χώρου.

Δηλαδή η σύγκλιση θα είναι γρηγορότερη και ταυτόχρονα θα πρέπει να εξασφαλίζει ότι οι επαναλήψεις δεν θα είναι υπολογιστικά πολύ ακριβές υπολογιστικά. Οι παραπάνω απαιτήσεις για την επιλογή ενός σωστού προρρυθμιστή θα πρέπει να βρίσκονται σε ισορροπία μεταξύ τους. Το ιδανικό θα ήταν οι πράξεις να γίνονται γρήγορα χωρίς υπερβολή σε υπολογιστικό χώρο και κόστος. Ο στόχος της προρρύθμισης (preconditioning) λοιπόν είναι, να αυξήσει την σύγκλιση της επαναληπτικής μεθόδου μειώνοντας τον αριθμό των επαναλήψεων που απαιτούνται για την σύγκλιση, χωρίς να αυξηθούν οι υπολογισμοί ανά επανάληψη [7].

Παρακάτω παρουσιάζονται οι τεχνικές προρρύθμισης $ILU0$, $ILUK$, $ILUT$ οι οποίες χρησιμοποιήθηκαν στον αλγόριθμο μας. Στα επόμενα κεφάλαια παρουσιάζονται τα πλεονεκτήματα και τα μειονεκτήματα τους ανάλογα με το πρόβλημα που υπάρχει προς λύση σε κάθε περίπτωση. Παρουσιάζεται ο τρόπος με τον οποίο πρέπει να γίνετε επιλογή τους καθώς και πως επηρεάζονται με τη μεταβολή συγκεκριμένων παραμέτρων.

4.3.1 ILU

Για λόγους επίδοσης και κυρίως όταν ο πίνακας του συστήματος είναι ιδιαίτερα αραιός, χρησιμοποιείται συχνά μία παραλλαγή της παραγοντοποίησης LU . Η ατελής παραγοντοποίηση ILU . Ο πίνακας $M = LU$ εκφράζεται στη μορφή παραγοντοποίησης με τους πίνακες L και U κάτω και άνω τριγωνικούς αντίστοιχα. Ο πίνακας M επιλέγεται ώστε M^{-1} να αποτελεί ικανοποιητική προσέγγιση του A^{-1} . Η ονομασία ατελής παραγοντοποίηση έχει δοθεί διότι κατά την παραγοντοποίηση του πίνακα M έχουν αγνοηθεί τα στοιχεία που πριν την παραγοντοποίηση ήταν μηδενικά και κατά την εκτέλεση της έγιναν μη μηδενικά [15].

Η ατελής παραγοντοποίηση θεωρείται μία κατηγορία προρρυθμισμού, σχετίζεται με το επίπεδο γεμίσματος και δεν επηρεάζεται από τις τιμές του πίνακα αλλά από τη δομή του. Αυτό επιφέρει πολλά προβλήματα υπολογιστικού κόστους για ρεαλιστικά προβλήματα. Μερικές εναλλακτικές μέθοδοι βασίζονται στην απόρριψη στοιχείων κατά την απαλοιφή Gauss με την εισαγωγή μιας σειράς κανόνων απόρριψης μικρών σε αριθμητική τιμή στοιχείων.

4.3.2 *ILU0*

Η no-fill *ILU*, δηλαδή η *ILU(0)* μέθοδος προρρυθμισμού είναι σχετικά εύκολη στην εφαρμογή της. Το υπολογιστικό της κόστος είναι πολύ μικρό και για μικρά γραμμικά συστήματα είναι αποτελεσματική. Για τέτοιου είδους προβλήματα προορίζονται οι προρρυθμιστές. Όμως, για προβλήματα πιο ρεαλιστικά με μεγάλο υπολογιστικό κόστος οι no-fill παραγοντοποιήσεις δίνουν ως αποτέλεσμα ακατέργαστες προσεγγίσεις του πίνακα A και η χρήση προρρυθμιστών με κάποιο γέμισμα (fill in) είναι απαραίτητη. Το αρχικό γέμισμα του πίνακα ορίζεται ως εξής:

$$lev_{ij} = \begin{cases} 0, & \text{if } a_{ij} \neq 0 \text{ or } i = j \\ \infty, & \text{otherwise} \end{cases} \quad (4.4)$$

Όμως, το επίπεδο του γεμίσματος ανανεώνεται με βάση τον τύπο

$$lev_{ij} = \min\{lev_{ij}, lev_{ik} + lev_{kj} + 1\}. \quad (4.5)$$

Ορίζουμε l έναν μη αρνητικό ακέραιο αριθμό. Με την μέθοδο *ILU(l)*, όλα τα γεμίσματα τα οποία θα είναι μεγαλύτερα αριθμητικά από το l θα αγνοούνται. Οπότε, για $l = 0$, θα έχουμε τον preconditioner *ILU(0)*. Σε πολλές περιπτώσεις η *ILU(1)* είναι μία σημαντική βελτίωση της *ILU(0)*. Σπάνια είναι η χρήση μεγάλων τιμών του l , εκτός βέβαια ορισμένων δύσκολων προβλημάτων, διότι αυξάνονται τα υπολογιστικά κόστη καθώς αυξάνουμε την τιμή του l . Η τιμή του l ποικίλει ανάλογα με το πρόβλημα που υπάρχει προς λύση σε κάθε περίπτωση. Για παράδειγμα πίνακες που είναι πολύ μακριά από το να είναι διαγώνια ενισχυμένοι (diagonally dominant) η *ILU(l)*, απαιτεί γεμίσματα τα οποία θα

είναι πολύ μικρά σε απόλυτη τιμή και έτσι θα έχουμε προορύθμιση χαμηλής ποιότητας. Στην πραγματικότητα, γεμίσματα που συνεισφέρουν σημαντικά στην ποιότητα του preconditioner αποθηκεύονται και χρησιμοποιούνται.

Drop tolerance: Η παράμετρος αυτή είναι ένας θετικός αριθμός t που χρησιμοποιείται στα κριτήρια απόρριψης στοιχείων. Μία απόλυτη στρατηγική απόρριψης μπορεί να χρησιμοποιηθεί εάν δεχόμαστε καινούρια γεμίσματα του πίνακα μόνο αν είναι μεγαλύτερα του t . Για πίνακες που είναι κακώς σταθμισμένοι (badly scaled) μπορεί να μην αποδώσει τα αναμενόμενα αποτελέσματα και για αυτές τις περιπτώσεις το (relative drop tolerance) σχετικό όριο απόρριψης είναι προτιμότερο. Όταν γίνεται απαλοιφή της γραμμής i του πίνακα, θα δεχόμαστε ένα νέο γέμισμα του πίνακα μόνο στην περίπτωση που είναι μεγαλύτερο σε απόλυτη τιμή από το $t\|a_{i*}\|_2$ όπου a_{i*} είναι η i γραμμή του πίνακα A .

Ένα μειονέκτημα της παραπάνω εφαρμογής είναι η διαδικασία με την οποία θα γίνετε επιλογή της κατάλληλης τιμής για το drop tolerance, οπότε η στρατηγική είναι να γίνονται δοκιμές μέχρι την εύρεση της καταλληλότερης τιμής. Μία άλλη δυσκολία που πρέπει να αντιμετωπιστεί είναι στην πρόβλεψη της αποθήκευσης των ατελών παραγόντων LU . Σε κάθε βήμα εξάλειψης θα αγνοούνται τα γεμίσματα που είναι μικρότερα από το γινόμενο $(t\|a_i\|_2)$, έτσι θα κρατήσουμε τα στοιχεία p με το μεγαλύτερο μέγεθος ανά γραμμή. Η παραλλαγή του παραπάνω επιτρέπει σε κάθε γραμμή των ατελών παραγόντων να περιέχει το πολύ p μη μηδενικά στοιχεία στις θέσεις που ήδη υπήρχαν μη μηδενικά στον αρχικό πίνακα A . Αυτό μπορεί να εφαρμοστεί και να δώσει νόημα σε προβλήματα των οποίων στο πίνακα A τα μη μηδενικά στοιχεία δεν είναι κατανομημένα ομοιόμορφα. Ο προορρυθμιστής από τα παραπάνω ορίζεται ως $ILUT(t, p)$ με t το drop tolerance και p το επιτρεπόμενο επίπεδο γεμίσματος και είναι ιδιαίτερα ισχυρός. Σε περίπτωση που αποτύχει για συγκεκριμένες τιμές των t, p τότε αρκεί να δώσουμε μικρότερη τιμή στο t ή μεγαλύτερη στο p .

Αλγόριθμος ILU(0)

Για $i = 2, \dots, n$

Για $k = 1, \dots, i - 1$ και για $(i, k) \in Nz(A)$

Υπολόγισε $a_{ik} = a_{ik}/a_{kk}$

Για $j = k + 1, \dots, n$ και για $(i, j) \in Nz(A)$

Υπολόγισε $a_{i,j} := a_{i,j} - a_{ik}a_{kj}$

Τέλος Για

Τέλος Για

Τέλος Για

4.3.3 ILU(k)

Σε μερικά προβλήματα η $ILLU(0)$ μπορεί να είναι ανεπαρκής για την σύγκλιση. Χρησιμοποιούμε την $ILU(k)$ η οποία εισάγει την έννοια του επιπέδου γεμίματος. Το k είναι ένας μη μηδενικός θετικός ακέραιος. Κάθε στοιχείο της απαλοιφής Gauss έχει ένα επίπεδο γεμίματος το οποίο καθορίζεται από τη σχέση 4.5. Στην περίπτωση που ο πίνακας είναι ισχυρά διαγώνιος όσο μεγαλύτερο είναι το επίπεδο γεμίματος τόσο μικρότερος θα είναι το στοιχείο. Σε κάθε περίπτωση η απόρριψη γίνεται με βάση την τιμή του επιπέδου γεμίματος του κάθε στοιχείου. Στη γενική περίπτωση το μέγεθος δεν εξαρτάται απαραίτητα από το επίπεδο γεμίματος.

ALGORITHM ILU(k)

Για όλα τα μη μηδενικά στοιχεία a_{ij} ορίζουμε $lev(a_{ij}) = 0$

Για $i = 2, \dots, n$

Για κάθε $k = 1, \dots, i - 1$ και για $lev(a_{ik}) \leq k$

Υπολόγισε $a_{ik} = a_{ik}/a_{kk}$

Υπολόγισε $a_{i*} = a_{i*} - a_{ik}a_{k*}$

Ανανεώνω $lev_{ij} = \min\{lev_{ij}, lev_{ik} + lev_{kj} + 1\}$

Τέλος Για

Αντικαθιστώ κάθε της στήλης i με $lev(a_{ij}) > k$ με μηδέν

Τέλος Για

Ο παραπάνω αλγόριθμος έχει αρκετά μειονεκτήματα. Πρώτον, το σύνολο των στοιχείων γεμίματος και ο υπολογιστικός χρόνος μαζί με τον αλγόριθμο ILU(k) δεν μπορούν να προβλεφθούν για $k > 0$. Δεύτερον, το υπολογιστικό κόστος για την ανανέωση του επιπέδου κάθε στοιχείου μπορεί να είναι πολύ μεγάλο.

4.3.4 ILUT

Στον γενικό αλγόριθμο του προορυθμιστή Incomplete LU με ανοχή (threshold) (*ILUT*) εισάγεται μία σειρά κανόνων για την απόρριψη μικρών στοιχείων. Η απόρριψη ενός στοιχείου σημαίνει απλά την αντικατάσταση του με μηδέν, όταν ικανοποιεί το συγκεκριμένο κριτήριο. Ο κανόνας απόρριψης μπορεί να εφαρμοστεί σε μια ολόκληρη γραμμή του πίνακα ελέγχοντας με τον ίδιο κανόνα όλα τα στοιχεία της και απορρίπτοντας τα ανάλογα την τιμή τους.

Τα θεωρήματα που ισχύουν στην *ILUT* είναι όμοια με αυτά της ατελούς παραγοντοποίησης LU. Εάν τα διαγώνια στοιχεία του αρχικού πίνακα είναι θετικά και τα στοιχεία εκτός διαγωνίου αρνητικά τότε κάτω από ορισμένες συνθήκες της αυστηρής κυρίαρχης διαγωνίου, οι πίνακες που παράγονται κατά την παραγοντοποίηση θα έχουν τις ίδιες ιδιότητες [30].

Αλγόριθμος ILUT

Για $i = 1, \dots, n$

$w := a_i^*$

Για $k = 1, \dots, i - 1$ και όταν $w_k \neq 0$

$w_k := w_k / a_{kk}$

Ορίζουμε ένα *dropping rule* για το w_k

Αν $w_k \neq 0$ τότε

$$w := w - w_k * u_{k*}$$

Τέλος Αν

Τέλος Για

Ορίζουμε ένα *dropping rule* για την γραμμή του w_k

$$l_{i,j} := w_j \text{ για } j = 1, \dots, i - 1$$

$$u_{i,j} := w_j \text{ για } j = 1, \dots, n$$

$$w := 0$$

Τέλος Για

4.4 Αναδιάταξη Αγνώστων/Εξισώσεων

Αρκετές φορές η δομή των εξισώσεων του γραμμικού συστήματος οδηγεί σε πίνακες με υψηλό δείκτη κατάστασης. Για αυτό το λόγο είναι αναγκαίος ο μετασχηματισμός του συστήματος με τις κατάλληλες μεταθέσεις, σε ένα άλλο με καλύτερες ιδιότητες. Η διαδικασία μετασχηματισμού του συστήματος ονομάζονται διαδικασία αναδιάταξης αγνώστων και εξισώσεων και χρησιμοποιείται για να

- μειωθούν τα fill in (γεμίσματα),
- για να γίνει εισαγωγή στην εφαρμογή παραλληλοποίησης,
- να αποδείξουν την σταθερότητα των ατελών παραγοντοποιήσεων (incomplete factorization),
- και σε πολλές περιπτώσεις να επηρεαστεί ο ρυθμός σύγκλισης των μεθόδων προ-ρύθμισης του υπόχωρου Krylov [7], [6].

Έστω P και Q μεταθετικοί πίνακες του συστήματος $Ax = b$, το σύστημα μπορεί να αντικατασταθεί από το ισοδύναμο (reordered) $PAQy = Pb$, με $x = Qy$ και υπολογισμό του PAQ . Παρακάτω αναλύονται οι αναδιατάξεις Cuthill-McKee (CMK) και reverse Cuthill-McKee (RCM).

4.4.1 CMK

Η αναδιάταξη Cuthill-McKee (CMK) είναι ένας αλγόριθμος που μετατρέπει ένα αραιό πίνακα συμμετρικής αραιής δομής σε ένα πίνακα μορφής ζώνης (banded) με μικρότερο εύρος. Στον παρακάτω αλγόριθμο γίνεται επιλογή του αρχικού κόμβου για να ξεκινήσει ο πρώτος έλεγχος. Το διάνυσμα p περιέχει την λίστα με του κόμβους με την σειρά που βρίσκονται και ελέγχονται, δηλαδή εκεί φυλάσσονται τα επίπεδα διαδοχής.

Αλγόριθμος CMK

Βρίσκουμε ένα αρχικό κόμβο u

Αρχικοποιούμε $S = \{u\}$, $seen = 1$, $p(seen) = u$ Σημείωσε το u

Αν $seen < n$, τότε

$$S_{new} = \emptyset$$

Για κάθε κόμβο u κάνε

$$p(++seen) = u$$

Για κάθε μη σημειωμένο w σε $adj(u)$ πηγαίνουμε από το μικρότερο σε μεγαλύτερο βαθμό και κάνουμε

$$Add\ w\ to\ S_{new}$$

Σημείωσε το w

Τέλος για

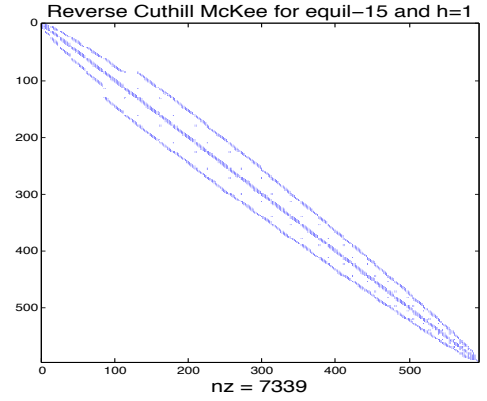
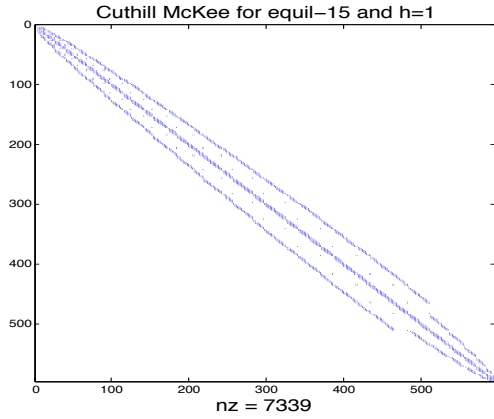
$$S := S_{new}$$

Τέλος για

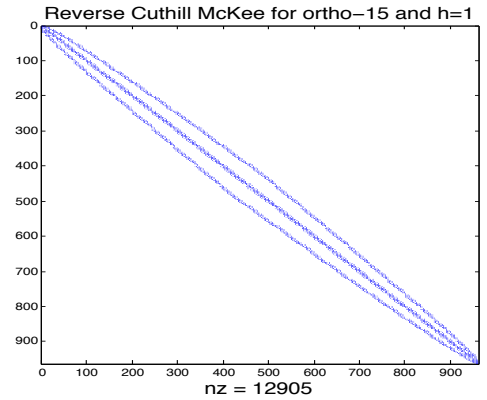
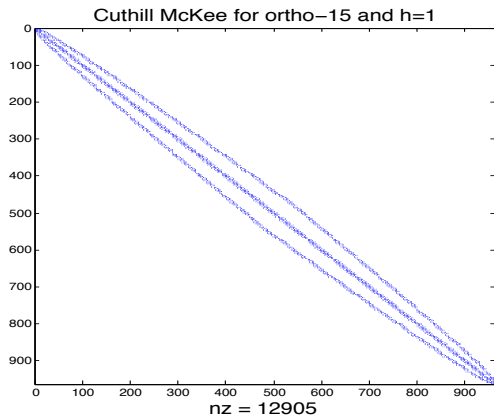
Τέλος αν

4.4.2 RCM

Ο αλγόριθμος αναδιάταξης reverse Cuthill-McKee είναι ακριβώς ο ίδιος αλγόριθμος με την Cuthill-McKee με την διαφορά ότι τα αποτελέσματα που έχουν προκύψει έχουν



Σχήμα 4.1: Μορφή πίνακα τύπου I μετά την χρήση της CMK και της RCM σε μέση στάθμη $h = 1$ με $N_x = 15$.



Σχήμα 4.2: Μορφή πίνακα τύπου II μετά την χρήση της CMK και της RCM σε μέση στάθμη $h = 1$ με $N_x = 15$.

αντιστραμμένους όρους. Έρευνες έχουν δείξει ότι η αναδιάταξη reverse Cuthill-McKee (RCM) δίνει τα καλύτερα αποτελέσματα σε σύγκριση με τις υπόλοιπες μεθόδους, ειδικά όταν εφαρμοστεί με προρρυθμιστή *ILUT* για προσεγγίσεις πεπερασμένων διαφορών για πίνακες αριθμητικά μη συμμετρικούς αλλά με δομή συμμετρικού [7].

Η RCM είναι καλύτερη σε σχέση με τις υπόλοιπες όχι μόνο για την απόδοση της, αλλά και γιατί είναι πολύ ισχυρή με μικρό αριθμό αποτυχιών. Με λίγα λόγια θα μπορούσε κανείς να πει ότι πρέπει να χρησιμοποιείται η RCM ως δεδομένο όταν κάποιος θέλει να χρησιμοποιήσει διαδικασία αναδιάταξης. Η RCM είναι μία μέθοδος που βασίζεται μόνο στην αραιή δομή του πίνακα και όχι στις αριθμητικές τιμές των στοιχείων του και υπερέρχει

της CMK στους τομείς της αποθήκευσης και της υπολογιστικής πολυπλοκότητας.

Στο σχήμα 4.1 φαίνεται η μορφή του πίνακα τύπου I (equilateral) με δομή 2.10 μετά τη χρήση της αναδιάταξης Cuthill-McKee και reverse Cuthill-McKee για μέση στάθμη $h = 1$ και $N_x = 15$. Όμοια μορφή έχουν όλοι οι πίνακες αυτού του τύπου ανεξαρτήτως της τιμής της μέσης στάθμης. Όλοι οι υπόλοιποι τύποι πλεγμάτων μορφής 2.9, 2.15, 2.18 μετά την χρήση της αναδιάταξης έχουν μορφή 4.2.

Κεφάλαιο 5

Αριθμητικά Αποτελέσματα

Σε αυτό το κεφάλαιο γίνεται σύγκριση της συμπεριφοράς σύγκλισης μεθόδων για την επίλυση γραμμικών συστημάτων που παράγονται για τους πίνακες τύπων I,II,III,IV που προκύπτουν από τα αντίστοιχα πλέγματα που παρουσιάστηκαν στην παράγραφο 2.4. Αρχικά, μελετήθηκε η περίπτωση με σταθερό βάθος $h = 1$ και κριτήριο απόρριψης το $tol = 10^{-5}$ για τους προορρυθμιστές και στη συνέχεια εξετάζονται διαφορετικές τιμές της μέσης στάθμης. Ορίζουμε μέση στάθμη την απόσταση της τοπογραφίας από ένα σταθερό σημείο αναφοράς (βλ. Σχήμα 2.1). Οι πίνακες προκύπτουν από το υπολογιστικό χωρίο $[0, 1] \times [0, 1]$. Στον πίνακα 5.1 παρουσιάζονται η διάσταση των πινάκων $2N \times 2N$ και το σύνολο των μη μηδενικών στοιχείων (Nz) για τα πλέγματα τύπου I,II,III και IV με συνεπή πύκνωση $N_x = 15$, $N_x = 30$ και $N_x = 60$ για μέση στάθμη $h = 1m$, $h = 10m$, και $h = 100m$.

5.1 Μελέτη πινάκων για μέση στάθμη $h = 1m$.

Στους Πίνακες 5.2 παρουσιάζονται οι απαιτούμενες επαναλήψεις για την σύγκλιση των επαναληπτικών μεθόδων GMRES, BiCGSTAB, CG και TFQMR, οι οποίες εφαρμόστηκαν σε όλους τους τύπους πλεγμάτων με συνεπή πύκνωση του κάθε πλέγματος. Για την ορθότερη σύγκριση των μεθόδων ανά τύπο πλέγματος παρουσιάζονται και τα αποτελέσματα, σε παρένθεση, με χρήση προορρυθμιστή ILUT(300) και κριτήριο απόρριψης στοιχείων το 10^{-5} .

Τύποι πλεγμάτων για μέση στάθμη $h = 1m$, $2N \times 2N$ (Nz)				
N_x	Τύπος I	Τύπος II	Τύπος III	Τύπος IV
15	352836 (7339)	925444 (12905)	262144 (6430)	352836 (7715)
30	4857616 (28436)	13853284 (50806)	3694084 (25163)	4857616 (29746)
60	74132100 (118293)	214388164 (201931)	55383364 (99258)	74132100 (118293)
Τύποι πλεγμάτων για μέση στάθμη $h = 10m$, (Nz)				
N_x	Τύπος I	Τύπος II	Τύπος III	Τύπος IV
15	(7339)	(12905)	(6430)	(7715)
30	(28436)	(50806)	(25163)	(29746)
60	(114196)	(201931)	(99258)	(118293)
Τύποι πλεγμάτων για μέση στάθμη $h = 100$, (Nz)				
N_x	Τύπος I	Τύπος II	Τύπος III	Τύπος IV
15	(7457)	(12909)	(6422)	(7713)
30	(28839)	(50828)	(25167)	(29741)
60	(114987)	(202073)	(99253)	(7714)

Πίνακας 5.1: Διάσταση πινάκων για διάφορα βάθη.

Από τους πίνακες φαίνεται ότι καθώς πυκνώνουμε το πλέγμα οι επαναλήψεις αυξάνονται, δεδομένο που το περιμέναμε καθώς ο πίνακας έχει πλέον μεγαλύτερη διάσταση και για την επίλυση του θα χρειαστούν περισσότερες επαναλήψεις.

Επίσης, παρατηρούμε ότι ο αριθμός των επαναλήψεων διαφέρει στις διάφορες επαναληπτικές μεθόδους. Αρχικά, στους πίνακες τύπου I (equilateral) στο αραιότερο πλέγμα με την μέθοδο GMRES είχαμε το μικρότερο αριθμό επαναλήψεων και με την CG τον μεγαλύτερο αριθμό, ενώ στο πυκνότερο παράδειγμα η BiCGSTAB έκανε τις λιγότερες επαναλήψεις και η GMRES τις περισσότερες. Με την χρήση προρρυθμιστή λοιπόν έχουμε τεράστια διαφορά στον αριθμό των επαναλήψεων, όπως και ήταν αναμενόμενο. Υπάρχουν περιπτώσεις όμως, που με τη χρήση προρρυθμιστή δεν επιτυγχάνεται σύγκλιση, συγκεκριμένα σε πίνακες τύπου II,IV.

Παρακάτω στους πίνακες 5.3 φαίνονται οι επαναλήψεις για τα πλέγματα τύπου II, IV με χρήση προρρυθμιστή ILU0. Παρατηρούμε ότι δεν βελτιώνει επιθυμητά τον αριθμό των επαναλήψεων. Παράλληλα, έγιναν δοκιμές για τα πλέγματα τύπου I, III και

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος I, $h = 1$, $ILUT$ $lfil = 300$ $drop = 10^{-5}$	15	279 (4)	281 (5)	385 (5)	301 (5)
	30	810 (5)	66 (5)	1712 (6)	631 (7)
	60	2572 (7)	1459 (7)	1203 (11)	1213 (9)

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος II, $h = 1$, $ILUT$ $lfil = 300$ $drop = 10^{-5}$	15	262 (15)	219 (17)	337 (-)	197 (17)
	30	536 (-)	315 (-)	- (-)	311 (-)
	60	2272 (-)	599 (-)	- (-)	611 (-)

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος III, $h = 1$, $ILUT$ $lfil = 300$ $drop = 10^{-5}$	15	267 (4)	213 (5)	230 (4)	3617 (5)
	30	727 (4)	495 (5)	427 (4)	- (5)
	60	2450 (7)	1029 (7)	869 (7)	914 (7)

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος IV, $h = 1$, $ILUT$ $lfil = 300$ $drop = 10^{-5}$	15	1999 (4)	581 (5)	- (5)	649 (5)
	30	- (6)	- (7)	- (7)	- (7)
	60	- (8)	3365 (9)	- (-)	- (9)

Πίνακας 5.2: Πλήθος επαναλήψεων χωρίς και με τη χρήση προρρυθμιστή σε παρένθεση $ILUT$ για μέση στάθμη $h = 1m$ στα διάφορα πλέγματα.

σημειώθηκε ότι ο προρρυθμιστής δεν συγκλίνει για καμία επαναληπτική μέθοδο. Η επιλογή του προρρυθμιστή $ILU0$ μειώνει σημαντικά το συνολικό χρόνο υπολογισμού με το μειονέκτημα μιας κακής προσέγγισης του αντίστροφου πίνακα.

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος II, $h = 1$, $ILU0$	15	-	34957	-	-
	30	-	-	-	-
	60	-	-	-	-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος IV, $h = 1$, $ILU0$	15	206	181	-	187
	30	499	363	-	361
	60	1326	729	-	777

Πίνακας 5.3: Επαναλήψεις με τη χρήση Preconditioner $ILU0$, για μέση στάθμη $h = 1$.

Στους πίνακες 5.4 παρουσιάζονται τα αποτελέσματα των μεθόδων με επιλογή προρρυθμιστή ILU_k με $k = 300$ ($k :=$ επίπεδο γεμίσματος).

Διαπιστώνεται ότι οι μέθοδοι με προρρυθμιστή τον $ILUT$ (πίνακας 5.2) εκτελούν μεγαλύτερο αριθμό επαναλήψεων σε σχέση με την επιλογή της προρρυθμίστης ILU_k , αφού η προσέγγιση του πίνακα A^{-1} είναι καλύτερη στη δεύτερη περίπτωση. Αυτό οφείλεται στο γεγονός ότι ο προρρυθμιστής ILU_k δεν έχει επιπλέον στρατηγική απόρριψης στοιχείων. Φυσικά το κόστος υπολογισμού στο χρόνο εκτέλεσης του προρρυθμιστή είναι ανάλογο. Για παράδειγμα η ILU_k στο πλέγμα τύπου III για $N_x = 60$ χρειάζεται $2.24s$ με τη μέθοδο GMRES ενώ στην αντίστοιχη περίπτωση ο $ILUT$ χρειάζεται $1.58s$. Στους Πίνακες 5.5 παρουσιάζεται η χρονομέτρηση του προρρυθμιστή ILU_k για κάθε επαναληπτική μέθοδο.

Έπειτα, στους πίνακες 5.6 έγιναν πειράματα για τις επαναλήψεις που χρειάστηκαν για την επίλυση του προβλήματος με την χρήση μόνο Reordering CMK. Ο αριθμός των επαναλήψεων είναι μεγάλος και σε πλέγματα τύπου IV παρατηρούμε ότι δεν επιτυγχάνεται ούτε καν η σύγκλιση. Γενικά, ο χρόνος που χρειάζεται για να ολοκληρωθεί η μέθοδος είναι αρκετά μικρός και έτσι οδηγούμαστε στο συμπέρασμα ότι η χρήση του Reordering είναι μία σχεδόν ανέξοδη διαδικασία που σε συνδυασμό με preconditioning θα μπορέσει

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος I, $h = 1$, $ILUK(300)$	15	2	3	2	3
	30	2	3	2	3
	60	2	3	2	3

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος II, $h = 1$, $ILUK(300)$	15	2	3	2	3
	30	2	3	2	3
	60	-	-	-	-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος III, $h = 1$, $ILUK(300)$	15	2	3	2	3
	30	2	3	2	3
	60	2	3	2	3

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος IV, $h = 1$, $ILUK(300)$	15	2	3	2	3
	30	2	3	2	3
	60	3	3	3	3

Πίνακας 5.4: Επαναλήψεις με Preconditioner ILUK για όλους τους τύπους πλεγμάτων με μέση στάθμη $h = 1m$.

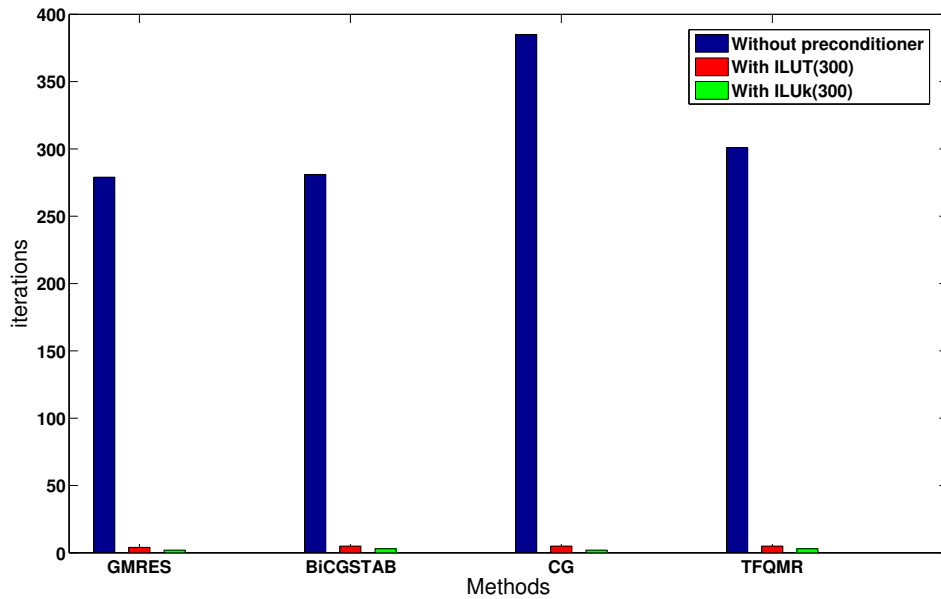
να φέρει τα επιθυμητά αποτελέσματα όπως θα παρατηρήσουμε παρακάτω.

Παρακάτω φαίνονται οι επαναλήψεις και ο χρόνος που χρειάζεται για την επίλυση του γραμμικού συστήματος με την εφαρμογή των δύο καλύτερων μεθόδων με βάση τα αποτελέσματα μας, της GMRES και της μεθόδου BiCGSTAB. Με T συμβολίζουμε τον συνολικό χρόνο (total time) που χρειάστηκε ο αλγόριθμος μας για να επιλύσει το

Περιγραφή	N_x	Τύπος I	Τύπος II	Τύπος III	Τύπος IV
Χρόνος (s) για, $ILUT(300)$, $drop = 10^{-5}$	15	0.013	1.006	0.01	0.015
	30	0.13	-	0.11	0.16
	60	1.84	-	1.58	1.95

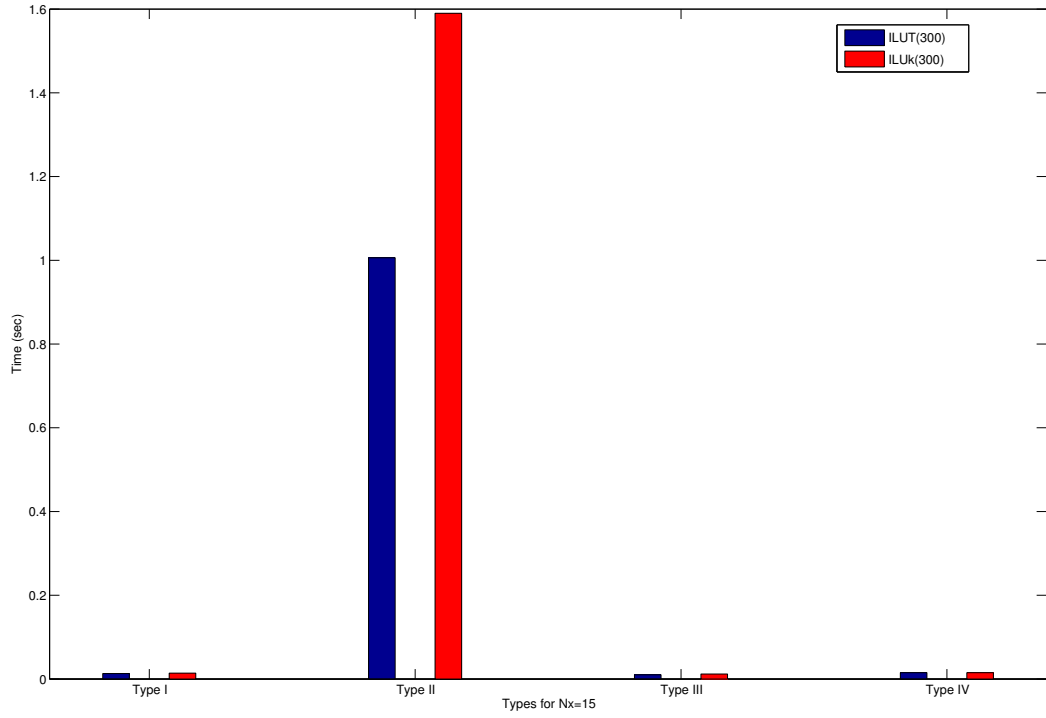
Περιγραφή	N_x	Τύπος I	Τύπος II	Τύπος III	Τύπος IV
Χρόνος (s) για, $ILUK(300)$	15	0.014	1.59	0.012	0.015
	30	0.19	87.3	0.15	0.19
	60	2.62	-	2.24	2.69

Πίνακας 5.5: Υπολογιστικός χρόνος για Preconditioner ILUK, $lfil = 300$



Σχήμα 5.1: Συγκριτικό ραβδόγραμμα των προρρυθμιστών ILUT και ILUK για $N_x = 15$ και μέση στάθμη $h = 1m$.

γραμμικό σύστημα και με TP τον χρόνο που χρειάζεται ο προρρυθμιστής και με TR τον χρόνο που θέλει το reordering. Το reordering βελτίωσε κατά πολύ τις ιδιότητες



Σχήμα 5.2: Συγκριτικό ραβδόγραμμα για το χρόνο εκτέλεσης των προρρυθμιστών ILUT και ILUK για $N_x = 15$ και μέση στάθμη $h = 1m$.

των πινάκων αφού με χρήση του RCM συγκλίνουν όλες οι μέθοδοι σε κάθε τύπο και διάσταση.

5.2 Μελέτη πινάκων για μέση στάθμη $h = 10m$.

Στους παρακάτω πίνακες έχουν δοκιμαστεί όλες οι δυνατές επιλογές για προρρύθμιση και reordering για όλους τους πίνακες που έχουν προκύψει από τα διαφορετικά είδη πλεγμάτων σε διάφορες τιμές μέσης στάθμης. Αρχικά, δοκιμάστηκαν όλοι οι πίνακες που προκύπτουν για μέση στάθμη $h = 100$ χρησιμοποιώντας την *ILUT* σε συνδυασμό με την *RCM* μελετώντας τις παραμέτρους *lfil* και *drop tolerance*. Παρατηρούμε ότι ανεξαρτήτως του *lfil* για τιμές του *drop tolerance* 10^{-5} και μεγαλύτερες κανένας πίνακας δεν συγκλίνει. Αυτό οφείλεται στον κακό δείκτη κατάστασης του πίνακα, καθώς και στο

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος I, $h = 1$, CMK	15	279	271	385	301
	30	814	647	1712	647
	60	2179	1527	1199	1209

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος II, $h = 1$, CMK	15	195	151	-	161
	30	536	311	-	311
	60	2401	621	-	615

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος III, $h = 1$, CMK	15	267	227	230	422
	30	728	489	427	285
	60	2463	987	870	-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος IV, $h = 1$, CMK	15	1952	595	-	645
	30	-	-	-	-
	60	-	-	-	-

Πίνακας 5.6: Επαναλήψεις με τη χρήση CMK για μέση στάθμη $h = 1$.

φάσμα των ιδιοτιμών.

Δοκιμάζοντας όμως μικρότερο *drop tolerance* το 10^{-10} και κρατώντας σταθερές τις υπόλοιπες παραμέτρους παρατηρούμε σύγκλιση για κάθε μέθοδο όπου φαίνονται αναλυτικά στον πίνακα. Σύμφωνα με την εργασία του [14] ο προρρυθμιστής *ILUT* δουλεύει αρκετά καλά σε ρηχά σε σύγκριση με βαθιά νερά. Όσο η μέση στάθμη αυξάνει χάνει την αποτελεσματικότητά του. Αυτό το συμπέρασμα προκύπτει και από τα δικά μας αποτελέσματα.

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος I, $h = 1$, CMK	15	3.03E-4	2.98E-4	2.98E-4	3.13E-4
	30	1.06E-3	1.087E-3	1.09E-3	1.07E-3
	60	4.24E-3	4.34E-3	4.52E-3	4.29E-3

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος II, $h = 1$, CMK	15	5.07E-4	5.19E-4	-	4.9E-4
	30	2.71E-3	1.91E-3	-	1.88E-3
	60	7.4E-3	7.54E-3	-	7.49E-3

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος III, $h = 1$, CMK	15	4.28E-4	4.37E-4	4.3E-4	4.44E-4
	30	1.59E-3	1.71E-3	1.56E-3	1.26E-3
	60	4.76E-3	3.57E-3	3.63E-3	-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος IV, $h = 1$, CMK	15	4.85E-4	4.87E-4	-	4.87-4
	30	-	-	-	-
	60	-	-	-	-

Πίνακας 5.7: Χρόνος που χρειάζεται η CMK για πίνακα μέσης στάθμης $h = 1$.

Για μέση στάθμη $h = 100$, η μέθοδος δε συγκλίνει σε καμία σχεδόν περίπτωση και για κανένα τύπο πλέγματος. Όμως, για βάθος $h = 10$ τα αποτελέσματα της μεθόδου είναι τα επιθυμητά. Λόγω των προβλημάτων της $ILUT$ για μεγαλύτερα βάθη, μελετήθηκε και χρησιμοποιήθηκε η $ILLUk$, μία μέθοδος προρρυθμισμού που δε περιέχει *drop tolerance*. Με αυτή τη μέθοδο παρατηρήθηκε ότι οι πίνακες μας συγκλίνουν για μεγάλα βάθη και έτσι αποτελεί την εναλλακτική μέθοδο για την επίλυση γραμμικών συστημάτων που προ-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος I, $h = 1$, RCM	15	279	291	385	299
	30	811	661	1712	653
	60	2416	1427	1199	1007

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος II, $h = 1$, RCM	15	195	159	-	161
	30	536	315	-	313
	60	2237	595	-	617

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος III, $h = 1$, RCM	15	264	215	230	-
	30	725	445	426	-
	60	2426	983	870	-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Τύπος IV, $h = 1$, RCM	15	1875	515	-	643
	30	-	-	-	-
	60	-	-	-	-

Πίνακας 5.8: Επαναλήψεις που χρειάζεται το RCM για μέση στάθμη $h=1$.

κύπτουν από μεγάλα h .

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Χρόνος (s) για :, Τύπος I, $h = 1$, RCM	15	3.13E-4	5.10E-4	5.08E-4	5.08E-4
	30	1.84E-3	1.86E-3	1.84E-3	1.84E-3
	60	7.2E-3	7.37E-3	5.64E-3	7.32E-3

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Χρόνος (s) για :, Τύπος II, $h = 1$, RCM	15	4.99E-4	8.80E-4	-	4.94E-4
	30	2.07E-4	3.26E-4	-	3.16E-4
	60	9.93E-4	7.75E-4	-	8.00E-4

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Χρόνος (s) για :, Τύπος III, $h = 1$, RCM	15	4.43E-4	4.44E-4	2.74E-4	-
	30	1.63E-3	1.67E-3	9.69E-4	-
	60	3.74E-4	6.38E-4	6.29E-4	-

Περιγραφή	N_x	GMRES	BiCGSTAB	CG	TFQMR
Χρόνος (s) για :, Τύπος IV, $h = 1$, RCM	15	3.07E-4	5.41E-4	-	3.11E-4
	30	-	-	-	-
	60	-	-	-	-

Πίνακας 5.9: Χρόνος για RCM σε μέση στάθμη $h = 1$.

5.3 Μελέτη πινάκων για βάθος $h = 100m$.

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος I, $h = 1$, $tol = 10^{-5}$, $lfil = 300$, $ILUT + RCM$	15	4	0.021	0.018	4.67E-4	5	0.018	0.015	5.1E-4
	30	5	0.16	0.150	1.85E-3	5	0.19	0.17	1.87E-3
	60	7	2.20	2.07	4.86E-3	7	2.23	2.11	6.02E-3

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος II, $h = 1$, $tol = 10^{-5}$, $lfil = 300$, $ILUT + RCM$	15	4	0.040	0.03	1.17E-3	5	0.040	0.038	5.52E-4
	30	5	0.45	0.42	3.13E-3	5	0.45	0.42	3.20E-3
	60	7	6.69	6.42	0.01	7	7.29	7.02	0.015

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος III, $h = 1$, $tol = 10^{-5}$, $lfil = 300$, $ILUT + RCM$	15	4	0.015	0.013	2.8E-4	5	0.022	0.020	4.4E-4
	30	5	0.21	0.19	1.07E-3	5	0.197	0.18	1.63E-3
	60	7	2.69	2.57	6.1E-3	7	2.68	2.59	3.6E-3

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος IV, $h = 1$, $tol = 10^{-5}$, $lfil = 300$, $ILUT + RCM$	15	4	0.017	0.015	5.1E-4	5	0.026	0.023	5.04E-4
	30	5	0.17	0.15	1.8E-3	5	0.17	0.16	1.85E-3
	60	8	2.36	2.22	7.2E-3	9	2.60	2.44	7.05E-3

Πίνακας 5.10: RCM+ILUT για όλους τους τύπους πλεγμάτων

Μελετώντας τα αποτελέσματα για βάθος $h = 100$ με προρρύθμιση ILUk και reordering RCM παρατηρούμε ότι ο αριθμός των επαναλήψεων είναι σχετικά μικρός.

Για παράδειγμα στους πίνακες που προκύπτουν από πλέγματα μορφής III για $N_x = 60$ πραγματοποιούνται 3 επαναλήψεις με BiCGSTAB και 3 με GMRES. Στην αντίστοιχη περίπτωση εάν άλλαζε ο προρρυθμιστής και χρησιμοποιούσαμε τον ILUT η μέθοδος δε συγκλίνει καν.

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος I, $h = 10$, $tol = 10^{-5}$, $l_{fil} = 300$, $ILUT + RCM$	15	7	0.030	0.015	5.160E-4	7	0.020	0.019	5.40E-4
	30	12	0.18	0.15	1.85E-3	13	0.18	0.15	1.87E-3
	60	28	2.55	2.04	7.03E-3	29	2.76	2.80	7.330E-3
Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος II, $h = 10$, $tol = 10^{-5}$, $l_{fil} = 300$, $ILUT + RCM$	15	9	0.044	0.035	8.4E-4	9	0.039	0.032	5.24E-4
	30	13	0.57	0.50	3.16E-3	15	0.50	0.43	3.12E-3
	60	453	22.7	6.20	8.61E-3	295	16.6	6.19	8.75E-3
Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος III, $h = 10$, $tol = 10^{-5}$, $l_{fil} = 300$, $ILUT + RCM$	15	7	0.022	0.019	4.3E-4	9	0.018	0.014	4.470E-4
	30	13	0.210	0.17	1.60E-4	15	0.21	0.170	1.59E-3
	60	57	3.29	2.38	6.270E-3	53	3.25	2.39	6.320E-3
Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος IV, $h = 10$, $tol = 10^{-5}$, $l_{fil} = 300$, $ILUT + RCM$	15	9	0.019	0.015	5.11E-4	11	0.017	0.012	3.1E-4
	30	24	0.21	0.015	1.92E-3	23	0.21	0.015	1.93E-3
	60	781	15.3	2.22	7.5E-3	225	6.14	2.48	7.36E-3

Πίνακας 5.11: $ILUT + RCM$ για βάθος $h = 10$ με $tol = 10^{-5}$.

Για παράδειγμα στο ίδιο τύπου πλέγμα (III) για βάθος 10 οι επαναληπτικές μέθοδοι GMRES και BiCGSTAB με προρρύθμιση ILUT και reordering RCM χρειάζονται 7-9 επαναλήψεις και συνολικό χρόνο 0.018s χρησιμοποιώντας κριτήριο απόρριψης 10^{-5} και 4-5 σε χρόνο 0.044s. Στην ίδια περίπτωση η ILUk με RCM χρειάζεται το μισό αριθμό επαναλήψεων 2-3 με χρόνο 0.020s. Στα ραβδογράμματα Σχήμα 5.3 και 5.4 γίνεται

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος I, $h = 10$, $tol = 10^{-10}$, $lfil = 300$, $ILUT + RCM$	15	3	0.021	0.019	5.11E-4	3	0.020	0.017	5.37E-4
	30	5	0.17	0.16	2.02E-3	5	0.17	0.15	1.85E-3
	60	5	2.30	2.21	7.47E-3	5	2.33	2.24	7.65E-3

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος II, $h = 10$, $tol = 10^{-10}$, $lfil = 300$, $ILUT + RCM$	15	3	0.044	0.040	8.38E-4	3	0.043	0.039	9.3E-4
	30	5	0.45	0.43	2.6E-3	5	0.45	0.42	1.9E-3
	60	13	6.67	6.19	7.47E-3	15	6.75	6.19	0.012

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος III, $h = 10$, $tol = 10^{-10}$, $lfil = 300$, $ILUT + RCM$	15	4	0.019	0.017	4.48E-4	5	0.018	0.015	4.68E-4
	30	4	0.20	0.19	1.66E-3	5	0.24	0.22	2.13E-3
	60	5	3.28	3.19	5.39E-3	5	3.33	3.23	3.87E-3

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος IV, $h = 10$, $tol = 10^{-10}$, $lfil = 300$, $ILUT + RCM$	15	4	0.022	0.019	5.35E-4	5	0.023	0.021	5.86E-4
	30	5	0.17	0.16	1.88E-3	5	0.17	0.15	1.93E-3
	60	6	2.61	2.50	7.42E-3	7	2.36	2.23	7.4E-3

Πίνακας 5.12: $ILUT + RCM$ για βάθος $h = 10$ με $tol = 10^{-10}$.

σύγκριση για βάθη 10 των προρρυθμιστών ILUT και ILUk όπου παρατηρείται η αναγκαιότητα της χρήσης κάποιου είδους προρρύθμισης για την μείωση των επαναλήψεων, είτε με την επαναληπτική μέθοδο GMRES είτε με την BiCGSTAB.

Παράλληλα, γίνεται σύγκριση των επαναληπτικών μεθόδων χωρίς και με τη χρήση προρρυθμιστή. Παρατηρούμε ότι η ILUk χρειάζεται μικρότερο αριθμό επαναληπτικών βημάτων.

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος I,	15	3	0.017	0.015	3.25E-4	3	0.024	0.022	5.17E-4
$h = 10$	30	3	0.21	0.20	1.09E-3	3	0.22	0.21	1.99E-3
$lfil = 300$	60	3	3.10	3.04	6.54E-3	3	3.38	3.23	4.39E-3
$ILUK + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος II,	15	3	0.055	0.051	8.3E-4	3	0.045	0.041	5.09E-4
$h = 10$	30	3	0.59	0.57	1.86E-3	3	0.68	0.60	3.18E-3
$lfil = 300$	60	3	8.69	8.57	7.74E-3	3	8.92	8.80	8.93E-3
$ILUK + RCM$									

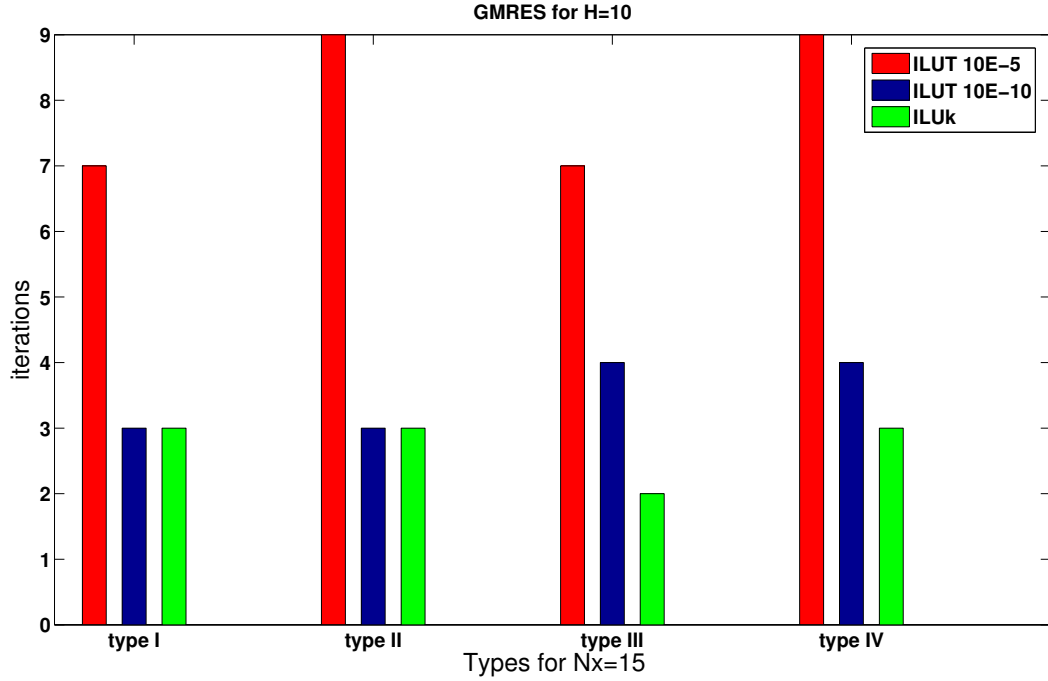
Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος III,	15	2	0.020	0.018	2.67E-4	3	0.020	0.019	2.65E-4
$h = 10$	30	3	0.285	0.275	9.66E-4	3	0.39	0.37	1.31E-4
$lfil = 300$	60	3	4.63	4.56	6.12E-3	3	4.76	4.69	3.95E-3
$ILUK + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος IV,	15	3	0.019	0.017	3.08E-4	3	0.019	0.017	5.31E-4
$h = 10$	30	3	0.22	0.21	1.25E-3	3	0.21	0.208	1.34E-3
$lfil = 300$	60	4	4.46	3.38	7.45E-3	3	3.45	3.39	4.38E-3
$ILUK + RCM$									

Πίνακας 5.13: $ILUK + RCM$ γαι βάθος $h = 10$

5.4 Μελέτη πινάκων με τοπογραφία.

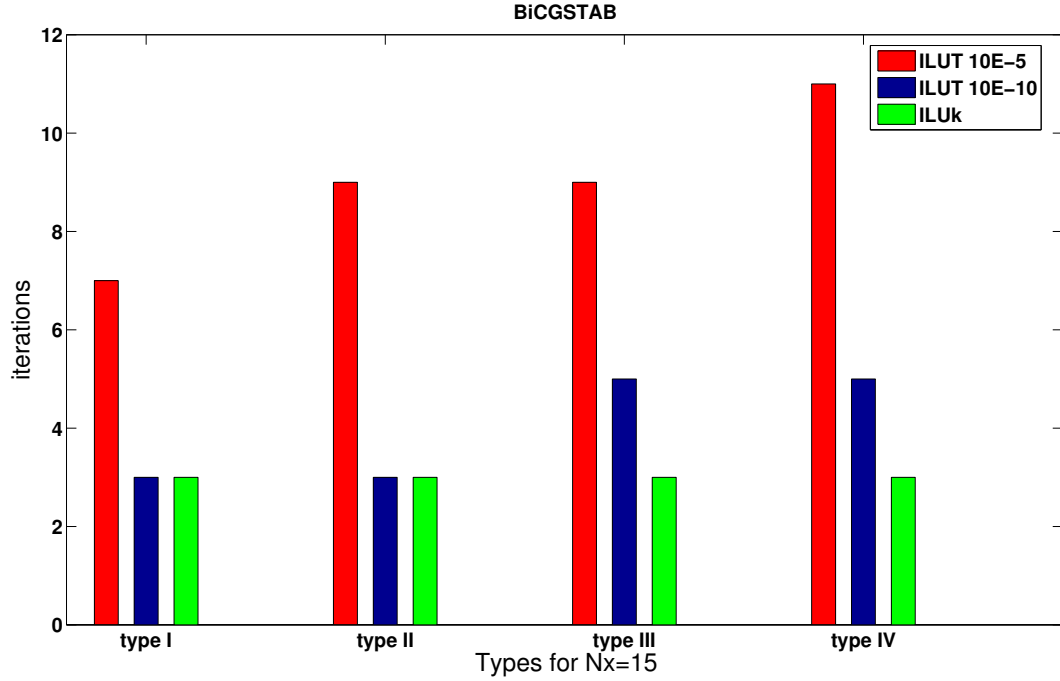
Παρακάτω παρουσιάζονται τα αποτελέσματα που προκύπτουν από ένα πείραμα χωρίου διάστασης $x \in [-10, 29]$ και $y \in [0, 0.8]$ για πλέγμα τύπου I. Μελετήθηκαν τα βάθη 0.1, 10, 100 με $N_x = 180$ τοποθετώντας και τοπογραφία της μορφής που φαίνεται στο σχήμα 5.6. Όταν έχουμε τοποθέτηση τοπογραφίας το h είναι πλέον μεταβλητό και αλλάζουν οι



Σχήμα 5.3: Συγκριτικό ραβδόγραμμα της GMRES και των προρρυθμιστών ILUT και ILUk για $N_x = 15$ και βάθος $h = 10m$.

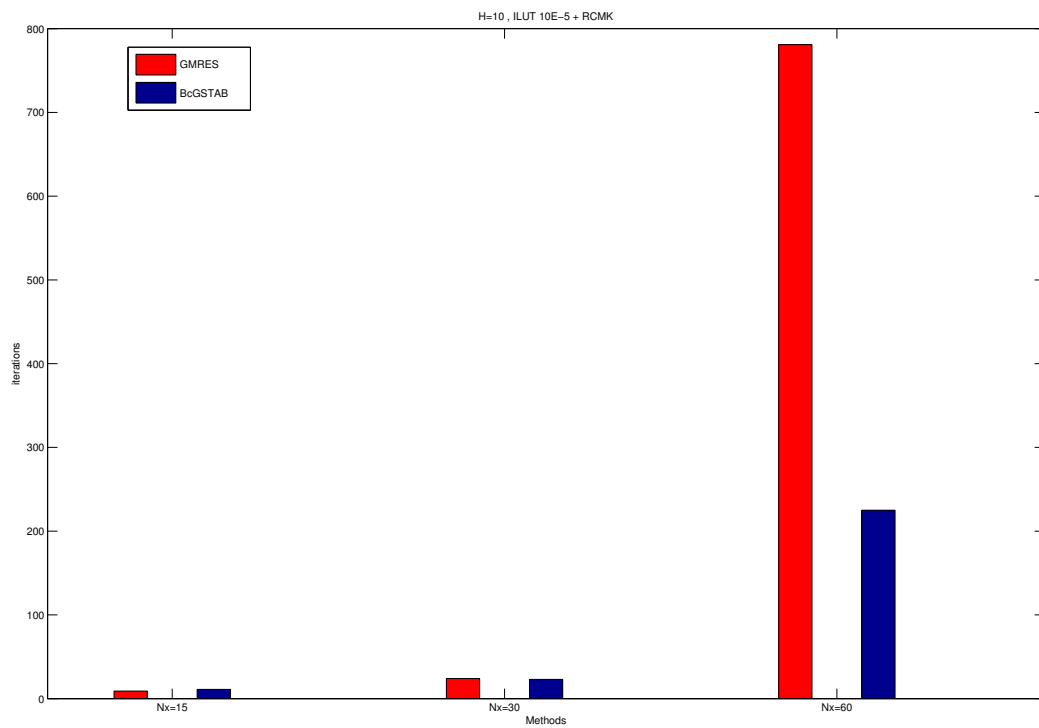
τιμές του πίνακα A . Αρχικά, για βάθος 0.1 χωρίς τοπογραφία οι επαναληπτικές μέθοδοι GMRES και BiCGSTAB χρειάστηκαν 46 και 53 επαναλήψεις αντίστοιχα για να ολοκληρώσουν τη μέθοδο. Χρησιμοποιώντας όμως τον προρρυθμιστή ILUT και RCM. Είδαμε λοιπόν ότι οι επαναλήψεις μειώθηκαν στις 3 και ο συνολικός χρόνος είναι 1.67s. Στη περίπτωση που χρησιμοποιηθεί ο προρρυθμιστής ILUk μπορεί να έχουμε μείωση 1-2 επαναλήψεις αλλά ο χρόνος θα αυξηθεί αρκετά. Το αποτέλεσμα αυτό είναι αναμενόμενο καθώς ο ILUk είναι ένας ακριβός προρρυθμιστής και όπως έχουμε συμπεράνει είναι αποδοτικότερος σε γραμμικά συστήματα που προκύπτουν για μεγαλύτερα βάθη. Στο πείραμα με την τοπογραφία και βάθος $h = 0.1$ αυξήθηκε το κριτήριο απόρριψης στοιχείων του προρρυθμιστή από 10^{-5} σε 10^{-10} για τον προρρυθμιστή ILUT και παρατηρήθηκε μείωση των επαναλήψεων όμως αυξήθηκε ο χρόνος.

Έπειτα, για βάθος 10 στο ίδιο τύπο πλέγματος χωρίς τοπογραφία οι επαναληπτικές μέθοδοι GMRES και BiCGSTAB χρειάζονται 90.000 και 9.339 επαναλήψεις αντίστοιχα.



Σχήμα 5.4: Συγκριτικό ραβδόγραμμα της BiCGSTAB και των προρρυθμιστών ILUT και ILUK για $N_x = 15$ και βάθος $h = 10m$.

Αν γίνει χρήση προρρυθμιστή ILUT ή ILUK και RCM οι επαναλήψεις πέφτουν στις 3 και ο συνολικός χρόνος από 751s και 48s χρειάζεται μόνο 2-3s. Παρόμοια αποτελέσματα έχουμε και στην περίπτωση που υπάρχει τοπογραφία. Τέλος, έχουμε αποτελέσματα για βάθος 100 χωρίς να υπάρχει τοπογραφία. Σε αυτή την περίπτωση οι επαναληπτικές μέθοδοι GMRES και BiCGSTAB δε συγκλίνουν. Χρησιμοποιήθηκαν και συνδυάστηκαν ο προρρυθμιστής ILUT με reordering RCM και για drop tolerance 10^{-5} η μέθοδος αποκλίνει. Έτσι χρησιμοποιήθηκε η τιμή 10^{-10} και παρατηρήθηκε σύγκλιση σε 4 επαναλήψεις με την GMRES σε συνολικό χρόνο 2.23 και 5 επαναλήψεις με την BiCGSTAB σε 2.27s. Στις παραπάνω περιπτώσεις ο χρόνος μόνο για τον προρρυθμιστή είναι 2s. Αν αλλάξουμε τον προρρυθμιστή και χρησιμοποιήσουμε τον ILUK παρατηρούμε μείωση 1 με 2 επαναλήψεις με συνολικό χρόνο 2.79s και χρόνο προρρύθμισης 2.61s. Με τον ILUK οι επαναλήψεις μειώνονται χωρίς να χρειάζεται να μεταβάλλουμε το drop tolerance.



Σχήμα 5.5: Συγκριτικό ραβδόγραμμα των μεθόδων GMRES και BicGSTAB μέσης στάθμης $h = 10m$.

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος I,	15	3	0.018	0.016	5.7E-4	3	0.023	0.020	5.05E-4
$h = 100$	30	3	0.21	0.20	1.86E-3	3	0.21	0.20	1.85E-3
$lfil = 300$	60	3	3.23	3.16	7.44E-3	3	3.09	3.02	7.37E-3
$ILUK + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος II,	15	3	0.054	0.051	8.25E-4	3	0.051	0.047	8.34E-4
$h = 100$	30	4	0.59	0.57	2.40E-3	5	0.60	0.57	3.15E-3
$lfil = 300$	60	7	8.81	8.53	8.03E-3	9	8.85	8.51	8.68E-3
$ILUK + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος III,	15	3	0.029	0.026	4.60E-4	3	0.021	0.019	4.68E-4
$h = 100$	30	3	0.32	0.31	1.61E-3	3	0.28	0.27	1.61E-3
$lfil = 300$	60	3	4.40	4.33	3.99E-3	3	4.38	4.32	5.22E-3
$ILUK + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος IV,	15	3	0.019	0.017	5.4E-4	3	0.021	0.019	5.099E-4
$h = 100$	30	4	0.22	0.21	1.88E-3	5	0.22	0.21	1.86E-3
$lfil = 300$	60	4	3.45	3.38	5.35E-3	5	3.16	3.06	7.50E-3
$ILUK + RCM$									

Πίνακας 5.14: $ILUK + RCM$ για βάθος $h = 100$.

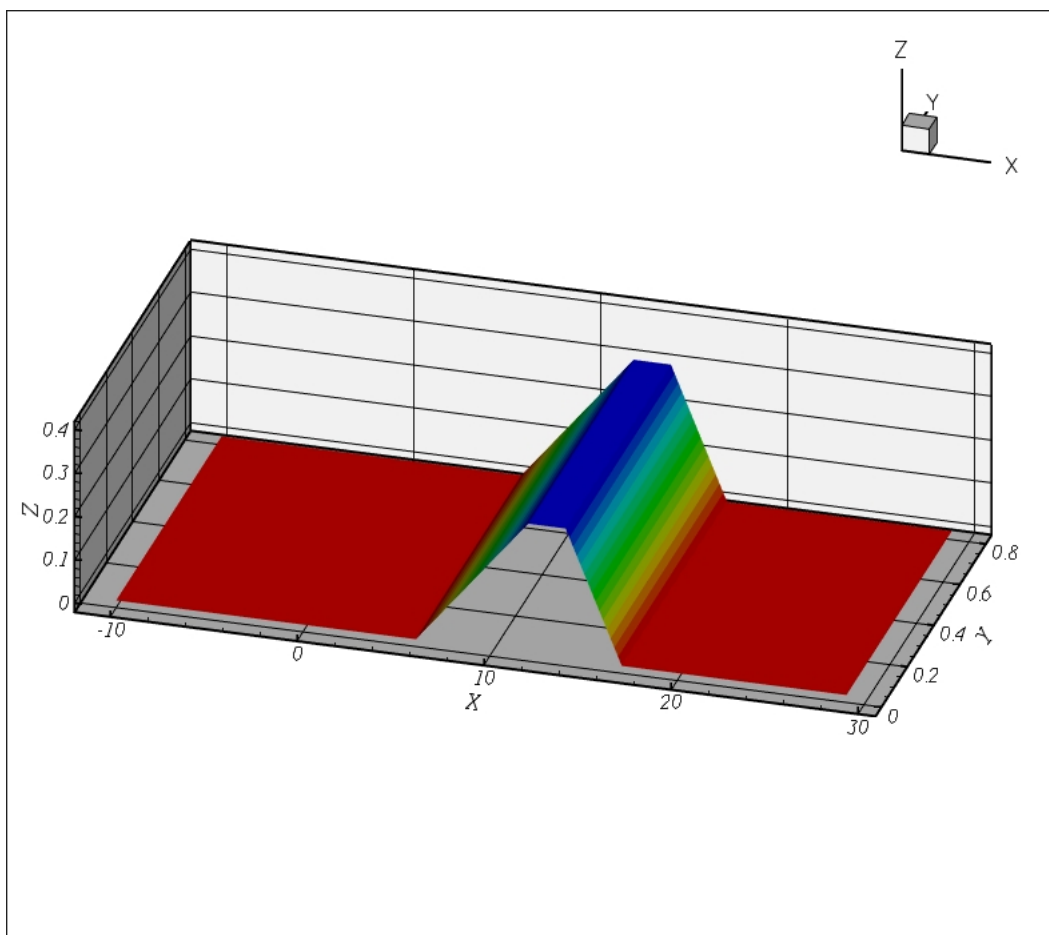
Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος Ι, $tol = 10^{-10}$ $h = 100$ $l_{fil} = 300$ $ILUT + RCM$	15	3	0.014	0.012	5.26E-4	3	0.018	0.016	5.4E-4
	30	5	0.16	0.15	1.82E-3	5	0.19	0.18	1.84E-3
	60	5	2.30	2.20	7.21E-3	5	2.29	2.20	7.07E-3

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος ΙΙ, $tol = 10^{-10}$ $h = 100$ $l_{fil} = 300$ $ILUT + RCM$	15	3	0.07	0.035	8.18E-4	3	0.03	0.032	8.49E-4
	30	5	0.52	0.495	3.17E-3	5	0.515	0.48	3.1E-3
	60	13	7.23	6.74	7.34E-3	15	6.78	6.22	0.012

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος ΙΙΙ, $tol = 10^{-10}$ $h = 100$ $l_{fil} = 300$ $ILUT + RCM$	15	4	0.018	0.016	4.32E-4	5	0.017	0.015	4.3E-4
	30	4	0.20	0.19	1.62E-3	5	0.20	0.19	1.62E-3
	60	5	3.027	2.92	6.23	5	3.026	2.92	6.35E-3

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος ΙV, $tol = 10^{-10}$ $h = 100$ $l_{fil} = 300$ $ILUT + RCM$	15	4	0.01	0.015	5.32E-4	5	0.01	0.013	5.25E-4
	30	5	0.17	0.15	1.890E-3	5	0.174	0.15	1.83E-3
	60	6	2.33	2.22	4.55E-3	7	2.38	2.25	7.32E-3

Πίνακας 5.15: $ILUT + RCM$ για βάθος $h = 100$.



Σχήμα 5.6: Μορφή τοπογραφίας

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος I, $tol = 10^{-5}$	15	103	0.05	0.018	5.02E-4	99	0.04	0.01	5.00E-4
$h = 100$	30	-	-	-	-	-	-	-	-
$lfil = 300$	60	-	-	-	-	-	-	-	-
$ILUT + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος II, $tol = 10^{-5}$	15	-	-	-	-	-	-	-	-
$h = 100$	30	-	-	-	-	-	-	-	-
$lfil = 300$	60	-	-	-	-	-	-	-	-
$ILUT + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος III, $tol = 10^{-5}$	15	-	-	-	-	-	-	-	-
$h = 100$	30	-	-	-	-	-	-	-	-
$lfil = 300$	60	-	-	-	-	-	-	-	-
$ILUT + RCM$									

Περιγραφή	N_x	$Gmres$				$Bicgstab$			
		it	T	TP	TR	it	T	TP	TR
Τύπος IV, $tol = 10^{-5}$	15	125	0.063	0.01	5.01E-4	91	0.05	0.01	4.99E-4
$h = 100$	30	-	-	-	-	-	-	-	-
$lfil = 300$	60	-	-	-	-	-	-	-	-
$ILUT + RCM$									

Πίνακας 5.16: $ILUT + RCM$ για βάθος $h = 100$.

Περιγραφή	<i>Methods</i>	<i>ILUT + RCM</i>				<i>ILUK + RCM</i>			
		<i>it</i>	<i>T</i>	<i>TP</i>	<i>TR</i>	<i>it</i>	<i>T</i>	<i>TP</i>	<i>TR</i>
Τύπος I $l_{fil} = 300$ $tol = 10^{-5}$ $h = 0.1$ $N_x = 180$	<i>Gmres</i>	3	1.67	1.50	0.029	2	2.75	2.61	0.029
	<i>Bicgstab</i>	3	1.63	1.47	0.026	3	2.87	2.68	0.03
	<i>Gmres</i>	3	1.24	1.11	0.021	2	2.31	2.11	0.024
	<i>Bicgstab</i>	3	1.24	1.11	0.025	3	2.21	2.06	0.021
	με τοπογραφία								

Περιγραφή	<i>Methods</i>	<i>ILUT + RCM</i>				<i>ILUK + RCM</i>			
		<i>it</i>	<i>T</i>	<i>TP</i>	<i>TR</i>	<i>it</i>	<i>T</i>	<i>TP</i>	<i>TR</i>
Τύπος I $l_{fil} = 300$ $tol = 10^{-5}$ $h = 10$ $N_x = 180$	<i>Gmres</i>	2	2.29	2.1	0.02	3	3.62	3.38	0.029
	<i>Bicgstab</i>	3	2.32	2.13	0.030	3	2.88	2.70	0.025
	<i>Gmres</i>	9	2.39	1.95	0.025	3	2.79	2.61	0.025
	<i>Bicgstab</i>	9	2.37	1.95	0.025	3	2.79	2.61	0.025
	με τοπογραφία								

Πίνακας 5.17: Επαναλήψεις και χρόνος για πίνακες τύπου I με βάθος $h = 0.1, h = 1$ με $N_x = 180$ χωρίς και με τοπογραφία.

Περιγραφή	<i>Methods</i>	<i>ILUT + RCM</i>				<i>ILUK + RCM</i>			
		<i>it</i>	<i>T</i>	<i>TP</i>	<i>TR</i>	<i>it</i>	<i>T</i>	<i>TP</i>	<i>TR</i>
Τύπος I $l_{fil} = 300$ $tol = 10^{-10}$ $h = 100$ $N_x = 180$	<i>Gmres</i>	4	2.23	2.01	0.025	3	2.79	2.61	0.025
	<i>Bicgstab</i>	5	2.27	2.01	0.025	3	2.79	2.61	0.025
	χωρίς τοπογρ.								

Πίνακας 5.18: Επαναλήψεις και χρόνος για πίνακες τύπου I με βάθος $h = 100$ με $N_x = 180$ χωρίς τοπογραφία με $tol = 10^{-10}$.

Κεφάλαιο 6

Συμπεράσματα

Στη παρούσα μεταπτυχιακή εργασία παρουσιάστηκαν τα αποτελέσματα της συμπεριφοράς από την αριθμητική επίλυση αραιών γραμμικών συστημάτων που προκύπτουν με χρήση πεπερασμένων όγκων από διακριτοποίηση εξισώσεων τύπου Boussinesq.

Αρχικά, μελετήθηκαν οι αραιές μορφές αποθήκευσης πινάκων όπως είναι η COO, CSR, MSR και CSC. Το πρόγραμμα που κατασκευάστηκε σε γλώσσα προγραμματισμού Fortran και έχοντας χρησιμοποιήσει υποπρογράμματα διαδικασιών από τη βιβλιοθήκη γραμμικής άλγεβρας SPARSKIT, η οποία δέχεται ένα πίνακα σε δομή αποθήκευσης μορφής COO και τον μετατρέπει σε CSR μορφή. Από την ίδια βιβλιοθήκη χρησιμοποιήθηκαν υποπρογράμματα διάφορων επαναληπτικών μεθόδων, διαδικασιών αναδιάταξης και προρρύθμισης.

Ο κώδικας, έχει τη δυνατότητα επιλογής μεταξύ επαναληπτικών μεθόδων που ανήκουν στον υπόχωρο Krylov (Gmres, BiCGSTAB, TFQMR, CG), των διαδικασιών αναδιάταξης CMK,RCM και των προρρυθμιστών ILUO, ILUT, ILUK. Έπειτα, έγινε συστηματική αριθμητική μελέτη για πλέγματα των τεσσάρων τύπων που προέρχονται από τριγωνισμούς που χρησιμοποιήθηκαν (Distorted, Equilateral, Orthogonal, Orthogonal τύπου II) και των ανάλογων πινάκων. Σε κάθε πλέγμα έγινε συνεπή πύκνωση για τη μελέτη της συμπεριφοράς του αποτελέσματος, όσο αφορά τη σύγκλιση, το χρόνο και το υπολογιστικό κόστος. Παράλληλα, παρουσιάζονται στο κεφάλαιο 5 υπολογιστικά πειράματα, στα οποία έχουν μεταβληθεί οι φυσικοί παράμετροι του προβλήματος, όπως είναι η μέση στάθμη

και η διάσταση του υπολογιστικού χωρίου.

Στη μελέτη των τεσσάρων τύπων πλεγμάτων έγινε συνεπή πύκνωση $N_x = 15, 30, 60$ για μέση στάθμη μεγέθους $h = 0.1m, h = 10m, h = 100m$ και δοκιμάστηκαν οι πιθανοί συνδυασμοί επαναληπτικής μεθόδου με αναδιάταξη και προρρυθμηση. Επίσης, μεταβλήθηκε το κριτήριο απόρριψης στοιχείων και το επίπεδο γεμίσματος για να επιτευχθεί γρηγορότερη σύγκλιση σε χρόνο και επαναλήψεις. Δοκιμάστηκαν πίνακες που έχουν παραχθεί από περιπτώσεις χωρίς τοπογραφία και με τοπογραφία μορφής του σχήματος 5.6.

Συμπερασματικά, από τα πειράματα για τις επαναληπτικές μεθόδους προκύπτει ότι οι δύο γρηγορότερες σε σύγκλιση, αριθμό επαναλήψεων και χρόνου για την επίλυση του γραμμικού αραιού συστήματος είναι η Gmres και η BiCGSTAB. Αν γίνει συνδυασμός της επαναληπτικής μεθόδου και της διαδικασίας προρρυθμησης παρατηρούμε ότι ο ILUT προρρυθμιστής είναι καλύτερος σε θέμα χρόνου και επαναλήψεων σε μικρότερου μεγέθους μέση στάθμη. Σε αυτή τη περίπτωση ο ILUk είναι χρονικά λίγο ακριβότερος. Στην περίπτωση όμως που η μέση στάθμη αυξηθεί παρατηρούμε ότι η ILUT χάνει την αποτελεσματικότητά της γιατί σε πολλές περιπτώσεις δεν επιτυγχάνεται σύγκλιση. Τη τελευταία παρατήρηση για την αναποτελεσματικότητά της ILUT σε μεγάλες τιμές της μέσης στάθμης, διατύπωσαν στις εργασίες τους [14] που έλυσαν ένα γραμμικό σύστημα που προκύπτουν από τη χρήση της μεθόδου των πεπερασμένων διαφορών.

Στα επόμενα πειράματα μαζί με την προρρυθμηση έγινε και αναδιάταξη στον πίνακα. Μεταξύ των CMK και RCM, η δεύτερη ήταν αποδοτικότερη, όπως ήταν αναμενόμενο, αφού έχει λιγότερα γεμίσματα κατά την εκτέλεση του αλγορίθμου. Αναλυτικότερα, ο καλύτερος συνδυασμός για την επίλυση του γραμμικού συστήματος μελετώντας χρόνο και επαναλήψεις είναι :

- Για συστήματα που παράγονται για προβλήματα μικρής και μεσαίας μέσης στάθμης η GMRES ή η BiCGSTAB με προρρυθμιστή ILUT και αναδιάταξη RCM είναι ο καλύτερος συνδυασμός. Οι συγκεκριμένες επαναληπτικές μέθοδοι διαφέρουν

στο γεγονός ότι η Gmres χρειάζεται λιγότερες επαναλήψεις, ενώ η BiCGSTAB, χρειάζεται λιγότερο χρόνο.

- Για μεγαλύτερες τιμές της μέσης στάθμης οι επαναληπτικές μέθοδοι σε συνδυασμό με ILUT με $tol = 10^{-5}$ δε συγκλίνουν, οπότε ο χρήστης θα πρέπει να μεταβάλλει το tol σε 10^{-10} για να έχουμε αποτελέσματα ανάλογα με τη χρήση του προρρυθμιστή ILUK.

Όμοια αποτελέσματα έχουμε και για γραμμικά συστήματα που έχουν παραχθεί χωρίς και με τοπογραφία με μεγάλο N_x . Η GMRES με αναδιάταξη RCM και προρρύθμιση ILUT έχει γρηγορότερη σύγκλιση, καθώς και μικρότερο αριθμό επαναλήψεων στα μέσης στάθμης μεγέθη $h = 0.1m, h = 10m$. Για μέσο βάθος $h = 100m$ ο πίνακας που προκύπτει χωρίς τοπογραφία δε συγκλίνει με τον ILUT προρρυθμιστή και $tol = 10^{-5}$. Εάν αλλάξουμε το $tol = 10^{-10}$, τότε θα συγκλίνει η μέθοδος, με μικρές διαφορές από την ILUK, η οποία συγκλίνει χωρίς να χρειάζεται να μεταβληθούν καθόλου οι παράμετροι της.

Βιβλιογραφία

- [1] <http://www-users.cs.umn.edu/saad/software/ITSOL/>.
- [2] <http://www-users.cs.umn.edu/saad/software/SPARSKIT/>.
- [3] Barth, T. J. Aspects of Unstructured grids and finite volume solvers for the Euler and Navier-Stokes equations. In Special Course on Unstructured Grid Methods for advection Dominated Flows, AGARD report 787, 1992.
- [4] Barth, T. J. A 3-D upwind Euler solver for unstructured meshes. *AIAA*, CP, paper 91-1548.
- [5] Barth, T. J, Ohlberger, M., Stein, E. and de Borst, R. and Huges, T.R. Finite volume methods: foundation and analysis. John Wiley and Sons Ltd., 2004. Encyclopedia of Computational Mechanics.
- [6] Belem Para. Reordering effects on preconditioned Krylov methods in AMR solutions of flow and transport. Proc. of the XXVII Iberian Latin-American Congress on Computational Methods in Engineering CILAMCE 2006 Brazilian Assoc. for Comp. Mechanics & Latin American Assoc. of Comp. Methods in Engineering Belem, Para, Brazil.
- [7] Benzi M. Preconditioning Techniques for Large Linear Systems: A survey. *Mathematics and Computer Science Department, Emory University, Atlanta, Georgia* 30322., 2002, 418-476.

- [8] Nikolos, I. K. και Delis, A. I. An unstructured node-centered finite volume scheme for shallow water flows with wet/dry fronts over complex topography. *Comput. Methods Appl. Mech. Engrg.*, 198, 3723-3750, 2009.
- [9] Engsing-Karup. A.P., Hesthaven J.S. Warburton, T. DG-FEM solution for nonlinear wave structure interaction using Boussinesq-type equations. *Coastal. Engineering*, 55,197, 2008.
- [10] Erduran, K.S. Further application of hybrid solution to another form of Boussinesq equations and comparisons. *International Journal for Numerical Methods in Fluids*, 53,827-849, 2007.
- [11] Eskilsson C., Sherwin, J. Spectral/hp discontinuous Galerkin methods for modelling 2D boussinesq equations. *Journal of Computational Physics*, 210,566, 2006.
- [12] Freund, R. A transpose-free quasi minimal residual algorithm for non Hermitian linear systems. *Siam J. Sci. Comput.*, 14,470, 1993.
- [13] Freund, R. Nachtigal, N.M. QMR:A quasi-minimal residual method for non Hermitian linear systems. *Numer. Math.*, 60,351, 1991.
- [14] Fuhrman, R. D. Numerical solutions of fully non-linear and highly dispersive Boussinesq equations in two horizontal dimensions. *International Journal for numerical methods in Fluids.*, 44,231-255, 2003.
- [15] Gustafsson, I. and Lindskog, G. A preconditioning technique based on element matrix factorizations. *Comp. Methods ppl. Mech. Eng.*, 55,201-220, 1986.
- [16] Henk A. van der Vorst. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of non-symmetric linear systems. *Siam J.Sci. Stat.Comput.*, 12,631, 1992.

- [17] Hestenes, M R. Stiefel, E.L. Methods of conjugate gradients for solving linear systems. *J.Res. Natl.Bur.Stand*, 49,409, 1952.
- [18] Delis, A. I. και Nikolos, I. K και Kazolea, M. Performance and comparison of cell-centered and node-centered unstructured finite volume discretizations for shallow water free surface flows. *Archives of Computational Methods in Engineering*, 18, 57, 2011.
- [19] Kazolea M. and Delis, A. I. A well-balanced shock-capturing hybrid finite volume-finite difference numerical scheme for extended 1D Boussinesq models. *Applied Numerical Mathematics*., In press.
- [20] Lanczos, C. Solutions of systems of linear equations by minimized iteration. *J.Res. Natl.Bur.Stand*, 82,47, 1952.
- [21] Liggett, J. A. *Fluid Mechanics*. Inc McGrawHill, 1994.
- [22] Madsen, P. A. and Sørensen, O. R. A new form of the Boussinsq equations with improved linear dispersion characteristics. Part 2: A slowing varying bathymetry. *Coast. Eng.*, 18,183-204, 1992.
- [23] Nwogu, O. An alternative form of the Boussinesq equations for nearshore wave propagation . *Journal of Waterway, Port, Coastal, and Ocean Engineering* , 119, 618-638, 1994.
- [24] Paige, C.C, Saunders, M.A. Solutions of sparse indefinite systems of linear equations. *Siam J. Nymmer. Anal.*, 12,617, 1975.
- [25] Peregrine, D.H. Long waves on a beach. *Journal of Fluid Mechanics* 27, 1967, 27,815.

- [26] Reid, J.K. On the methods of conjugate gradients for the solution of large sparse systems of linear equations, in *Large Sparse Sets of Linear Equations. edited by J.K.Reid*, 231, 1971.
- [27] Roeber, V. and Cheung, K. F. and Kobayashi, M. H. Shock-capturing Boussinesq-type model for nearshore wave processes. *Coast. Engrg.*, 57, 407, 2010.
- [28] Roeber, V., Cheung, K. Boussinesq-type model for energetic breaking waves in fringing reef environment. *δασιαβλ. Ενγινεερινγ*, 57,407-423.
- [29] Roe, P.L. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*, 43,357, 1981.
- [30] Saad Y. *Iterative Methods for Sparse Linear Systems*. siam, 2003.
- [31] Saad Y. Henk A. van de Vorst. *Iterative Solution of Linear Systems in the 20-th Century*.
- [32] Saad Y. Schultz, M. H. GMRES:A generalized minimal residual algorithm for solving nonsymmetric linear systems. *Siam J.Sci. Stat.Comput.*, 7,856, 1986.
- [33] Kazolea, και M. Delis, A. I. και Nikolos, I. K και Synolakis, C.E. An unstructured finite volume numerical scheme 2D Boussinesq-type equations. . *Coastal engineering*, 69, 42-66, 2011.
- [34] Tonelli, M. Petti, M. Hubrid finite-volume-finite-difference scheme for 2D extended Boussinesq equations in the surf zone. *Coastal. Engineering*, 37,567, 2010.
- [35] Walkey, M., Berzins, M. A finite element method for the two-dimensional extended Boussinesq equations. *International Journal for Numerical Methods in Fluids*, 39,865, 2002.
- [36] Ακριβής Γ. Δ. και Δουγαλής Β. Α. *Εισαγωγή στην αριθμητική ανάλυση*. Πανεπιστημιακές εκδόσεις Κρήτης, 2004.