



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Εφαρμογή Μεθόδων Deep Equilibrium σε
Προβλήματα Αντίστροφης Απεικόνισης: Η
Περίπτωση της Ανακατασκευής Εικόνων MRI

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΚΑΤΣΑΡΟΥ ΑΝΔΡΕΑ

Επιβλέποντες: Κωνσταντίνος Μπερμπερίδης
Γκίλλας Αλέξανδρος

Πάτρα, Οκτώβρης 2025



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Εφαρμογή Μεθόδων Deep Equilibrium σε
Προβλήματα Αντίστροφης Απεικόνισης: Η
Περίπτωση της Ανακατασκευής Εικόνων MRI

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΚΑΤΣΑΡΟΥ ΑΝΔΡΕΑ

Επιβλέποντες: Κωνσταντίνος Μπερμπερίδης
Γκίλλας Αλέξανδρος

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 15η Οκτωβρίου 2025.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Κωνσταντίνος Μπερμπερίδης
Καθηγητής

.....
Ψαράκης Εμμανουήλ
Καθηγητής

.....
Κοσμόπουλος Δημήτρης
Καθηγητής

Πάτρα, Οκτώβρης 2025



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Copyright ©—All rights reserved Κατσαρός Ανδρέας, 2025.
Με την επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέχρινε.

Τπεύθυνη Δήλωση

Βεβαιώνω ότι είμαι συγγραφέας αυτής της πτυχιακής εργασίας, και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην πτυχιακή εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης, βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του προγράμματος σπουδών του τμήματος Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληρογραφικής του Πανεπιστημίου Πατρών.

(Τπογραφή)

.....

Κατσαρός Ανδρέας

Περίληψη

Στην παρούσα εργασία αντιμετωπίζουμε την ανακατασκευή εικόνων μαγνητικής τομογραφίας MRI ως αντίστροφο πρόβλημα μορφής $y = Ax + \epsilon$ υπό υποδειγματοληψία στον χώρο k (k-space), με στόχο τη μείωση χρόνου σάρωσης χωρίς απώλεια διαγνωστικής πληροφορίας. Αναπτύσσουμε ένα ολοκληρωμένο σύστημα ανακατασκευής από τα ακατέργαστα δεδομένα γονάτου προερχόμενα από μονό πηνίο (single-coil) του σύνολου fastMRI. Η ανάλυση καλύπτει τη ροή από την πολυτομική φόρτωση δεδομένων έως και την αξιολόγηση. Η αξιολόγηση γίνεται υπό καρτεσιανές μάσκες μεταβλητής πυκνότητας με επιταχύνσεις $4\times$ και $8\times$. Χρησιμοποιούμε αρχικοπόίηση με μηδενική πλήρωση (zero-filled) και μετράμε την απόδοση με Λόγο Σήματος προς Θόρυβο (PSNR) και Δείκτης Δομικής Ομοιότητας (SSIM).

Ο πυρήνας της μεθόδου είναι ένα σχήμα Βαθιάς Ισορροπίας με Μέθοδο Εγγύς Κλίσης (Deep Equilibrium Proximal Gradient Descent, DE-Prox), το οποίο εναλλάσσει τον όρο συνέπειας δεδομένων στον χώρο k (k-space) με έναν προεκπαιδευμένο αποθορυβοποιητή DnCNN με υπολειμματική αρχιτεκτονική (residual) και φασματική κανονικοποίηση (spectral normalization). Ένα μοντέλο βαθιάς Ισορροπίας αξιοποιεί ένα έμμεσο (implicit) επίπεδο, του οποίου η έξοδος ορίζεται ως το σταθερό σημείο (fixed point) μιας παραμετροποιημένης μη γραμμικής συνάρτησης. Για την αναπαράσταση χρησιμοποιούνται Δίκτυα με Κοινά Βάρη (weight-tied networks), όπου το ίδιο σύνολο παραμέτρων επαναχρησιμοποιείται σε όλες τις επαναληπτικές εφαρμογές του μετασχηματισμού. Επιπλέον, διερευνούμε μια νέα συνάρτηση κόστους για μιγαδικές εικόνες, την κάθετη απώλεια (perpendicular loss, \perp -loss), η οποία αποσυνθέτει ρητά το σφάλμα σε γωνιακό (φάση) και μετρικό (μέτρο), αντιμετωπίζοντας θεωρητικά τη μεροληψία συρρίκνωσης του μέσου τετραγωνικού σφάλματος MSE σε περιπτώσεις φασικής ασυμφωνίας.

Συνολικά, η εργασία καταδεικνύει ότι ένα προσεκτικά σταθεροποιημένο δίκτυο βαθιάς Ισορροπίας για MRI, προσφέρει σταθερή, χρήσιμα οπτική βελτίωση, ενώ η \perp -loss αποτελεί θέμα προς συζήτηση, αιτιολογημένη κατεύθυνση για μάθηση στο μιγαδικό πεδίο, με χώρο για βελτιστοποίηση.

Λέξεις Κλειδιά

Μαγνητική Τομογραφία (MRI), Ανακατασκευή Εικόνας, Αντίστροφα Προβλήματα, Deep Equilibrium (DEQ), Optimization, Data-Consistency,, DnCNN, Agorithmic Unrolling, Implicit Differentiation, Spectral Normalization, Forward Solver, Backward Solver, k-space, slices, singlecoil, fastMRI, Masking, PSNR, SSIM, Perpendicular Loss (\perp -loss).

Abstract

In this work, we cast MRI image reconstruction as an inverse problem of the form $y = Ax + \varepsilon$ under undersampling in k-space, aiming to reduce scan time without loss of diagnostic information. We develop an end-to-end reconstruction system from raw single-coil knee data from the fastMRI dataset. The analysis covers the pipeline from multi-slice data loading through evaluation. Evaluation is conducted with Cartesian variable-density masks at acceleration factors of $4\times$ and $8\times$. We use zero-filled initialization and assess performance with Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

The core of the method is a Deep Equilibrium Proximal Gradient Descent (DE-Prox) scheme that alternates a k-space data-consistency step with a pretrained DnCNN denoiser featuring a residual architecture and spectral normalization. A deep equilibrium model employs an implicit layer whose output is defined as the fixed point of a parameterized nonlinear function. For the representation, we use weight-tied networks, in which the same set of parameters is reused across all iterative applications of the transformation. In addition, we investigate a new loss function for complex-valued images, the perpendicular loss (\perp -loss), which explicitly decomposes the error into angular (phase) and radial (magnitude) components, theoretically addressing the shrinkage bias of mean squared error (MSE) in cases of phase mismatch.

Overall, the work demonstrates that a carefully stabilized deep equilibrium network for MRI provides consistent, useful visual improvement, while the \perp -loss remains a topic for discussion—a justified direction for learning in the complex domain, with room for optimization.

Keywords

Magnetic Resonance Imaging (MRI), Image Reconstruction, Inverse Problems, Deep Equilibrium (DEQ), Optimization, Data-Consistency, Regularizer, DnCNN, Algorithmic Unrolling, Implicit Differentiation, Spectral Normalization, Forward Solver, Backward Solver, Pre-processing, k-space, Slices, Single-coil, fastMRI, Masking, PSNR, SSIM, Perpendicular Loss (\perp -loss).

στους γονείς μου

Ευχαριστίες

Θα ήθελα καταρχήν να ευχαριστήσω τον καθηγητή κ.Μπερμπερίδη για την εποπτεία και καθοδήγηση αυτής της διπλωματικής εργασίας και για την ευκαιρία που μου έδωσε να την εκπονήσω στο εργαστήριο Σημάτων και Τηλεπικοινωνιών. Ιδιαίτερες ευχαριστίες οφείλω στον Δρ. Αλέξανδρο Γκίλλα για την καθοριστική συμβολή, την υποστήριξη και την εξαιρετική συνεργασία που είχαμε.Θα ήθελα, επίσης, να ευχαριστήσω θερμά τους γονείς μου για τη συνεχή συμπαράσταση και τη στήριξή τους όλα αυτά τα χρόνια, καθώς και την αδελφή μου Ελευθερία και τους φίλους μου για την ενθάρρυνση και την παρουσία τους.Τέλος, θα ήθελα να ευχαριστήσω ιδιαίτέρως τους φίλους και συμφοιτητές μου, Νίκο και Γρηγόρη, για τη βοήθειά τους στις δοκιμές, την ανάλυση των δεδομένων και τη γόνιμη ανταλλαγή απόψεων κατά την πορεία της εργασίας.

Περιεχόμενα

Περίληψη	i
Abstract	iii
Ευχαριστίες	vii
Περιεχόμενα	x
1 Εισαγωγή	2
2 Βασικές Αρχές Μαγνητικής Τομογραφίας (MRI) και Διατύπωση του Αντίστροφου Προβλήματος	5
2.1 Η Μαγνητική Τομογραφία και η Κλινική της Σημασία	5
2.1.1 Φυσικές αρχές : Βασικές Έννοιες και Μαγνητισμός Συντονισμός	6
2.1.2 MRI και χώρος k (k-space): Από την Απόκτηση Δεδομένων στην Εικόνα	10
2.2 Βασικές έννοιες των αντίστροφων προβλημάτων	12
2.2.1 Σύνδεση με το πρόβλημα της ανακατασκευής MRI	15
3 Σχετική Έρευνα στην Ανακατασκευή MRI	20
3.1 Χρήσιμο Θεωρητικό Υπόβαθρο για Μεθόδους Αραιής Αναπαράστασης	20
3.2 Προσεγγίσεις με μάθηση	23
3.2.1 Μέθοδος Ξετυλίγματος Αλγορίθμου και δίκτυα για ανακατασκευή MRI	31
3.3 Η ιδέα των μοντέλων Ισορροπίας DEQ	34
3.4 Οι λειτουργίες του μοντέλου Ισορροπίας :	37
3.4.1 Σταθερό σημείο μοντέλου Ισορροπίας με επαναληπτικό χάρτη	37
3.4.2 DEQ οικογένειες επιλυτών και θεωρητικά κριτήρια	40
4 Το Σύνολο Δεδομένων	43
4.1 Ο χώρος k (k-space)	43
4.2 Μάσκες υποδειγματοληψίας και ρυθμίσεις επιτάχυνσης (π.χ. 4× / 8×)	46
4.3 Το σύνολο δεδομένων (fastMRI)	48
4.4 Μηχανισμός φόρτωσης και Προεπεξεργασία Δεδομένων	53

5 Μεθοδολογία: ανακατασκευή MRI με μέθοδο Βαθιάς Ισορροπίας με Μέθοδο Εγγύς Κλίσης	57
5.1 Βασικές διατυπώσεις μοντέλου Ισορροπίας:	57
5.2 Προσεγγιση και Αλγορίθμική Λογική μεθόδου:	59
5.3 Χάρτης βαθιάς Ισορροπίας με μέθοδο εγγύς κλίσης	63
6 Αποθρυβοποιητής (Denoiser) και Αξιολόγηση Μοντέλου Ισορροπίας	66
6.1 Προεκπαιδευμένος Αποθρυβοποιητής	66
6.1.1 Προεκπαίδευση με λευκό Γκαουσιανό θόρυβο	67
6.1.2 Αρχιτεκτονική DnCNN	68
6.1.3 Αξιολόγηση του αποθρυβοποιητή	68
6.2 Υλοποίηση του μοντέλου Ισορροπίας	72
7 Συνάρτηση Απώλειας \perp loss	77
7.1 Εισαγωγή	77
7.2 Η συνάρτηση μέσου τετραγωνικού σφάλματος και η μεροληψία του	79
7.3 Περιγραφή της συνάρτησης απώλειας \perp loss	79
7.3.1 Η γωνιακή συνιστώσα της προτεινόμενης απώλειας	79
7.3.2 Η λεία συνέχεια για μεγάλες γωνίες	80
7.3.3 Μετρική συνιστώσα που αποκαθιστά το σωστό μήκος	80
7.3.4 Η πλήρης απώλεια και η αμεροληψία ως προς το μέτρο	80
7.3.5 Αριθμητική σταθερότητα και ειδικές περιπτώσεις	81
7.4 Ενσωματωση στην πράξη	81
7.4.1 Σχολιασμός αποτελεσμάτων	83
8 Συμπέρασμα και Μελλοντική Εργασία	89
8.1 Συμπεράσματα και κατευθύνσεις	89
8.2 Περίληψη συνεισφορών και αποτελεσμάτων	89
Α' Παράρτημα	97
Α'.1 Εξηγηση	97

Κεφάλαιο 1

Εισαγωγή

Η μαγνητική τομογραφία (MRI) προσφέρει υψηλή αντίθεση μαλακών ιστών και ευελιξία πρωτοχόλλων χωρίς ιονίζουσα ακτινοβολία[1][2]. Το τίμημα είναι ο χρόνος σάρωσης: για να απεικονίσουμε λεπτομέρεια πρέπει να δειγματοληπτήσουμε πυκνά τον χώρο k [45] , κάτι που αυξάνει τη διάρκεια, το κόστος και την ευαισθησία σε κίνηση. Η κατευθυνόμενη υποδειγματοληψία (undersampling) συντομεύει τον χρόνο, αλλά μετατρέπει την ανακατασκευή σε κακώς τεθειμένο (ill posed) αντίστροφο πρόβλημα[22][24]. Η τυπική διατύπωση είναι $y = Ax + \varepsilon$ με A τον τελεστή μέτρησης (FFT + μάσκα), και λύνεται ως κανονικοποιημένη βελτιστοποίηση

$$\min_x \frac{1}{2} \|Ax - y\|_2^2 + R(x).$$

Τα τελευταία χρόνια, εκτός από προσεγγίσεις στην επίλυση του συγκεκριμένου προβλήματος όπως compressed sensing[28], ISTA/FISTA, ADMM, αναδείχθηκε η μέθοδος ξετυλίγματος αλγορίθμου (algorithmic unrolling)[40] μια μεθοδολογία που μετασχηματίζει επαναληπτικούς αλγορίθμους βελτιστοποίησης σε βαθιά νευρωνικά δίκτυα, αντιστοιχίζοντας κάθε επανάληψη σε ένα αντίστοιχο επίπεδο του δικτύου. Συγχρόνως με τα ξεδιπλωμένα δίκτυα, τα Μοντέλα Βαθιάς Ισορροπίας (DEQs)[42] έχουν αναδειχθεί ως συμπληρωματική προσέγγιση στη βαθιά μάθηση, εισάγοντας μια διαφορετική οπτική στον σχεδιασμό νευρωνικών δικτύων. Τα μοντέλα βαθιάς Ισορροπίας DEQs συνιστούν μια τεχνική, όπου, αντί της στοίβαξης πεπερασμένου αριθμού επιπέδων, το μοντέλο υιοθετεί ένα και μόνο implicit επίπεδο, το οποίο προσδιορίζει την έξιδο του ως το Σταθερό Σημείο [42] μιας παραμετροποιημένης μη γραμμικής συνάρτησης. Για προβλήματα ανακατασκευής εικόνας, αναλύεται ο πρακτικός χάρτης Βαθιάς Ισορροπίας [41] με Μέθοδο Εγγύς Κλίσης(DE-Prox)[22]: εναλλάσσει τον όρο συνέπειας δεδομένων με ένα δίκτυο αποθορυβοποίησης (DnCNN) (R_θ) στο χώρο εικόνας. Επιπρόσθετα, διερευνάται μια νέα συνάρτηση κόστους, η κάθετη απώλεια(\perp loss)[53] σε σύνθετο πεδίο που αποσκοπεί στη ρητή ευθυγράμμιση φάσης και στη σωστή αποκατάσταση μέτρου, ως εναλλακτική του μέσου τετραγωνικού σφάλματος (MSE).

1.1 Σημασία του Προβλήματος

1. Κλινική ανάγκη για ταχύτερη μαγνητική τομογραφία (MRI). Μικρότερος χρόνος εξέτασης σημαίνει χαμηλότερο κόστος/φόρτο μηχανήματος και λιγότερα λάθη λόγω κινήσε-

ων(motion artefacts), με άμεσο όφελος για τη διαγνωστική αξιοπιστία[41]. Ωστόσο, η υποδειγματοληψία στον χώρο k (k-space) μειώνει την πληροφορία και καθιστά τον A μη πλήρους βαθμίδας τελεστή, άρα το πρόβλημα ασταθές[18][27].

2. Ανάγκη για μεθόδους με φυσική ερμηνεία και αποδοτικότητα. Τα ξεδιπλωμένα δίκτυα ενσωματώνουν το προωθητικό (forward) μοντέλο και επιτυγχάνουν καλές ανακατασκευές, όμως η μνήμη τους κλιμακώνεται με τα βήματα[40][41]. Τα μοντέλα βαθιάς Ισορροπίας διατηρούν τον όρο συνέπειας των δεδομένων (data-consistency), του φυσικού μοντέλου και εκπαιδεύουν απευθείας το σταθερό σημείο, με σταθερό κόστος μνήμης και σαφή μαθηματική θεμελίωση[41] (ιδιότητες σύγκλισης, Lipschitz έλεγχος του R_θ , κατάλληλο βήμα η).
3. Σύνθετη φύση του σήματος MRI. Τα δεδομένα είναι μιγαδικά: η φάση επηρεάζει την ευθυγράμμιση στον χώρο k και το μέτρο καθορίζει την ένταση. Το κλασικό μέσο τετραγωνικό σφάλμα, μπορεί να δημιουργήσει μεροληψία συρρίκνωσης του μέτρου όταν υπάρχει φασικό σφάλμα. Αυτό υποκινεί την έρευνα για συναρτήσεις κόστους [53] προσαρμοσμένες στα μιγαδικά δεδομένα, ώστε να τιμωρούν χωριστά την απόκλιση φάσης και μέτρου.

1.2 Στόχοι της Διπλωματικής Εργασίας

- A. Πλαίσιο και δεδομένα.** – Διατύπωση του MRI ως αντίστροφου προβλήματος ($(y = Ax + \varepsilon)$), ανάλυση χώρου k[45]/μασκών και προεπεξεργασία fastMRI (single coil knee). Υλοποίηση πορείας mask → IFFT (zero fill) → crop/normalization → metrics.
- B. Ανακατασκευή με Υλοποίηση Βαθιάς Ισορροπίας.** – Προεκπαίδευση denoiser σε λευκό Gaussian θόρυβο ($\sigma = 0.30/0.10/0.01$) για το στάδιο Ισορροπίας με συγχριμένη αρχιτεκτονική. Σύγκριση έναντι zero filled σε $4\times/8\times$ με PSNR/SSIM και συνοδευτικούς διαγνωστικούς ελέγχους.
- Γ. Νέα συνάρτηση κόστους (\perp loss).** – Θεωρητική και πρακτική διερεύνηση μιας συνάρτησης κόστους[53] που αποσυνδέει γωνιακό σφάλμα (φάση) από το σφάλμα μέτρου (μήκος). Στόχος είναι να καθιδηγεί τον αποθορυβοποιητή denoiser προς οπτικά πιο πιστή ευθυγράμμιση (άκρα, λεπτομέρειες), ακόμη κι αν οι κλασικές μετρικές δεν βελτιώνονται ομοιόμορφα. Συγκρίνεται ποσοτικά έναντι του μέσου τετραγωνικού σφάλματος, με ποιοτική αποτίμηση σε τομές γονάτου.
- Δ. Πλήρης, αναπαραγώγιμη σχεδίαση.** – Χρησιμοποιείται το σύνολο δεδομένων fastMRI με δεδομένα γονάτου μονού πηνίου και καρτεσιανές μάσκες μεταβλητής πυκνότητας για $4\times/8\times$.

Κεφάλαιο 2

Βασικές Αρχές Μαγνητικής Τομογραφίας (MRI) και Διατύπωση του Αντίστροφου Προβλήματος

Στο Κεφάλαιο 2 παρουσιάζονται οι βασικές αρχές της μαγνητικής τομογραφίας και το μαθηματικό πλαίσιο της ανακατασκευής εικόνας: Αρχικά (2.1) εξηγούμε τη διαδικασία της μαγνητικής τομογραφίας, το φαινόμενο του μαγνητικού συντονισμού, τους χρόνους χαλάρωσης T_1/T_2 , τους παραμέτρους ακολουθιών (TR/TE) και τα πηνία λήψης (coils). Στη συνέχεια (2.2) διατυπώνουμε τις βασικές έννοιες αντίστροφων προβλημάτων $y = Ax + \varepsilon$, αναλύοντας φαινόμενα όπως γιατί η υποδειγματοληψία και ο ύσροβος οδηγούν σε κακώς ορισμένα συστήματα. Τέλος δείχνουμε πώς αυτές οι ιδέες αφορούν και συνδέονται στη πράξη με το πρόβλημα ανακατασκευής MRI.

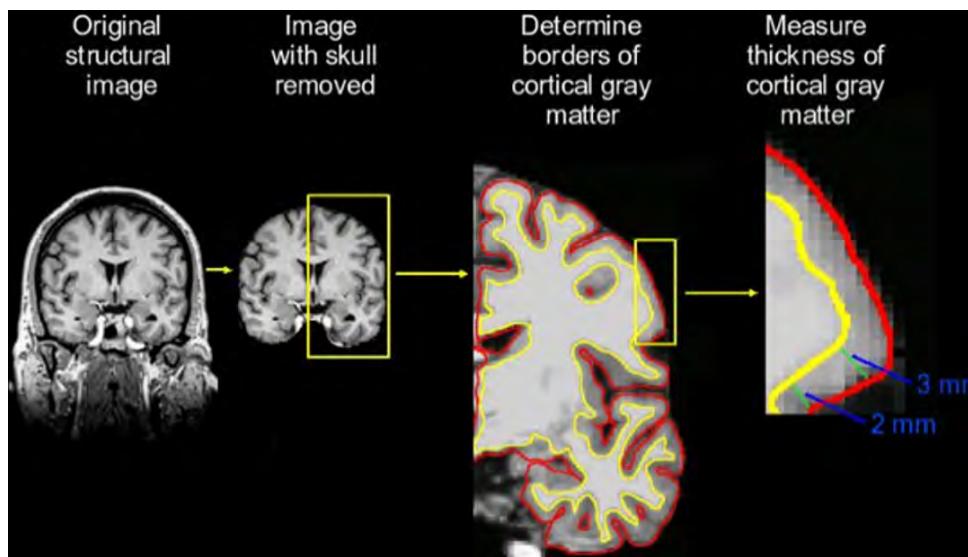
2.1 Η Μαγνητική Τομογραφία και η Κλινική της Σημασία

Η μαγνητική τομογραφία (MRI) είναι μια απεικονιστική τεχνική ιατρικής διάγνωσης που εκμεταλλεύεται τις μαγνητικές ιδιότητες των πυρήνων ατόμων, χυρίως των πρωτονίων του υδρογόνου, για να παραγάγει εικόνες υψηλής αντίθεσης, μη επεμβατικές, των εσωτερικών δομών του σώματος. Σε αντίθεση με τις ακτινογραφίες και την αξονική τομογραφία, που χρησιμοποιούν ιονίζουσες ακτινοβολίες, το MRI είναι μη ιονίζουσα μέθοδος [1]. Κλινικά, οι MRI εικόνες χρησιμοποιούνται για τη διάγνωση πλήθους παθολογιών: ανεύρεση και χαρακτηρισμός όγκων σε εγκέφαλο/σώμα, αξιολόγηση τραυματισμών αρθρώσεων, απεικόνιση καρδιακής λειτουργίας, απεικονίσεις του νωτιαίου μυελού και, επίσης, για λειτουργικές/δομικές μελέτες (π.χ. fMRI), όπου το MRI παρέχει πληροφορίες που είναι κρίσιμες για θεραπευτικές αποφάσεις και για τον προγραμματισμό χειρουργικών επεμβάσεων [2].

Το σύστημα της μαγνητικής τομογραφίας, δημιουργεί ένα ισχυρό, ομοιογενές πεδίο για

πόλωση των πυρήνων, εφαρμόζει ελεγχόμενες μεταβολές πεδίου (gradients) για χωρική κωδικοποίηση και χρησιμοποιεί παλμούς ραδιοσυχνοτήτων RF για διέγερση.[3] Το σήμα που λαμβάνεται από τα πηνία (coils) μετατρέπεται σε ψηφιακές μετρήσεις στον χώρο και από εκεί ανασυντίθεται σε εικόνα μέσω αντίστροφου μετασχηματισμού Fourier.

Η πρακτική αξία του MRI συνοψίζεται στην ικανότητα παραγωγής τρισδιάστατων τομών με εξαιρετική αντίθεση μεταξύ διαφορετικών μαλακών ιστών[3][4] κάτι που σε πολλές κλινικές περιπτώσεις αντικαθιστά ή συμπληρώνει CT και υπέρχο. Επιπλέον, επειδή ο ιατρός μπορεί να επιλέξει πρωτόκολλα (T1-weighted, T2-weighted, FLAIR, PD, κ.ά.) που τονίζουν διαφορετικές φυσιοχημικές ιδιότητες, το MRI είναι ιδιαίτερα ευέλικτο ως προς τη στοχευμένη απεικόνιση συγκεκριμένων ιστών ή παθολογιών.[5]

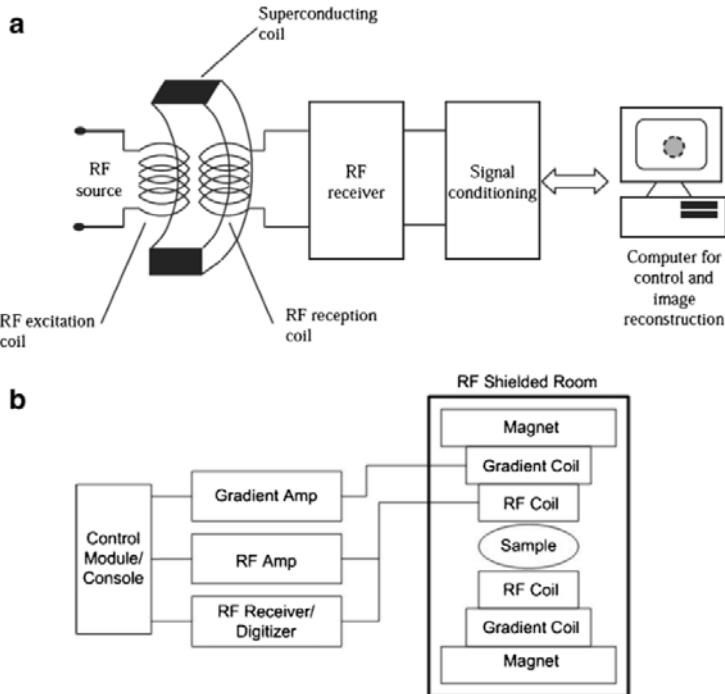


Σχήμα 2.1: Δομικές/ανατομικές εικόνες Μαγνητικής Τομογραφίας (MRI) της εγκεφαλικής ανατομίας. Τρέχουσα εφαρμογή: μέτρηση του πάχους του εγκεφαλικού φλοιού. Μεταβολές στο πάχος του φλοιού παρατηρούνται μετά από παθολογικές καταστάσεις ή έπειτα από επανειλημμένη έκθεση σε συγκεκριμένες εμπειρίες.[10]

2.1.1 Φυσικές αρχές : Βασικές Έννοιες και Μαγνητισμός Συντονισμός

Η μαγνητική τομογραφία στηρίζεται στο φαινόμενο του μαγνητικού συντονισμού, το οποίο προκύπτει από την κβαντική ιδιότητα spin των πυρήνων (χυρίως των πρωτονίων του υδρογόνου), και στην αλληλεπίδρασή τους με εξωτερικά μαγνητικά πεδία και ελεγχόμενους παλμούς ραδιοσυχνοτήτων (RF). Ιστορικά, η τεχνική προήλθε από την Πυρηνική Μαγνητική Αντήχηση (nuclear magnetic resonance, NMR)[4] [6].Το αρχικό «nuclear» στο ακρωνύμιο αφαιρέθηκε στις κλινικές εφαρμογές τη δεκαετία του 1980, καθώς προκαλούσε αβάσιμους φόβους περί πυρηνικής ακτινοβολίας. Σε πρακτικό επίπεδο, το MRI είναι μη ιονίζουσα μέθοδος και μπορεί να εφαρμοστεί επανειλημμένα χωρίς

τον ίδιο κίνδυνο που συνδέεται με την έκθεση σε ακτίνες X[6].

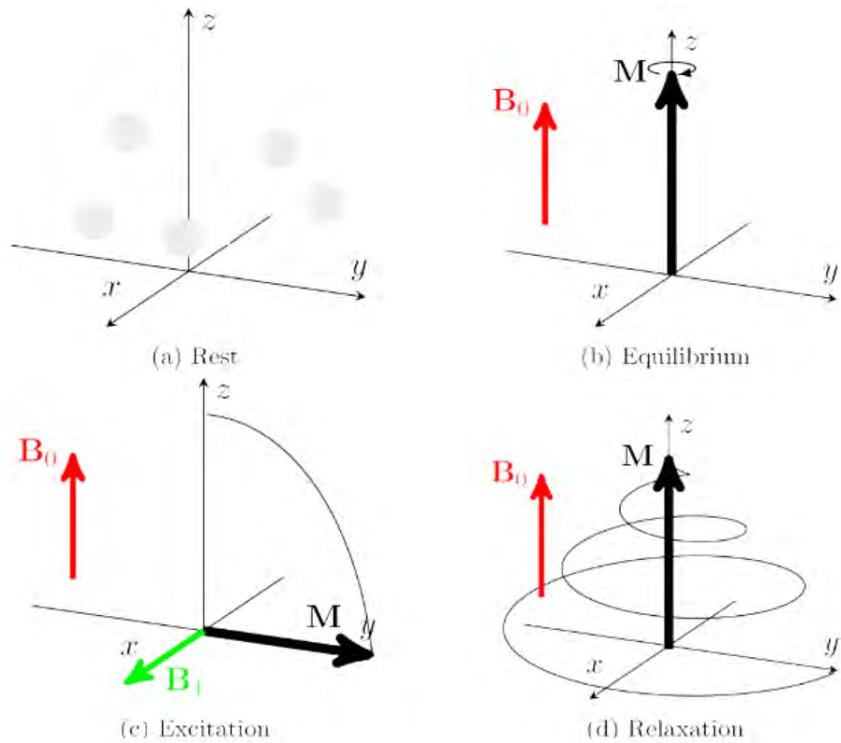


Σχήμα 2.2 Διάγραμμα βαθμίδων (block diagram) ενός τυπικού συστήματος Μαγνητικής Τομογραφίας – MRI (Abdullah, 2008). (β) Διάγραμμα βαθμίδων συστήματος MRI (Jin, 1989)[7]

Το θεμελιώδες υλικό που ανιχνεύει η MRI είναι το πρωτόνιο του υδρογόνου. Το πρωτόνιο έχει την κβαντική ιδιότητα spin, η οποία του προσδίδει μαγνητική ροπή μ.[5]. Στην παρουσία ενός ομογενούς εξωτερικού πεδίου B_0 , κάθε μαγνητικό δίπολο τείνει να ευθυγραμμιστεί με το πεδίο. Λόγω όμως της ιδιότητας της στροφορμής, αντί να ευθυγραμμιστεί άμεσα, παρουσιάζει κίνηση γύρω από το B_0 με συχνότητα Larmor ω_0 που ικανοποιεί την εξίσωση

$$\omega_0 = \gamma B_0,$$

όπου γ είναι ο γυρομαγνητικός λόγος (gyromagnetic ratio) του πρωτονίου[8] [9]. Η συλλογική επίδραση δισεκατομμυρίων πρωτονίων παράγει ένα μακροσκοπικό διάνυσμα μαγνήτισης \mathbf{M} που μπορεί να περιγραφεί κλασικά και να επιφέρει επαγόμενη τάση στα πηγιά λήψης όταν μεταβάλλεται στο χρόνο. Η λειτουργία του ραδιοσυχνικού παλμού (RF) είναι να μετατοπίζει τμήμα του μαγνητισμού M των πρωτονίων από τη διαμήκη συνιστώσα (κατά τον άξονα z) προς την εγκάρσια συνιστώσα. Η γωνία με την οποία ο ραδιοσυχνικός παλμός στρέφει τον μαγνητισμό M ονομάζεται γωνία κλίσης (flip angle). Από αυτό προκύπτει η εξάρτηση του σήματος από τη γωνία κλίσης: μικρή $\alpha \rightarrow$ μικρό σήμα· για $\alpha = 90^\circ$ επιτυγχάνεται μέγιστη εγκάρσια συνιστώσα.



Σχήμα 2.3 Απεικόνιση του πειράματος NMR. (a) Χωρίς εφαρμογή μαγνητικού πεδίου, τα spin είναι τυχαία προσανατολισμένα. (β) Όταν εφαρμοστεί στατικό μαγνητικό πεδίο B_0 κατά τον άξονα z . (γ) Η μαγνήτιση Ισορροπίας μετατοπίζεται προς το εγκάρσιο επίπεδο xy από την επίδραση του B_1 στη συχνότητα συντονισμού. (δ) Όταν παύσει η διέγερση RF, η μαγνήτιση χαλαρώνει προς την τιμή Ισορροπίας με συχνότητα μετάπτωσης που εξαρτάται από τις μαγνητικές ιδιότητες του εξεταζόμενου στοιχείου.[12]

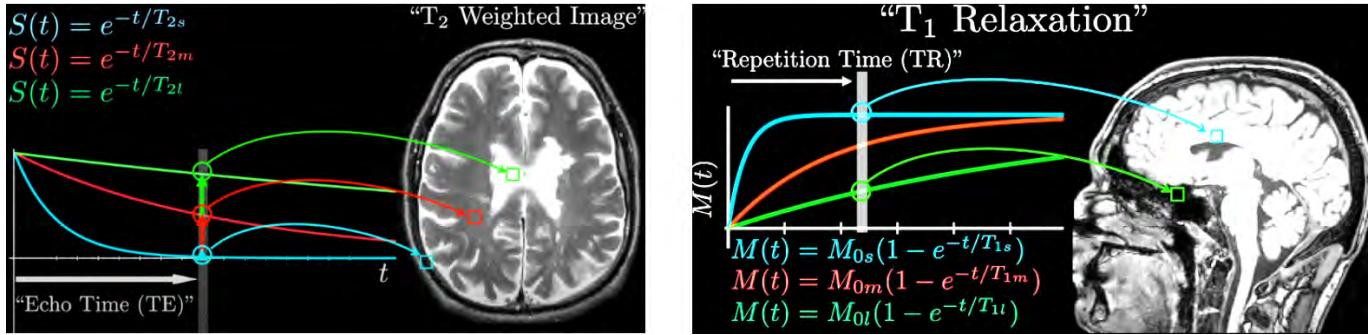
Η αύξηση της έντασης του κύριου πεδίου B_0 οδηγεί σε άνοδο τόσο της συχνότητας Larmor όσο και της έντασης του σήματος (SNR), όμως συνοδεύεται από τεχνικούς περιορισμούς όπως η υψηλότερη ειδική απορρόφηση ενέργειας (SAR) και αυξημένες ανομοιογένειες πεδίου, καθώς και αυξημένο κόστος λειτουργίας.[12]

Free Induction Decay (FID), αποσυγχρονισμός και T_2 , T_2^* . Αμέσως μετά τη διέγερση, οι πυρήνες εκπέμπουν ένα ταλαντούμενο σήμα που φύνει εκθετικά επειδή οι φάσεις τους διασπείρονται: οι μικρές διαφορές τοπικού πεδίου (δημιουργούμενες από ανωμαλίες, γειτονικές μαγνητικές ροπές, ανομοιογένειες του B_0) προκαλούν σταδιακό αποσυγχρονισμό.

Το παρατηρούμενο σήμα μπορεί ιδεατά να γραφεί ως

$$S(t) = A \cos(\omega_0 t) e^{-t/T_2},$$

όπου A περιέχει το $\sin \alpha$, το N και άλλους παράγοντες. Στην πραγμάτικότητα, λόγω ανομοιογενειών, ο ρυθμός αποσυγχρονισμού που παρατηρούμε είναι ταχύτερος και ονομάζεται T_2^* (με $T_2^* \leq T_2$). Το αρχικό σήμα μετά την παλιμκή διέγερση ονομάζεται FID[12].



Σχήμα 2.4 Οι χρόνοι χαλάρωσης T_1 και T_2 είναι εγγενείς υλικές παράμετροι των ιστών και καθορίζουν τη χρονική εξέλιξη του σήματος που παράγεται. [14]

Η χρονική εξέλιξη της εγκάρσιας μαγνήτισης (χωρίς επιπλέον διεγέρσεις) δίνεται από

$$M_{xy}(t) = M_{xy}(0) e^{-t/T_2}, \quad (2.1)$$

ενώ η επαναφορά μαγνητισμού κατά τον άξονα z (μετά από διέγερση) είναι

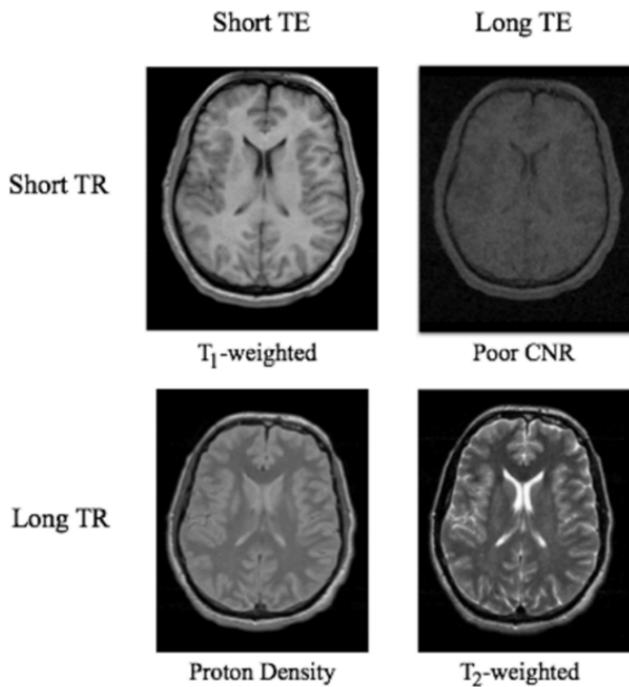
$$M_z(t) = M_0(1 - e^{-t/T_1}), \quad (2.2)$$

όπου M_0 είναι ο μαγνητισμός Ισορροπίας.

Σημείωση: Η μορφή $S(t) = A \cos(\omega_0 t) e^{-t/T_2}$ είναι απλοποιημένη. Πιο φυσικά το σήμα είναι σύνθετο και γράφεται γραμμικά:

$$S(t) = A e^{-i\omega_0 t} e^{-t/T_2^*}. \quad (2.3)$$

Οι εικόνες MRI αποκτούν βάρος σε T_1 ή T_2 επιλέγοντας το πότε (χρονικά) θα μετρήσουμε το σήμα. Δύο κρίσιμοι χειρισμοί του scanner είναι ο χρόνος επανάληψης (repetition time (TR)), ο χρόνος ανάμεσα σε επαναλαμβανόμενες διεγέρσεις — και ο χρόνος ηχούς echo time (TE), ο χρόνος από τη διέγερση μέχρι την καταγραφή του σήματος/ηχού. Γενικά, εάν επιλέξουμε TR πολύ μικρό σε σχέση με T_1 , η διαφορά εξαρτάται ισχυρά από T_1 (T1-weighted)· εάν επιλέξουμε TE της τάξης του T_2 , η διαφορά προκύπτει από (T2-weighted)· εάν $TR \rightarrow \infty$ και $TE \rightarrow 0$, τότε η εικόνα προκύπτει κυρίως από την spin density M_0 (spin-density weighting)[12].



Σχήμα 2.5 Η επιλογή των TE (χρόνος ηχούς) και TR (χρόνος επανάληψης) καθορίζει τον τύπο στάθμισης/αντίθεσης που φαίνεται στην τελική εικόνα.[12]

Σε πολλές κλινικές ακολουθίες, μετά τις απλοποιήσεις (π.χ. 90°), μπορούμε να χρησιμοποιήσουμε την μορφή σαν συνάρτηση TR και TE :

$$S(\text{TR}, \text{TE}) = (1 - e^{-\text{TR}/T_1}) M_0 e^{-\text{TE}/T_2}$$

Αυτή η έκφραση εμφανίζει τον ρόλο της επαναφοράς $(1 - e^{-\text{TR}/T_1})$ και της εκθετικής απόσβεσης $e^{-\text{TE}/T_2}$. Από αυτήν την εξίσωση γίνεται φανερό πως με την κατάλληλη επιλογή TR και TE μπορούμε να επιλέξουμε την εικόνα προς T_1 ή T_2 αντίστοιχα[12].

2.1.2 MRI και χώρος k (k-space): Από την Απόκτηση Δεδομένων στην Εικόνα

Τα πηνία (coils) Τα πηνία λήψης (coils) ενεργούν σαν κεραίες με τοπική ευαισθησία. Κάθε coil βλέπει την εικόνα μέσα από έναν χάρτη ευαισθησίας $C(\mathbf{r})$. Το σήμα από κάθε coil είναι το άθροισμα της τοπικής (σύνθετης) εγκάρσιας μαγνήτισης, κωδικοποιημένης στο χώρο των συχνοτήτων, συν θόρυβος. Η ύπαρξη πολλών coils επιτρέπει παράλληλη απεικόνιση: με διαφορετικές ευαισθησίες μπορούμε να ανακτήσουμε πληροφορία όταν το k -space είναι υποδειγματοληπτημένο.[15]

Τι είναι ο χώρος k —Συνοπτικά (αναλύεται στο κεφάλαιο δεδομένα)
Ο χώρος k είναι ο χώρος όπου αποθηκεύονται οι χωρικές συχνότητες της εικόνας.

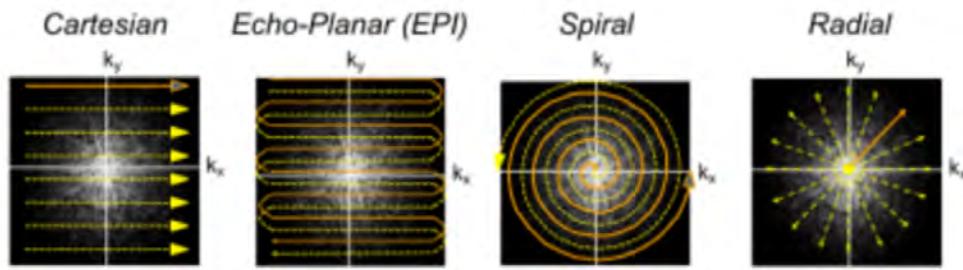
Το κέντρο του χρατάει το γενικό ανάγλυφο/αντίθεση δομών και τα άκρα χρατάνε τις λεπτομέρειες και τα όρια. Η εικόνα προκύπτει με αντίστροφο Fourier μετασχηματισμό του χώρου[17][18]. Σε συνεχή μορφή, εάν $f(\mathbf{r})$ είναι η πραγματική επιφανειακή κατανομή της μαγνήτισης (image), τότε το αντίστοιχο k -χώρος είναι

$$F(\mathbf{k}) = \int f(\mathbf{r}) e^{-i2\pi\mathbf{k}\cdot\mathbf{r}} d\mathbf{r},$$

και αντιστρόφως

$$f(\mathbf{r}) = \int F(\mathbf{k}) e^{i2\pi\mathbf{k}\cdot\mathbf{r}} d\mathbf{k}.$$

Ένα βασικό, χρηστικό σημείο: οι χαμηλές χωρικές συχνότητες (χοντά στο κέντρο του k -χώρου) μεταφέρουν την πληροφορία της γενικής έντασης και του SNR, ενώ οι υψηλές χωρικές συχνότητες (άκρα του k -χώρου) μεταφέρουν τις λεπτομέρειες και τα όρια [46]. Στην εικόνα παρουσιάζονται οι κυριότερες τροχιές δειγματοληψίας για την πλήρωση του k -χώρου, όπως η καρτεσιανή (γραμμική), η echo-planar, η σπειροειδής και η ακτινική.[16]



Σχήμα 2.6 Διαφορετικές τροχιές στον χώρο k : Καρτεσιανή, EPI (Echo-Planar Imaging), Ακτινική, Σπειροειδής. [18]

Μετάβαση σε Αντίστροφα Προβλήματα

Έχοντας περιγράψει την φυσική προέλευση του σήματος, την ψηφιοποίηση, τον ρόλο των coils και τη δομή του χώρου k , μπορούμε να μεταβούμε στο πλαίσιο των αντίστροφων προβλημάτων στο MRI. Στο 2.2 θα διατυπώσουμε την ανακατασκευή ως πρόβλημα βελτιστοποίησης.

2.2 Βασικές έννοιες των αντίστροφων προβλημάτων

Διατύπωση του αντιστρόφου προβλήματος

Έστω $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ και θόρυβος $\varepsilon \in \mathbb{R}^m$. Το **αντίστροφο πρόβλημα** συνίσταται στην ανάκτηση του x από το γραμμικό στοχαστικό μοντέλο

$$y = Ax + \varepsilon$$

όπου ο A είναι ο (γραμμικός) **προωθητικός τελεστής** και ε μοντελοποιεί στοχαστικές/συστηματικές αβεβαιότητες (θόρυβος μέτρησης και σφάλμα μοντέλου)[21].

Application	Forward model	Notes
Denoising [58]	$A = I$	I is the identity matrix
Deconvolution [58, 59]	$\mathcal{A}(x) = h * x$	h is a known blur kernel and $*$ denotes convolution. When h is unknown the reconstruction problem is known as blind deconvolution.
Superresolution [60, 61]	$A = SB$	S is a subsampling operator (identity matrix with missing rows) and B is a blurring operator corresponding to convolution with a blur kernel
Inpainting [62]	$A = S$	S is a diagonal matrix where $S_{i,i} = 1$ for the pixels that are sampled and $S_{i,i} = 0$ for the pixels that are not.
Compressive Sensing [63, 64]	$A = SF$ or $A =$ Gaussian or Bernoulli ensemble	S is a subsampling operator (identity matrix with missing rows) and F discrete Fourier transform matrix.
MRI [3]	$A = SFD$	S is a subsampling operator (identity matrix with missing rows), F is the discrete Fourier transform matrix, and D is a diagonal matrix representing a spatial domain multiplication with the coil sensitivity map (assuming a single coil acquisition with Cartesian sampling in a SENSE framework [65]).
Computed tomography [58]	$A = R$	R is the discrete Radon transform [66].
Phase Retrieval [67–70]	$\mathcal{A}(x) = Ax ^2$	$ \cdot $ denotes the absolute value, the square is taken elementwise, and A is a (potentially complex-valued) measurement matrix that depends on the application. The measurement matrix A is often a variation on a discrete Fourier transform matrix.

Σχήμα 2.7 Ενδεικτικά παραδείγματα αντίστροφων προβλημάτων[22]

Καλά/κακώς τεθειμένα προβλήματα Κατά Hadamard, το πρόβλημα είναι καλά τεθειμένο αν ισχύουν: (i) ύπαρξη λύσης, (ii) μοναδικότητα, (iii) συνεχής εξάρτηση από τα δεδομένα y . Διαφορετικά είναι κακώς τεθειμένο.

Ακριβής αντιστροφή. Αν $m = n$ και ο A είναι αντιστρέψιμος, τότε $x = A^{-1}y$ και η εξάρτηση είναι συνεχής με σταθερά Lipschitz $\|A^{-1}\|$. Η αριθμητική ευστάθεια ελέγχεται από τον **αριθμό κατάστασης** $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$.

Πλεονάζουσες/ελλιπείς μετρήσεις. Όταν $m \neq n$ ή/και A είναι ελλιπούς βαθμίδας, η απλή λύση δεν ορίζεται ή δεν είναι μοναδική. Τυπικά εξετάζουμε:

- Λύση ελαχίστων τετραγώνων

$$x = \arg \min_x \|Ax - y\|_2^2,$$

που ικανοποιεί τις κανονικές εξισώσεις $A^\top Ax = A^\top y$.

- Άπειρες λύσεις: επιλέγουμε, π.χ., την ελάχιστου κανόνα

$$x = \arg \min\{\|x\|_2 : Ax = y\}$$

Κανονικοποίηση

Για την αποκατάσταση ευστάθειας και καθοριστικότητας, εισάγουμε **όρο κανονικοποίησης** [22] $R : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ και θέτουμε

$$x^* \in \arg \min_{x \in \mathbb{R}^n} F(x), \quad F(x) = \frac{1}{2} \|Ax - y\|_2^2 + \lambda R(x), \quad \lambda > 0$$

- Αν ο R είναι κυρτός, κάτω ημι-συνεχής και συνεκτικός (ή αν ο όρος $\frac{1}{2} \|Ax - y\|_2^2$ ελέγχει τις κατευθύνσεις του πυρήνα), τότε υπάρχει λύση.
- Αν επιπλέον το F είναι ισχυρά κυρτό (π.χ. $A^\top A \succeq \mu I$ ή/και R είναι μ -ισχυρά κυρτός), η λύση είναι μοναδική.

Ισχυρή κυρτότητα & μοναδικότητα. Αν $A^\top A \succeq \mu I$ με $\mu > 0$ (πλήρης στηλοβαθμίδα) ή/και το R είναι μ -ισχυρά κυρτό, τότε το F είναι μ -ισχυρά κυρτό και η λύση είναι μοναδική. Για Tikhonov με πίνακα L : αν $\text{null}(A) \cap \text{null}(L) = \{0\}$, τότε

$$A^\top A + \lambda L^\top L \succ 0.$$

Ενδεικτικές επιλογές R .

- Tikhonov (ℓ_2): $R(x) = \lambda \|Lx\|_2^2$. Κλειστή μορφή (αν το άθροισμα είναι αντιστρέψιμο):

$$x^* = (A^\top A + \lambda L^\top L)^{-1} A^\top y.$$

- Lasso / αραιότητα (ℓ_1): $R(x) = \lambda \|x\|_1$, με

$$0 \in 2A^\top(Ax^* - y) + \lambda \partial\|x^*\|_1.$$

- Ολική μεταβολή (TV, σε πλέγμα):

$$R(x) = \lambda \|\nabla x\|_{2,1} = \lambda \sum_p \|\nabla x_p\|_2.$$

- Περιορισμοί: $R = \iota_C$ (δεικτοσυνάρτηση κυρτού κλειστού C):

$$\min_x \|Ax - y\|_2^2 \quad \text{c.t. } x \in C.$$

Αλγορίθμικά

- Βαθμωτή κάθοδος (διαφορίσιμο R , L -Lipschitz ∇F):

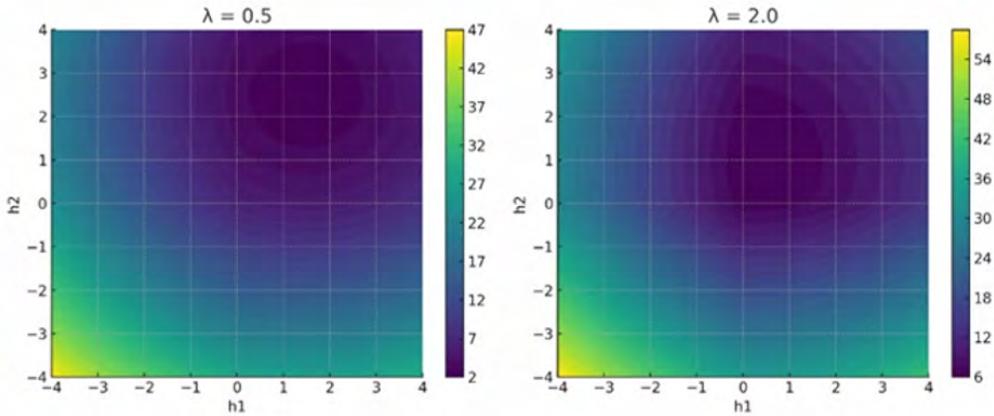
$$x_{k+1} = x_k - \tau \nabla F(x_k), \quad \tau \in (0, 2/L), \quad \nabla F = 2A^\top(Ax - y) + \nabla R(x).$$

Μέθοδοι Εγγυτητας:

$$x_{k+1} = \text{prox}_{\tau R}\left(x_k - 2\tau A^\top(Ax_k - y)\right), \quad \text{prox}_{\tau R}(z) = \arg \min_x \frac{1}{2}\|x - z\|_2^2 + \tau R(x).$$

Για $R(x) = \lambda \|x\|_1$: soft-thresholding.

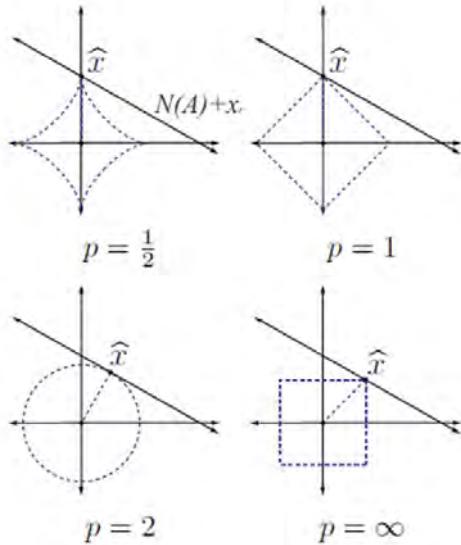
Κλιμάκωση. Η μορφή $\frac{1}{2}\|Ax - y\|_2^2 + \lambda R(x)$ Επιλογή του λ : αρχή της απόκλισης (discrepancy principle), καμπύλη L L-curve, διασταυρούμενη επικύρωση (cross-validation).



Σχήμα 2.8 Συμβιβασμός της ποινής μεταξύ αραιότητας (sparsity) και ανακατασκευής. Μικρότερη τιμή λ καλύτερη ανακατασκευή, αλλά μικρότερο όφελος από την αραιοποίηση.

Γεωμετρική ερμηνεία των l_p -νόρμων και σύνδεση με αραιότητα.
Μια ιδιαίτερα διδακτική ερμηνεία της επιλογής του κανονικοποιητή γίνεται μέσω της τομής μιας l_p -σφαίρας με το σύνολο λύσεων. Το πρώτο σημείο τομής της l_p -σφαίρας με τον χώρο λύσεων απεικονίζει τον τρόπο με τον οποίο οι διάφορες νόρμες προτιμούν διαφορετικές δομές λύσεων. Οι l_p -σφαίρες για $p > 1$ είναι στρογγυλεμένες και τείνουν να προωθούν πυκνές λύσεις, ενώ η l_1 -σφαίρα έχει γωνίες (χορνέρ) που συχνά συμπίπτουν με τους άξονες, οδηγώντας έτσι σε λύσεις με πολλούς μηδενικούς συντελεστές, μια συμπεριφορά που ευνοεί την αραιότητα. Στις εικόνες που παρατίθενται (Σχήμα 2.9), φαίνεται πώς η τομή με τις σφαίρες l_1, l_2, l_∞ και ακόμη και μη-κοίλων μορφών (π.χ. $l_{1/2}$) δίνει διαφορετικά σημεία τομής και επομένως διαφορετικές προτιμήσεις για την υποστήριξη της λύσης. Η l_0 νόρμα (μετρικά περισσότερο ένα πλήθος μη μηδενικών) ισοδυναμεί με ένα σταυρωτό

σχήμα που περιορίζει τις λύσεις πάνω στους άξονες, αλλά είναι μη-συνεχής και μη κυρτή, πράγμα που κάνει την ελαχιστοποίησή της (P_0) υπολογιστικά ανεπίτευκτη σε μεγάλα μεγέθη (NP-Hard) επομένως, συνήθως προτιμάται η ℓ_1 [25].



Σχήμα 2.9 Τομή της ℓ_p -σφαίρας με τον χώρο λύσεων της εξίσωσης $Ax = b$ για τις νόρμες ℓ_1 , ℓ_2 , ℓ_∞ και τη μη κυρτή (non-convex) $\ell_{1/2}$. Η ℓ_0 , ως ακραία περίπτωση της $\ell_{1/2}$, έχει μορφή σταυρού και έτσι τα σημεία της περιορίζονται αποκλειστικά στους άξονες. [25]

2.2.1 Σύνδεση με το πρόβλημα της ανακατασκευής MRI

Το πρόβλημα ανακατασκευής MRI μπορεί να διατυπωθεί ως αντίστροφο πρόβλημα. Το προωθητικό μοντέλο καρτεσιανής δειγματοληψίας με πολλαπλά πηνία διατυπώνεται με απλή συμβολική μορφή. Έστω $x(\mathbf{r})$ η (σύνθετη) εικόνα στο χώρο του αντικειμένου που θέλουμε να ανακτήσουμε. Για κάθε πηνίο coil $c = 1, \dots, C$ το παρατηρούμενο χώρο k είναι το αποτέλεσμα των ακόλουθων βημάτων: πολλαπλασιασμός του πεδίου εικόνας με το βάρος της ευαισθησίας $C_c(\mathbf{r})$, μετασχηματισμός Fourier \mathcal{F} , και εφαρμογή της μάσκας δειγματοληψίας M που μηδενίζει τις μη-παρατηρούμενες θέσεις του διακριτού k χώρου. Προσθέτουμε τον όρο θορύβου ε_c . Συμβολικά:

$$y_c = M \cdot \mathcal{F}\{C_c \cdot x\} + \varepsilon_c.$$

Σε διακριτό, πρακτικό επίπεδο, τα y_c είναι διανύσματα μετρήσεων (σειρές δειγμάτων κατά την κατεύθυνση k_x/k_y), το M είναι μια δυαδική μάσκα δειγματοληψίας και \mathcal{F} ο διακριτός Fourier μετασχηματισμός. Ενώνοντας τις καταγραφές όλων των πηνίων σε ένα διανύσμα $y = [y_1, \dots, y_C]^\top$, ο συνολικός προωθητικός τελεστής

Α μπορεί να γραφεί ως στοίχιση των επιμέρους τελεστών:

$$y = Ax + \varepsilon, \quad A \equiv \begin{bmatrix} M\mathcal{F}C_1 \\ \vdots \\ M\mathcal{F}C_C \end{bmatrix},$$

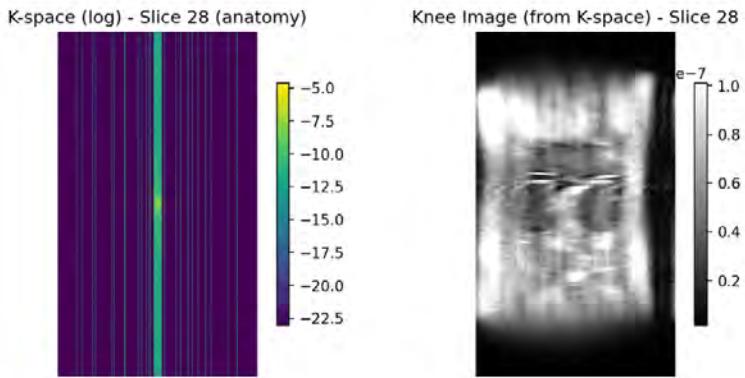
όπου κάθε $M\mathcal{F}C_c$ σημαίνει: πολλαπλασιασμός με τον χάρτη ευαισθησίας $C_c(\mathbf{r})$, Fourier μετασχηματισμός και εφαρμογή της μάσκας M .

Αυτή η συνοπτική μορφή παρουσιάζει τη Μαγνητική Τομογραφία (MRI) ως ένα αντίστροφο πρόβλημα δειγματοληψίας Fourier [27]. Στην ειδική περίπτωση συστήματος ενός πηνίου (single-coil), οι χάρτες ευαισθησίας C_c τίθενται ίσοι με 1 και το μοντέλο απλοποιείται σε

$$y = M\mathcal{F}x + \varepsilon.$$

Γιατί το MRI πρόβλημα γίνεται κακώς τεθειμένο πρόβλημα όταν υποδειγματοληπτούμε:

Η κλασική Nyquist αρχή λέει ότι για πλήρη ανακατασκευή χωρίς αναδίπλωση χρειάζεται επαρκής δειγματοληψία στο χώρο k . Όταν αραιώσουμε γραμμές ή σημεία στον χώρο αυτό (undersampling), χάνουμε πληροφορία. Το αντίστοιχο γραμμικό σύστημα γίνεται υποπροσδιορισμένο και ο χώρος λύσεων περιγράφεται από ένα υποχώρο με άπειρες λύσεις. Η αποδοτικότητα εξαρτάται από την πληροφορία των χάρτων και τη γεωμετρία της μάσκας M [26].



Σχήμα 2.10 Παράδειγμα χώρου- k με υποδειγματοληψία (undersampled) (αριστερά) και της αντίστοιχης ανακατασκευής με IFFT μετά από συμπλήρωση με μηδενικά (zero-filled) (δεξιά), που εμφανίζει απώλεια λεπτομέρειας.

Η αριστερή απεικόνιση δείχνει ότι η μάσκα δειγματοληψίας έχει αποκόψει πολλές κάθετες γραμμές στον χώρο k , και γι' αυτό η δεξιά εικόνα, που είναι ο αντίστροφος μετασχηματισμός Fourier του υποδειγματοληπτημένου k -χώρου, εμφανίζει απώλεια πληροφορίας. [26].

$$x_{\text{zf}} = \mathcal{F}^{-1}\{M \mathcal{F}\{x\}\} = x * \text{PSF}, \quad \text{PSF} = \mathcal{F}^{-1}\{M\}.$$

Επιπλέον χρήσιμες παρατηρήσεις: Κάποιες κρίσιμες ιδέες που συγχνά στο MRI ως inverse problem διατυπωμένες[27]:

Πρώτον, ο L2 όρος συνέπειας δεν χρησιμοποιείται επειδή έτσι συνηθίζεται, αλλά πηγάζει κατευθείαν από το μοντέλο θορύβου στον χώρο k . Αν ο θόρυβος είναι καλά προσεγγίσιμος ως i.i.d. Gaussian, τότε το MLE ισοδυναμεί με ελαχιστοποίηση τετραγωνικού σφάλματος:

$$\hat{x} = \arg \min_x \|s - Ex\|_2^2,$$

όπου s τα δεδομένα k χώρος και E ο encoding operator (Fourier, ευαισθησίες πηγών κ.λπ.). Αυτό είναι απλώς η στοχαστικά βέλτιστη επιλογή υπό Gaussian θόρυβο. Με άλλο μοντέλο θορύβου η συνάρτηση κόστους αλλάζει.

Δεύτερον, η σταθερότητα της λύσης ελέγχεται από το πόσο καλά ορισμένο είναι το E . Το μέτρο είναι ο αριθμός κατάστασης

$$\kappa(E) = \frac{\tau_{\max}}{\tau_{\min}} \quad (\geq 1),$$

που δίνει και δεσμό διάδοσης σφάλματος:

$$\frac{\|x - \hat{x}\|_2}{\|x\|_2} \lesssim \kappa(E) \frac{\|\varepsilon\|_2}{\|s\|_2}.$$

Με υποδειγματοληψία, η $\tau_{\min} \rightarrow 0$, άρα $\kappa(E) \rightarrow \infty$: το πρόβλημα γίνεται μη μοναδικό, ασταθές και η απλή λύση ελαχίστων τετραγώνων παθαίνει ενίσχυση θορύβου. Σε πλήρως δειγματοληπτημένο σύνολο μονού πηγίου, Καρτεσιανό σενάριο, $E = F$ και $\kappa(E) = 1$, οπότε $\hat{x} = F^{-1}s$.

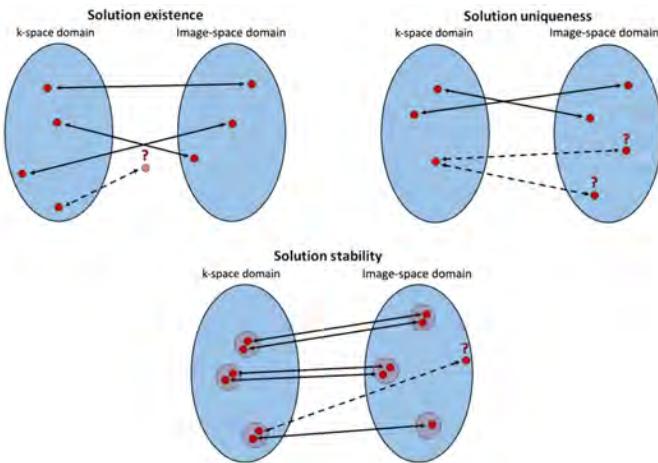
Τρίτον, όταν το πρόβλημα είναι υποπροσδιορισμένο προσθέτουμε πληροφορία (prior) ως περιορισμούς. Ισοδύναμες μορφές είναι η δεσμευμένη

$$\hat{x} = \arg \min_x R(x) \quad \text{s.t.} \quad \|s - Ex\|_2^2 \leq \varepsilon,$$

και η ποινικοποιηγένη μορφή

$$\hat{x} = \arg \min_x \|s - Ex\|_2^2 + \lambda R(x),$$

όπου $R(x)$ μπορεί να είναι $\frac{\lambda}{2} \|x\|_2^2$ (Tikhonov) για σταθεροποίηση ή $\lambda \|Wx\|_1$ (π.χ. wavelets/TV) για αραιότητα. Άρα το $R(x)$ μειώνει την αποτελεσματική διάσταση του χώρου λύσεων, κλείνοντας το χάσμα που άνοιξε το φαινόμενο της υποδειγματοληψίας, με ελεγχόμενο συμβιβασμός μεροληψίας–διασποράς.



Σχήμα 2.11 Ενδεικτικό διάγραμμα των συνθηκών που πρέπει να ικανοποιούνται ώστε ένα πρόβλημα να είναι καλώς ορισμένο well-posed. Υπαρξη λύσης σύμφωνα με το υποκείμενο μαθηματικό μοντέλο. Η λύση πρέπει να είναι μοναδική, δηλαδή ένα σύνολο μετρήσεων να αντιστοιχεί σε μία και μόνο λύση στον χώρο εικόνας. Η λύση σε μικρές διαταραχές στα μετρούμενα δεδομένα να προκαλούν μόνο μικρές διαταραχές στη λύση της εικόνας. [27]

Σύνδεση με μεθόδους: Με αυτή τη μαθηματική και πρακτική εικόνα, στο επόμενο τμήμα θα διατυπώσουμε ακριβέστερα το γραμμικό πρόβλημα A και θα περιγράψουμε μεθόδους επίλυσης προβλημάτων.

Κεφάλαιο 3

Σχετική Έρευνα στην Ανακατασκευή MRI

Στο Κεφάλαιο 3 θα συνοψίσουμε τη σχετική έρευνα στην ανακατασκευή MRI: αρχικά περιγράφουμε τις μεθόδους, από την απλή αρχικοποίηση με μηδενική πλήρωση αντίστροφου μετασχηματισμού Fourier ως τις τεχνικές με κανονικοποιητή και compressed sensing που χρησιμοποιούν ℓ_1/ℓ_2 όρους και επαναληπτικούς αλγορίθμους, εξηγώντας τα πλεονεκτήματα και τους περιορισμούς τους σε σχέση με χρόνο σάρωσης, ποιότητα εικόνας και υπολογιστικό κόστος. Στη συνέχεια παρουσιάζουμε τις προσεγγίσεις με μάθηση, όπου νευρωνικά δίκτυα και τεχνικές (ξεδίπλωσης αλγορίθμων) ενσωματώνουν τη δομή του προωθητικού μοντέλου στο δίκτυο, επιτρέποντας την εκμάθηση από δεδομένη και έτσι επιτυγχάνεται καλύτερος συμβιβασμός μεταξύ ποιότητας και ταχύτητας. Τέλος, επισημαίνουμε πρακτικά ζητήματα αξιολόγησης (οπτική/ποσοτική ποιότητα, χρόνος εκτέλεσης) που καθορίζουν την καταλληλότητα κάθε μεθόδου για κλινικές εφαρμογές και τη βασική ιδέα των μοντέλων Ισορροπίας.

3.1 Χρήσιμο Θεωρητικό Υπόβαθρο για Μεθόδους Αραιής Αναπαράστασης

Στην πράξη το κυρίαρχο περιοριστικό στοιχείο στο MRI είναι ο χρόνος σάρωσης. Η ποιότητα και η λεπτομέρεια μιας εικόνας συνδέονται άμεσα με την πυκνότητα δειγματοληφίας στο k -χώρο. Περισσότερα δείγματα σημαίνουν μεγαλύτερο χρόνο εξέτασης, αυξημένη ευαισθησία στο κίνδυνο κίνησης του ασθενούς και χαμηλότερη κλινική διαθεσιμότητα. Γι' αυτό όπως είδαμε, το να δούμε τον k -χώρο ως έναν υποδειγματοληπτέο χώρο μπορεί να ενταχθεί ως ένα αντίστροφο πρόβλημα και επιδιώκει να μειώσει το χρόνο χωρίς να υποβαθμίσει το διαγνωστικά κρίσιμο περιεχόμενο.

Στο μαθηματικό πλαίσιο που έχουμε ήδη θέσει, αυτή η πρακτική αντιστοιχεί στο να κάνουμε τον forward τελεστή A μη-πλήρη μέσω της μάσκας M , γράφοντας το

μοντέλο είτε σε ενός πηνίου μορφή

$$y = M\mathcal{F}x + \varepsilon,$$

είτε σε πολλαπλών πηνίων μορφή

$$y = Ax + \varepsilon, \quad A \equiv \begin{bmatrix} M\mathcal{F}C_1 \\ \vdots \\ M\mathcal{F}C_C \end{bmatrix}.$$

Zero-filled IFFT και PSF: Η ταχύτερη προσέγγιση είναι να λάβουμε το υποδειγματοληπτημένο k -χώρο, να μηδενίσουμε (zero-fill)[26][30] τα μη μετρημένα σημεία και να εφαρμόσουμε απευθείας αντίστροφο Fourier:

$$\widehat{x_{\text{zf}}} = \mathcal{F}^{-1}\{M \cdot \mathcal{F}\{x\}\}.$$

Αυτή η επανασύνθεση απαιτεί μία μόνο IFFT και είναι πρακτικά άμεσα διαθέσιμη. Μαθηματικά όμως ισχύει η ταυτότητα

$$\widehat{x_{\text{zf}}} = x * \text{PSF}, \quad \text{PSF} = \mathcal{F}^{-1}\{M\},$$

Η συγκεκριμένη στρατηγική είναι χρήσιμη σε εφαρμογές όπου η απώλεια πληροφορίας είναι αμελητέα.[26] Η zero-filled IFFT είναι επομένως χρήσιμη για γρήγορη οπτική εκτίμηση και ως warm-start για επαναληπτικούς αλγορίθμους, αλλά σπάνια επαρκεί μόνη της όταν η επιτάχυνση είναι σημαντική.

Compressed sensing MRI: Στο compressed sensing(Συμπιεστική ανίχνευση) η βασική ιδέα είναι ότι αντί να μετράμε άμεσα όλα τα pixel της εικόνας x , παίρνουμε πολύ λιγότερες γραμμικές προβολές του σήματος με μια μετρητική μήτρα Φ . Το απλό γραμμικό μοντέλο μέτρησης γράφεται

$$y = \Phi x,$$

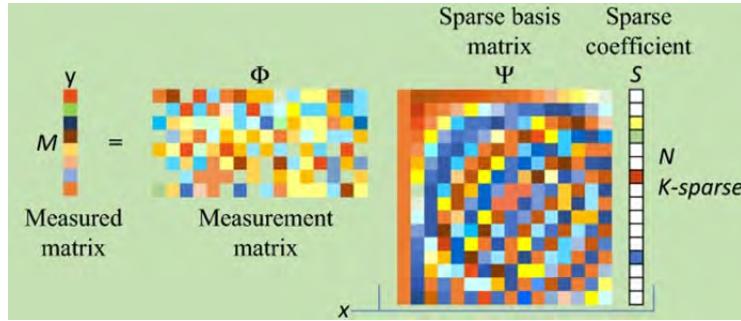
όπου y είναι το διάνυσμα των παρατηρήσεων (πολύ μικρότερης διάστασης από το x). Εάν το x εκφράζεται σε κάποια άλλη βάση ή λεξικό Ψ (π.χ. κυματομορφικές βάσεις — wavelets), και γράφουμε:

$$x = \Psi a,$$

όπου a είναι το διάνυσμα των συντελεστών στην βάση Ψ . Συνδυάζοντας τις δύο σχέσεις προκύπτει το ισοδύναμο μειωμένο σύστημα:

$$y = \Phi \Psi a = \Theta a,$$

όπου $\Theta \equiv \Phi \Psi$ [28]. Στην περίπτωση που το x είναι αραιό ή συμπιέσιμο στην Ψ -βάση, το a έχει λίγους μη μηδενικούς όρους. Αυτό είναι το κεντρικό σχήμα: αντί να ψάξουμε για οποιαδήποτε λύση του υποπροσδιορισμένου συστήματος, ψάχνουμε τη λιγότερο πυκνή (sparsest) λύση που ικανοποιεί τις μετρήσεις.



Σχήμα 3.1 Απεικόνιση της έννοιας της συμπιεστικής δειγματοληψίας (Compressed Sensing)[29]

Για να δουλέψει αυτό το σχήμα χρειάζονται δύο συμπληρωματικές ιδιότητες. Πρώτον, το σήμα πρέπει να είναι πραγματικά αραιό ή συμπιέσιμο σε κάποια Ψ -βάση — αν αυτό δεν ισχύει, το ℓ_1 δεν θα αναπαράγει καλά την αληθινή εικόνα. Δεύτερον, ο συνδυασμός της μήτρας μέτρησης Φ και της βάσης Ψ πρέπει να είναι «αρκετά μη-συγγενής» (incoherent). Στην πρακτική εφαρμογή του Cartesian MRI, η incoherence επιτυγχάνεται με τυχαιοποιημένα / variable-density μοτίβα υποδειγματοληψίας στο k -space, που κάνουν το aliasing στον image domain πιο κατάλληλο για ανάκτηση με αραιότητα.[28]

Οι αλγόριθμοι που υλοποιούν αυτά τα χριτήρια περιλαμβάνουν επαναληπτικές μεθόδους όπως ISTA ή ADMM, σπάει το αρχικό πρόβλημα σε απλούστερα υποβήματα (γραμμικό solve και prox step) και είναι ιδιαίτερα χρήσιμο για σύνθετους regularizers.

Συγκεκριμένα έχουμε:

$$(1) \text{Original problem: } \min_{x,z} F(x) + G(z) \quad \text{s.t.} \quad Ax + Bz = c$$

$$\Rightarrow y^{k+1} \rightarrow y^k + \rho (Ax^{k+1} + Bz^{k+1} - c)$$

$$\Rightarrow \text{Subproblem: } x^{k+1} \rightarrow \operatorname{argmin}_x F(x) + (y^k)^T (Ax + Bz^k - c) + \frac{\rho}{2} \|Ax + Bz^k - c\|_2^2$$

$$\Rightarrow \text{Subproblem: } z^{k+1} \rightarrow \operatorname{argmin}_z G(z) + (y^k)^T (Ax^{k+1} + Bz - c) + \frac{\rho}{2} \|Ax^{k+1} + Bz - c\|_2^2$$

Με το ADMM το δεσμευμένο πρόβλημα διασπάται σε δύο απλά υποβήματα και μία ενημέρωση πολλαπλασιαστή[37]. Η εναλλαγή των βημάτων μειώνει προοδευτικά το κατάλοιπο του περιορισμού και επιτρέπει να χειριστούμε χωριστά τον όρο πιστότητας και τον όρο κανονικοποίησης. Σε ανακατασκευή MRI, το πρώτο υποβήμα επιλύεται συνήθως με γρήγορο γραμμικό λύτη, ενώ το δεύτερο ισοδυναμεί με prox τελεστή για την αραιότητα. Η παράμετρος ρ ρυθμίζει πόσο αυστηρά επιβάλλεται ο περιορισμός σε κάθις επανάληψη.

(2) LASSO example: assume $b = Ax^{\text{true}} + \text{noise}$; recover x^{true} by optimization $x^{\text{lasso}} \leftarrow \underset{x}{\text{minimize}} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1$ also known as

ℓ_1 -regularized least-squares and compressed sensing [28]

$$x^{k+1} = \eta_{\lambda\alpha} \left(x^k - \alpha A^T (Ax^k - b) \right)$$

- convergence requires a proper stepsize α or line search
- the gradient-descent step reduces $\frac{1}{2} \|Ax - b\|^2$
- the soft-thresholding step $\eta_{\lambda\alpha}(\cdot)$ reduces $\lambda \|x\|_1$

Το παράδειγμα LASSO δείχνει πώς η κανονικοποίηση ℓ_1 προωθεί αραιές λύσεις, ενώ ο τετραγωνικός όρος κρατά τις εκτιμήσεις κοντά στις μετρήσεις. Η ενημέρωση συνδυάζει κίνηση gradient για την πιστότητα και soft-thresholding για την αραιότητα. Στην πράξη χρησιμοποιούνται μέθοδοι ISTA/FISTA ή proximal gradient για γρήγορη σύγκλιση. Με σωστή επιλογή βήματος και της λ , η διαδικασία αποδίδει σταθερά αραιές ανακατασκευές, ιδιαίτερα χρήσιμες σε compressed sensing MRI.[28]

3.2 Προσεγγίσεις με μάθηση

Πριν προχωρήσουμε στην μέθοδο ξεδιπλώματος και στην εφαρμογή του στα προβλήματα ανακατασκευής MRI, αξίζει να σταθούμε λίγο στις βασικές έννοιες της Μηχανικής Μάθησης και των Νευρωνικών Δικτύων. Αυτό που ακολουθεί είναι μια εισαγωγή: τι εννοούμε όταν μιλάμε για μάθηση από δεδομένα (learn from data)[32][34], πώς δομείται και μαθαίνει ένα νευρωνικό δίκτυο, γιατί οι παράμετροι αλλάζουν μεσόδους όπως η Καθοδική Κλίση Gradient Descent, τι ρόλο παίζουν οι συναρτήσεις ενεργοποίησης, η έννοια της οπισθοδιάδοσης και ποια πρακτικά θέματα πρέπει να προσέχουμε στην εκπαίδευση.

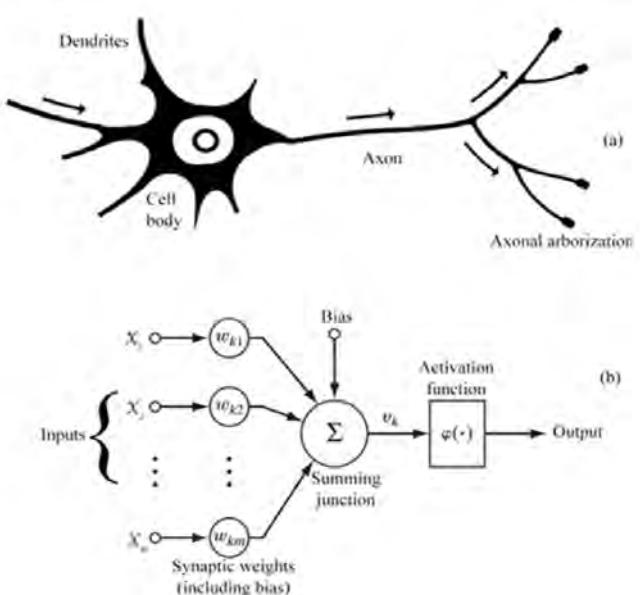
Βασικές Έννοιες Μάθησης και Νευρωνικών Δικτύων: Η Μηχανική Μάθηση, σε πολύ γενικές γραμμές, αφορά την εξαγωγή κανόνων, προτύπων και σχέσεων από δεδομένα με σκοπό την πρόβλεψη ή την απόφαση σε νέα, μη-εκπαιδευμένα παραδείγματα. Αντί να προγραμματίζουμε ρητά κανόνες, δίνουμε στο σύστημα ένα σύνολο παραδειγμάτων εισόδου-εξόδου και θέτουμε ως στόχο να μάθει μια συνάρτηση f που προσεγγίζει τον ζητούμενο μετασχηματισμό[32]. Η πρακτική μορφή αυτού του στόχου είναι η ελαχιστοποίηση μιας συνάρτησης κόστους $L(\theta)$ ως προς τις παραμέτρους θ ενός μοντέλου, που μετρά την αναντιστοιχία μεταξύ των προβλέψεων του μοντέλου και των πραγματικών τιμών στόχων στα δεδομένα εκπαίδευσης.

Μια από τις πιο ισχυρές και ευέλικτες οικογένειες μοντέλων είναι τα Νευρωνικά Δίκτυα. Ένα νευρωνικό δίκτυο αποτελείται από μονάδες (τεχνητούς νευρώνες)

οργανωμένες σε επίπεδα. Εμπνευσμένο από τη βιολογική φύση των νευρώνων[31], κάθε νευρώνας λαμβάνει έναν γραμμικό συνδυασμό εισόδων .Το κάθε στοιχείο εισόδου πολλαπλασιάζεται με ένα βάρος και σε όλα προστίθεται μια μεταβλητή μετατόπιση (bias) και στο αποτέλεσμα εφαρμόζεται μια μη-γραμμική συνάρτηση ενεργοποίησης. Η απλούστερη μαθηματική περιγραφή της εξόδου ενός νευρώνα είναι

$$y = f\left(\sum_i w_i x_i + b\right),$$

όπου x_i οι είσοδοι, w_i τα αντίστοιχα βάρη, b το bias και f η ενεργοποιητική συνάρτηση. Η μη-γραμμικότητα στην f είναι κρίσιμη: χωρίς αυτή, ακόμα και ένα πολυεπίπεδο δίκτυο θα ήταν ισοδύναμο με ένα απλό γραμμικό μοντέλο.

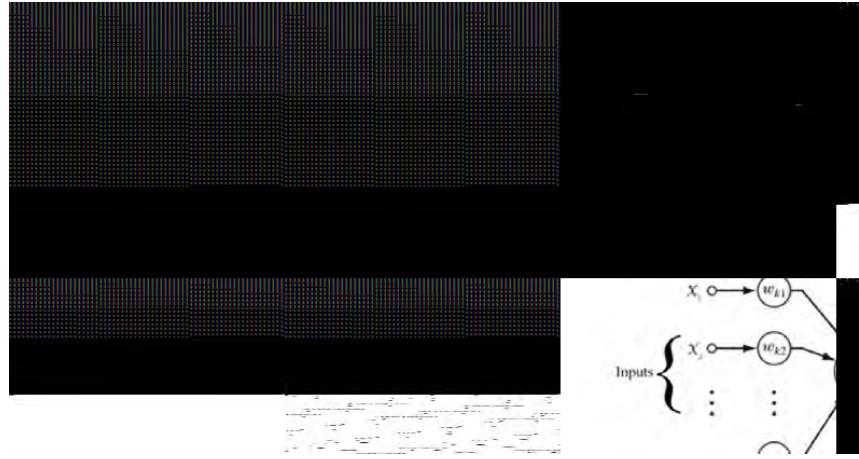


Σχήμα 3.2 Διάγραμα ενός βιολογικού/τεχνητού νευρώνα [31]

Στο επίπεδο οργάνωσης, ένα τυπικό βαθύ δίκτυο (deep neural network[38]) έχει ένα επίπεδο εισόδου, ένα ή περισσότερα χρυφά επίπεδα και ένα επίπεδο εξόδου. Οι τιμές που παράγονται από το ένα επίπεδο τροφοδοτούν το επόμενο κατά την εμπρόσθια διάδοση (forward pass). Συγκεκριμένα, αν $a^{(l)}$ είναι οι ενεργοποιήσεις στο επίπεδο l , τότε για το επόμενο επίπεδο ισχύει

$$z^{(l+1)} = W^{(l)} a^{(l)} + b^{(l)}, \quad a^{(l+1)} = \varphi(z^{(l+1)}),$$

όπου $W^{(l)}$ ο πίνακας βαρών, $b^{(l)}$ το διάνυσμα των bias και φ η στοιχειώδης μη-γραμμικότητα[38]. Η έξοδος του δικτύου συγχρίνεται με την επιθυμητή τιμή μέσω μιας συνάρτησης κόστους L . Η εκπαίδευση σημαίνει να βρούμε τις παραμέτρους $\{W^{(l)}, b^{(l)}\}$ που ελαχιστοποιούν το συνολικό κόστος πάνω στο σύνολο εκπαίδευσης.



Σχήμα 3.3 Βαθύ νευρωνικό δίκτυο (είσοδος κρυμμένα στρώματα, έξοδος)[39]

Συναρτήσεις κόστους (loss) Κανόνας μάθησης και επιλογή συνάρτησης απώλειας[32]

Θεωρούμε παραμετρικό μοντέλο $f(x; w)$ και δείγμα $\{(x_i, y_i)\}_{i=1}^n$. Η ελάχιστη προϋπόθεση μάθησης είναι ο ορισμός διαφορίσιμου λειτουργικού κόστους (εμπειρικού κινδύνου)

$$\mathcal{L}(w) = \frac{1}{n} \sum_{i=1}^n \ell(\hat{y}_i, y_i) + \gamma \mathcal{R}(w), \quad \hat{y}_i = f(x_i; w),$$

και ενός κανόνα ενημέρωσης παραμέτρων που να παρέχει κατεύθυνση καθόδου, π.χ.

$$w_{t+1} = w_t - \eta_t \nabla_w \mathcal{L}(w_t),$$

ή, σε μικροπαρτίδες(mini-batch) B_t ,

$$w_{t+1} = w_t - \eta_t \frac{1}{|B_t|} \sum_{i \in B_t} \nabla_w \ell(f(x_i; w_t), y_i).$$

Για επαρκώς μικρό ρυθμό μάθησης $\eta_t > 0$ ισχύει $\mathcal{L}(w_{t+1}) < \mathcal{L}(w_t)$.

(α) Πιλινδρόμηση (συνεχής έξοδος) Τετραγωνική απώλεια (μέση τετραγωνική απόκλιση):

$$\ell(\hat{y}, y) = \frac{1}{2} \|\hat{y} - y\|_2^2, \quad \nabla_{\hat{y}} \ell = \hat{y} - y.$$

Κατά συνέπεια, με κανόνα αλυσίδας,

$$\nabla_w \ell = J_f(x; w)^\top (\hat{y} - y),$$

όπου J_f ο Ιακωβιανός πίνακας του f ως προς w .

(β) Ταξινόμηση (πολυκλασική, softmax) Έστω logits $z \in \mathbb{R}^K$ και πιθανότητες κλάσεων μέσω softmax (εκθετικής κανονικοποίησης)

$$p_k(x; w) = \frac{e^{z_k(x; w)}}{\sum_{j=1}^K e^{z_j(x; w)}}, \quad k = 1, \dots, K,$$

με στόχο $y \in \{e_1, \dots, e_K\}$. Η διασταυρούμενη εντροπία (αρνητική λογαριθμική πιθανοφάνεια) είναι

$$\ell(z, y) = - \sum_{k=1}^K y_k \log p_k, \quad \frac{\partial \ell}{\partial z_k} = p_k - y_k.$$

Στη διυδική περίπτωση ($K = 2$) με σιγμοειδή πιθανότητα $p = \sigma(z) = \frac{1}{1+e^{-z}}$,

$$\ell(z, y) = -[y \log p + (1-y) \log(1-p)], \quad \frac{\partial \ell}{\partial z} = p - y.$$

Οι παραπάνω παράγωγοι, μέσω του κανόνα αλυσίδας, δίνουν τον απαιτούμενο μηχανισμό ενημέρωσης βαρών για τη μείωση του σφάλματος σε κάθε βήμα βελτιστοποίησης.

Θεωρούμε παραμετρικό μοντέλο $f(x; w)$ με παραμέτρους $w \in \mathbb{R}^d$ και συνάρτηση κόστους

$$L(w) = \frac{1}{n} \sum_{i=1}^n \ell(f(x_i; w), y_i) + \gamma R(w),$$

όπου ℓ είναι η στοιχειώδης απώλεια και R (π.χ. $\frac{1}{2} \|w\|_2^2$) κανονικοποιητής.

Βασικός κανόνας ενημέρωσης (καθοδική κλίση) Ο θεμελιώδης αλγόριθμος ελαχιστοποίησης είναι η καθοδική μέθοδος κλίσης[34]:

$$w_{t+1} = w_t - \eta_t \nabla L(w_t),$$

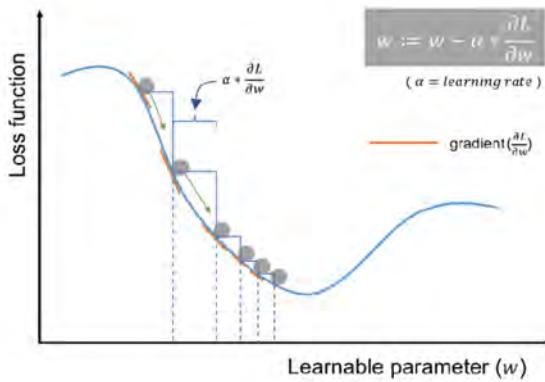
ή, σε επίπεδο μεμονωμένης παραμέτρου w_j ,

$$w_j \leftarrow w_j - \eta \frac{\partial L}{\partial w_j} [33]$$

Ο θετικός συντελεστής η_t είναι ο ρυθμός μάθησης. Στη στοχαστική εκδοχή (μηχροπαρτίδες B_t) χρησιμοποιείται αμερόληπτη εκτίμηση της κλίσης

$$g_t = \frac{1}{|B_t|} \sum_{i \in B_t} \nabla_w \ell(f(x_i; w_t), y_i) + \gamma \nabla R(w_t), \quad \mathbb{E}[g_t] = \nabla L(w_t),$$

και ο ενημερωτικός κανόνας γίνεται $w_{t+1} = w_t - \eta_t g_t$.



Φig. 3.4 Η διαισθητική ιδέα του gradient descent[33]

Αποτελεσματικός υπολογισμός παραγώγων-Οπισθοδιάδοση Για βαθιά δίκτυα ο υπολογισμός $\nabla L(w)$ υλοποιείται αποδοτικά με οπισθοδιάδοση[35] (κανόνας αλυσίδας).

Έστω πολυστρωματικό προωθητικό δίκτυο με στρώματα $\ell = 1, \dots, L$:

$$h^{(0)} = x, \quad z^{(\ell)} = W^{(\ell)} h^{(\ell-1)} + b^{(\ell)}, \quad h^{(\ell)} = \phi^{(\ell)}(z^{(\ell)}),$$

και έζοδο $\hat{y} = h^{(L)}$. Ορίζουμε τους συντελεστές σφάλματος

$$\delta^{(\ell)} \equiv \frac{\partial L}{\partial z^{(\ell)}} \in \mathbb{R}^{\dim(z^{(\ell)})}.$$

Τότε ισχύουν οι αναδρομικοί τύποι:

$$\delta^{(L)} = \frac{\partial L}{\partial h^{(L)}} \odot \phi^{(L)\prime}(z^{(L)}), \quad \delta^{(\ell)} = (W^{(\ell+1)\top} \delta^{(\ell+1)}) \odot \phi^{(\ell)\prime}(z^{(\ell)}), \quad \ell = L-1, \dots, 1,$$

όπου \odot είναι ο κατά-συνιστώσα πολλαπλασιασμός. Οι πραγματικές κλίσεις προϋπουν τοπικά:

$$\boxed{\frac{\partial L}{\partial W^{(\ell)}} = \delta^{(\ell)} h^{(\ell-1)\top}, \quad \frac{\partial L}{\partial b^{(\ell)}} = \delta^{(\ell)}}.$$

Οι υπολογισμοί αυτοί έχουν κόστος τάξης ενός ακόμη προωθητικού περάσματος και καθιστούν εφικτή την εκπαίδευση δικτύων με πολύ μεγάλο πλήθος παραμέτρων[35].

Ειδική απλοποίηση για ταξινόμηση με διασταυρούμενη εντροπία
Αν το τελικό στρώμα παράγει λογάριθμους πιθανοτήτων μέσω εκθετικής κανονικοποίησης και η απώλεια είναι διασταυρούμενη εντροπία, τότε

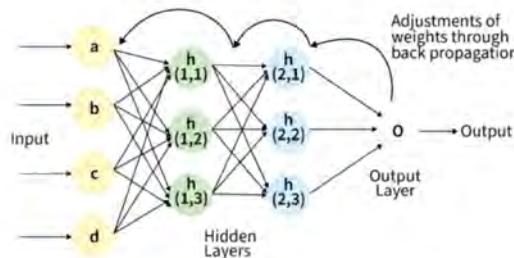
$$\delta^{(L)} = p - y,$$

όπου p το διάνυσμα προβλεπόμενων πιθανοτήτων και y η διανυσματική στόχευση.

Παρατηρήσεις ρύθμισης: (α) Ο ρυθμός μάθησης η πρέπει να ικανοποιεί κριτήρια ευστάθειας (π.χ. $\eta < 2/L_g$ σε ομαλές περιπτώσεις).

(β) Η προσθήκη ℓ_2 κανονικοποίησης ($R(w) = \frac{1}{2}\|w\|_2^2$) προσθέτει όρο γω στην κλίση.

(γ) Σε στοχαστική εκπαίδευση χρησιμοποιείται συχνά φθίνουσα ακολουθία η_t για σύγκλιση σε στάσιμα σημεία.



Σχήμα 3.5 Αναπαράσταση της οπισθοδιάδοσης[35]

Συναρτήσεις ενεργοποίησης: ρόλος, μαθηματικές ιδιότητες και επιδράσεις στην οπισθοδιάδοση

Οι συναρτήσεις ενεργοποίησης $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ εισάγουν τη μη γραμμικότητα στα στρώματα και καθορίζουν τη ροή των κλίσεων κατά την οπισθοδιάδοση. Σε προ-
ωθητικό δίκτυο

$$h^{(0)} = x, \quad z^{(\ell)} = W^{(\ell)} h^{(\ell-1)} + b^{(\ell)}, \quad h^{(\ell)} = \varphi^{(\ell)}(z^{(\ell)}),$$

οι συντελεστές σφάλματος ικανοποιούν

$$\delta^{(\ell)} = \frac{\partial L}{\partial z^{(\ell)}} = (W^{(\ell+1)})^\top \delta^{(\ell+1)} \odot \varphi^{(\ell)'}(z^{(\ell)}),$$

Παρακάτω συνοψίζονται οι πλέον καθιερωμένες ενεργοποιήσεις, με έμφαση σε πα-
ράγωγα και ιδιότητες που επηρεάζουν την οπισθοδιάδοση.

Σιγμοειδής

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \quad \sigma'(x) = \sigma(x)(1 - \sigma(x)) \in (0, \frac{1}{4}].$$

Φραγμένη ($0 < \sigma < 1$), κορεσμός για $|x| \gg 0 \Rightarrow \sigma'(x) \approx 0$ και έντονα εξαφανι-
ζόμενες κλίσεις σε βαθιά δίκτυα. Μη μηδενικό μέσο στην έξοδο.

Τπερβολική εφαπτομένη (tanh)

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad \tanh'(x) = 1 - \tanh^2(x) \in (0, 1].$$

Κεντραρισμένη στο μηδέν, αλλά επίσης κορεσμένη για $|x| \gg 0 \Rightarrow$ εξασθένηση
κλίσεων.

ReLU (γραμμική μονάδα αποκοπής)

$$\text{ReLU}(x) = \max(0, x), \quad \text{ReLU}'(x) = \begin{cases} 0, & x < 0, \\ 1, & x > 0, \end{cases} \quad \partial \text{ReLU}(0) = [0, 1].$$

Μη κορεσμένη για $x > 0$, επιτρέπει ισχυρή ροή κλίσεων και αραιές ενεργοποιήσεις

Διαρρέουσα ReLU (Leaky ReLU)

$$\text{LReLU}(x) = \max(\alpha x, x), \quad \text{LReLU}'(x) = \begin{cases} \alpha, & x < 0, \\ 1, & x \geq 0, \end{cases} \quad \alpha \in (0, 1) \ (\pi.\chi. 0.01).$$

Διατηρεί μικρή αλλά μη μηδενική κλίση στον αρνητικό ημιάξονα.

ELU (εκθετική μονάδα αποκοπής)

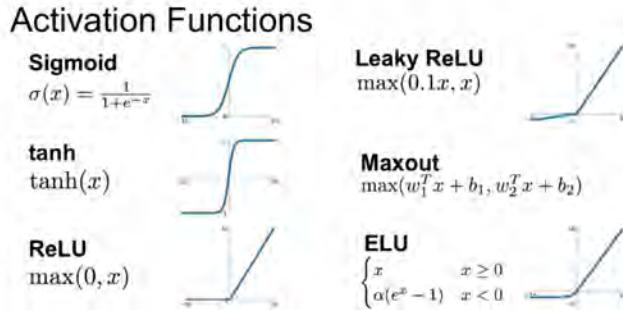
$$\text{ELU}(x) = \begin{cases} x, & x \geq 0, \\ \alpha(e^x - 1), & x < 0, \end{cases} \quad \text{ELU}'(x) = \begin{cases} 1, & x \geq 0, \\ \alpha e^x, & x < 0, \end{cases} \quad (\alpha > 0).$$

Παράγει αρνητικές εξόδους για $x < 0$ (κεντράρισμα γύρω από το 0) και ήπιο κορεσμό.

Maxout (μονάδα μέγιστου)

$$\text{Maxout}(x) = \max_{j=1,\dots,m} (w_j^\top x + b_j),$$

Οδηγίες επιλογής (καθιερωμένες) Κρυψμένα στρώματα: ReLU ή παραλλαγές της (Leaky/ELU) για βελτιωμένη ροή κλίσεων· έξοδος δυαδικής ταξινόμησης: σιγμοειδής (σ) με cross-entropy· έξοδος πολυκλασικής ταξινόμησης: softmax με cross-entropy.

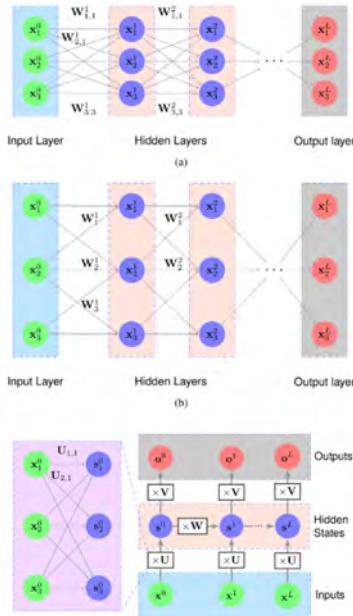


Σχήμα 3.6 Παραδείγματα διαφορετικών συναρτήσεων ενεργοποίησης[36]

Τέλος, μερικά χρίσιμα πρακτικά σημεία που επηρεάζουν την εκπαίδευση[34]:

- **Επιλογή ρυθμού μάθησης** (learning rate, η): κρίσιμος υπερπαράμετρος — μικρό $\eta \Rightarrow$ αργή σύγκλιση, μεγάλο $\eta \Rightarrow$ αστάθεια. Συχνά χρησιμοποιούνται χρονοπρογράμματα ρυθμού μάθησης (learning-rate schedules) ή μηχανισμοί (callbacks) που μειώνουν το η σταδιακά.
- **Αριθμός εποχών** (epochs): κάθε εποχή περιλαμβάνει την πρόσθια διάδοση (forward pass) όλων των δειγμάτων και την αντίστοιχη οπισθοδιάδοση.
- **Αρχικοποίηση βαρών** (weights initialization): κατάλληλη αρχικοποίηση σταθεροποιεί την κλίμακα ενεργοποίησεων και κλίσεων, μειώνοντας φαινόμενα εξαφάνισης/έκρηξης.
- **Πρόωρη διακοπή** (early stopping): παρακολουθούμε την απώλεια στο σύνολο επικύρωσης(validation set) και σταματάμε όταν πάύει να βελτιώνεται ή αρχίζει να επιδεινώνεται.

- **Εξαφάνιση/έκρηξη κλίσεων** (vanishing/exploding gradients): αντιμετώπιση με κατάλληλες συναρτήσεις ενεργοποίησης (π.χ. ReLU/Leaky ReLU), αποκόπη κλίσεων (gradient clipping), υπολειμματικές συνδέσεις(residual/skip connections), σωστή αρχικοποίηση και κανονικοποίηση.



Σχήμα 3.7 Διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων:
(πλήρη-συνδεδεμένο),(συνελεκτικό),(recurrent/ξεδιπλωμένες επαναλήψεις)[40].

Εδώ παρουσιάζονται τα εξής σχήματα:

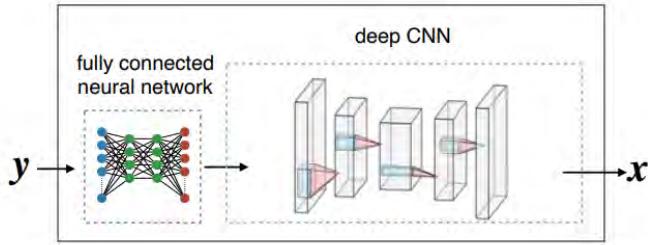
1. Ένα πλήρως συνδεδεμένο δίκτυο (fully-connected). Κάθε κόμβος ενός επιπέδου παίρνει είσοδο από όλους τους κόμβους του προηγούμενου επιπέδου. Κάθε νευρώνας εφαρμόζει έναν γραμμικό συνδυασμό και μετά μια μη-γραμμικότητα.
2. Η ιδέα των τοπικών συνδέσεων με επαναχρησιμοποιούμενα βάρη, που είναι το βασικό μοτίβο των συνελεκτικών δικτύων. Αντί για πλήρη σύνδεση χρησιμοποιούνται μικρά φίλτρα που εφαρμόζονται τοπικά, επιτυγχάνοντας οικονομία παραμέτρων και αποδοτική εκμάθηση χωρικών χαρακτηριστικών.
3. Ένα επαναληπτικό τύπο δικτύου (RNN / unrolled iteration): κάθε επανάληψη μοιάζει με ένα χρονικό βήμα — αν τα βάρη είναι κοινά (tied) μοιάζει με RNN, ενώ αν τα αφήσουμε διαφορετικά ανά βήμα μοιάζει με βαθύ δίκτυο που προσομοιώνει επαναλήψεις[40].

3.2.1 Μέθοδος Ξετυλίγματος Αλγορίθμου και δίκτυα για ανακατασκευή MRI

Ενα κανονικοποιημένο αντίστροφο πρόβλημα μπορεί να διατυπωθεί ως

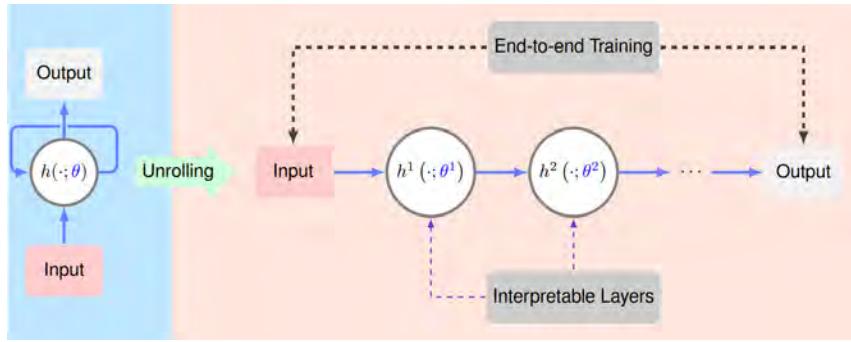
$$\hat{x} = \arg \min_x \frac{1}{2} \|Ax - y\|_2^2 + R(x),$$

όπου $R(x)$ είναι ο κανονικοποιητής [24] (Tikhonov, ℓ_1 σε wavelets κ.ά.). Υλοποιώντας αυτού του τύπου βελτιστοποίηση με κλασικούς αλγορίθμους (ISTA/FISTA, ADMM) προκύπτουν επαναληπτικές μέθοδοι. Οι πρόσφατες πρόοδοι στα μαθηματικά και στην υπολογιστική επιστήμη έδειξαν ότι μπορούμε να επιτύχουμε ικανοποιητικό συμβιβασμό ποιότητας–πολυπλοκότητας, αν ορισμένα από τα βήματα της βελτιστοποίησης γίνουν καθοδηγούμενα από τα δεδομένα. Από αυτή την παρατήρηση προέκυψε η ιδέα του ξετυλίγματος των αλγορίθμων.



Σχήμα 3.8 Όταν το προωθητικό μοντέλο A είναι άγνωστο, αλλά υπάρχουν επαρκή δεήγματα εκπαίδευσης, μπορεί να μάθει κανείς μια συνάρτηση ανακατασκευής χρησιμοποιώντας βαθιά νευρωνικά δίκτυα [24].

Η ιδέα του ξετυλίγματος είναι να πάρουμε έναν πεπερασμένο αριθμό επαναλήψεων ενός επαναληπτικού αλγορίθμου και να τον θεωρήσουμε ως ένα πεπερασμένο, βαθύ δίκτυο: κάθε επανάληψη αντιστοιχεί σε ένα στρώμα [40]. Στον παραδοσιακό αλγόριθμο οι πράξεις και οι υπερπαράμετροι (π.χ. φίλτρα, κατώφλια, βήμα) είναι σταθερές και προεπιλεγμένες. Με το ξετύλιγμα επιτρέπουμε σε μερικές (ή όλες) από αυτές τις ποσότητες να είναι εκπαιδεύσιμες [40] και εκπαιδεύουμε το σύνολο από όχρη σε όχρη πάνω σε ζεύγη (αραιά δειγματοληπτημένο \rightarrow σήμα/εικόνα αναφοράς). Με αυτόν τον τρόπο διατηρούμε τη δομή της κλασικής μεθόδου (συνέπεια ως προς τα δεδομένα), ενώ μαθαίνουμε βήματα προσαρμοσμένα στα δεδομένα της εφαρμογής μας, επιτυγχάνοντας καλύτερο συμβιβασμό ποιότητας–πολυπλοκότητας. Το ξετύλιγμα λειτουργεί ως γέφυρα ανάμεσα στη βελτιστοποίηση με βάση το μοντέλο και στη μάθηση καθοδηγούμενη από τα δεδομένα: διατηρεί τη δομή του φυσικού μοντέλου, αλλά επιτρέπει στο δίκτυο να μάθει τον καταλληλότερο τελεστή εγγύτητας ή αποθορυβοποιητή, ή τον κατάλληλο γραμμικό μετασχηματισμό σε κάθε επανάληψη.



Σχήμα 3.9 Συνοπτική απεικόνιση *algorithmic unrolling*. Δεδομένου επαναληπτικού αλγορίθμου με βήμα $h(\cdot; \theta)$ (αριστερά), παράγεται αντίστοιχο βαθύ δίκτυο (δεξιά) με διαδοχική στοίβαξη των επαναλήψεων σε *layers* $h^1(\cdot; \theta^1), h^2(\cdot; \theta^2), \dots$. Οι παράμετροι εκπαιδεύονται *end-to-end*, και τα *layers* κληρονομούν ερμηνευσιμότητα από τον αρχικό αλγόριθμο. Οι *learnable* παράμετροι επισημαίνονται με μπλε [40].

Στον αλγόριθμο ISTA, για παράδειγμα, κάθε επανάληψη περιλαμβάνει δύο στάδια: ένα βήμα καθόδου της κλίσης για τον όρο συνέπειας στα δεδομένα και ένα βήμα εγγύτητας με τελεστή μαλακού κατωφλίου, που επιβάλλει αραιότητα στον κατάλληλο χώρο αναπαράστασης. Για $R(x) = \lambda \|Wx\|_1$ ο ISTA επαναληπτικός τύπος γράφεται [40]

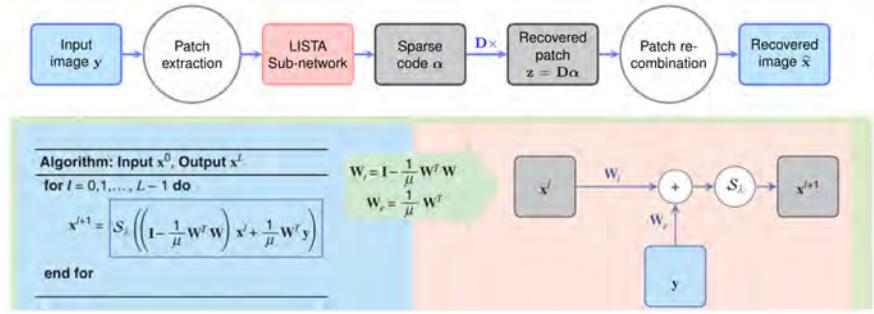
$$x^{(l+1)} = W^* \mathcal{S}_{\lambda/L} \left(W(x^{(l)} - A^T(Ax^{(l)} - y)) \right),$$

όπου W είναι ο (π.χ. wavelet) μετασχηματισμός, W^* ο αντίστροφος και \mathcal{S}_τ ο soft-thresholding operator. Κάθε επανάληψη περιέχει έναν όρο συνέπειας δεδομένων (γραμμικό) και ένα prior/denoising βήμα (μη-γραμμικό).

Στον αλγόριθμο LISTA, (Learned ISTA) [40] αντικαθιστούμε τους σταθερούς γραμμικούς τελεστές με εκπαιδεύσιμες μήτρες βαρών και τα κατώφλια με εκπαιδεύσιμες παραμέτρους κατωφλίωσης. Η ενημέρωση μπορεί να γραφεί

$$x^{(l+1)} = \mathcal{S}_{\lambda_l} (W_t^{(l)} x^{(l)} + W_e^{(l)} y),$$

όπου οι πίνακες $W_t^{(l)}, W_e^{(l)}$ και τα λ_l μαθαίνονται μέσω οπισθοδιάδοσης, ώστε λίγα επίπεδα να προσεγγίζουν την ακρίβεια πολλών επαναλήψεων ISTA σε πολύ μικρότερο αριθμό βημάτων. Το LISTA αναδεικνύει την κεντρική ιδέα: το ξετυλιγμένο δίκτυο διατηρεί τη λογική του αρχικού αλγορίθμου αλλά ορίζεται ως μάθηση από δεδομένα [40].



Σχήμα 3.10

LISTA (Learned ISTA)—ένας εκπαιδευόμενος νευρωνικός αλγόριθμος που μημείται τα βήματα του *ISTA*, αλλά αντικαθιστά τους σταθερούς γραμμικούς μετασχηματισμούς/κατώφλια με εκπαιδεύσιμες παραμέτρους.[40]

Στο πλαίσιο της MRI με Καρτεσιανή δειγματοληψία γράφουμε

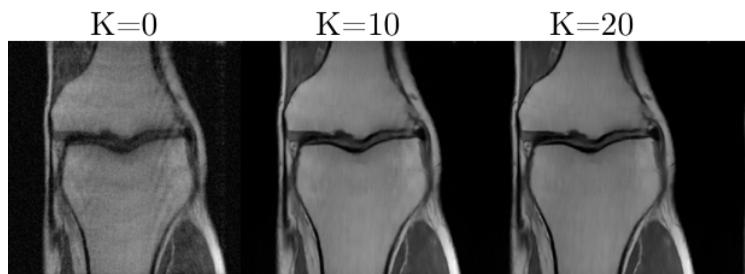
$$A = M \mathcal{F} C,$$

όπου M είναι η μάσκα δειγματοληψίας στον χώρο k , \mathcal{F} ο διακριτός μετασχηματισμός Fourier (και ο αντίστροφός του \mathcal{F}^{-1}), και C οι χάρτες ευαισθησίας των πηγών (σημειακός πολλαπλασιασμός). Η προσέγγιση του ξετυλίγματος είναι φυσική εδώ: το βήμα συνέπειας ως προς τα δεδομένα εφαρμόζεται άμεσα στον χώρο k , ενώ οι φυσικές πράξεις (Γρήγορος Μετασχηματισμός Fourier/αντίστροφος) παραμένουν μη εκπαιδεύσιμες.

Ξεκινάμε από την κλασική δομή εμπρόσθιας τροφοδότησης ενός στρώματος. Σε ένα τυπικό δίκτυο, κάθε στρώμα υλοποιεί τον μετασχηματισμό

$$\tilde{x}^{(l)} = D_{\theta_l}(x^{(l)}), \quad x^{(l+1)} = \tilde{x}^{(l)} + A^\dagger W_l(y - A \tilde{x}^{(l)}),$$

όπου D_{θ_l} είναι εκπαιδεύσιμος αποθηκευτικής στον χώρο εικόνας, A^\dagger ο συζυγής τελεστής του A (που ενσωματώνει τους χάρτες ευαισθησίας και τον αντίστροφο μετασχηματισμό Fourier) και W_l συντελεστής βήματος, συνήθως είτε βαθμωτός είτε εκπαιδεύσιμος διαγώνιος τελεστής. Η μορφή αυτή ενσωματώνει άμεσα τη φυσική του συστήματος.



Σχήμα 3.11 Παράδειγμα έπειτα από 20 επαναλήψεις της μεθόδου ξεδιπλώματος[41]

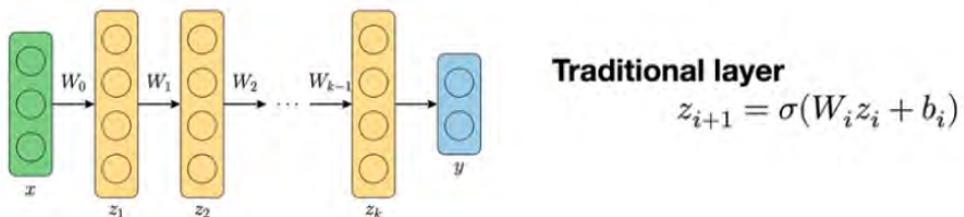
Στην ανακατασκευή MRI, το ξετύλιγμα μοντελοποιείται ως πεπερασμένη ακολουθία K επαναλήψεων που εναλλάσσουν βήμα συνέπειας ως προς τα δεδομένα, επιβολή των μετρήσεων στον χώρο k και βήμα κανονικοποίησης/αποθορυβοποίησης με εκπαιδεύσιμες παραμέτρους. Ο τερματισμός γίνεται με κριτήριο σύγχλισης:

$$|x^{(l+1)} - x^{(l)}| < \varepsilon.$$

Συνοπτικά, η μέθοδος ξετυλιγμένου αλγορίθμου, μετατρέπει έναν κλασικό επαναληπτικό αλγόριθμο (π.χ. ISTA/ADMM) σε βαθύ δίκτυο πεπερασμένου βάθους, όπου κάθε επανάληψη γίνεται επίπεδο με εκπαιδεύσιμες παραμέτρους. Με αυτόν τον τρόπο συνδυάζει γνώση μοντέλου με μάθηση από δεδομένα. Εφαρμόζεται σε αντίστροφα προβλήματα (MRI, CT, αποθορυβοποίηση, αποθόλωση) και λειτουργεί ως γέφυρα ανάμεσα στις μεθόδους βελτιστοποίησης και τα σύγχρονα deep networks.

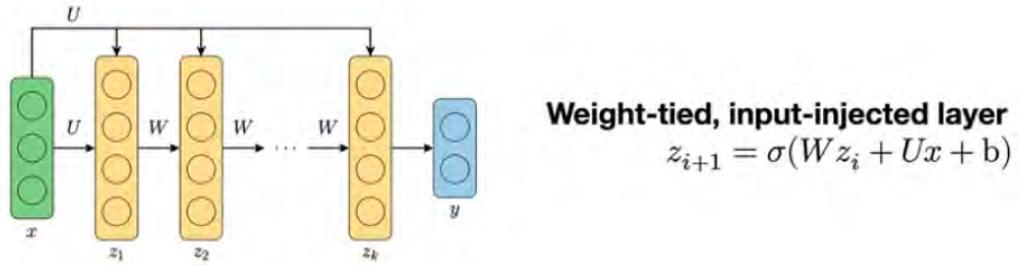
3.3 Η ιδέα των μοντέλων Ισορροπίας DEQ

Ξεκινάμε από την παραδοσιακή εμπρόσθια δομή ενός στρώματος δικτύου. Σε ένα τυπικό δίκτυο κάθε στρώμα μετασχηματίζει το διάνυσμα ενεργοποιήσεων z_i όπως στην εικόνα 3.12



Σχήμα 3.12 Παραδοσιακή περιγραφή [42]

Η συνήθης πρακτική είναι η χρήση διαφορετικών βαρών W_i σε κάθε επίπεδο, αυξάνοντας το βάθος μέσω διαδοχικών διαφορετικών δομικών μονάδων [42]. Δηλαδή, το βάθος ενός νευρωνικού δικτύου αντικατοπτρίζεται στο σύνολο των διαφορετικών W_i και των μη γραμμικοτήτων σ . Ωστόσο η ιδέα είναι να εφαρμόσουμε μια εναλλακτική προσέγγιση: τα βάρη ορίζονται ως συζευγμένα βάρη W (weight tied) και εισάγουμε το αρχικό σήμα σε κάθε επανάληψη με την ίδεα της σύζευξης βαρών. Συγκεκριμένα ορίζουμε το επαναληπτικό block της εικόνας 3.13:



Σχήμα 3.13 Περιγραφή με σύζευξη βαρών [42]

όπου σ είναι στοιχειώδης μη γραμμικότητα, W_i οι μήτρες βαρών του στρώματος και b_i οι μετατοπίσεις.

Ορίζουμε κοινή μήτρα W για όλα τα στρώματα και επανεισάγουμε την αρχική είσοδο σε κάθε επανάληψη. Συγκεκριμένα, το επαναληπτικό δομικό στοιχείο ορίζεται ως

$$z^{(k+1)} = \sigma(Wz^{(k)} + Ux + b),$$

όπου η ίδια μήτρα W εφαρμόζεται σε όλα τα επίπεδα (σύζευξη βαρών), το U είναι μήτρα που εισάγει την αρχική είσοδο x σε κάθε βήμα και το b σταθερή μετατόπιση. Παρότι η σύζευξη βαρών φαίνεται περιοριστική, ένα δίκτυο πρόσθιας τροφοδότησης με διαφορετικά βάρη μπορεί να προσεγγιστεί από κατάλληλα σχεδιασμένο δίκτυο με σύζευξη βαρών[42]. Στην πράξη, τα δίκτυα με σύζευξη βαρών είναι ιδιαίτερα αποδοτικά: μειώνουν σημαντικά τον αριθμό παραμέτρων και βελτιώνουν τη σταθερότητα της εκπαίδευσης [42].

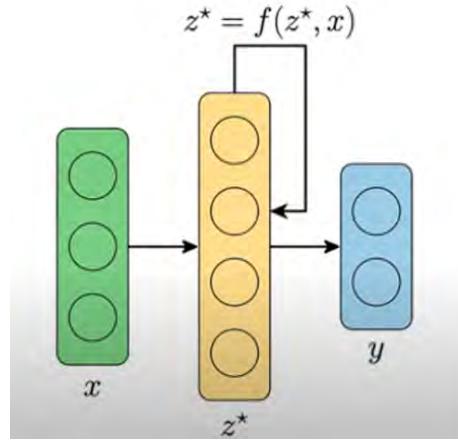
Παρατηρούμε τώρα ότι η παραπάνω επαναλαμβανόμενη εφαρμογή του ίδιου μη γραμμικού στρώματος δημιουργεί μια δυναμική επανάληψης:

$$z \mapsto f_\theta(z; x) \quad \text{με} \quad f_\theta(z; x) = \sigma(Wz + Ux + b)$$

Αν εφαρμόσουμε το ίδιο block πολλές φορές, θεωρητικά έως το άπειρο[41] , τότε, υπό κατάλληλες συνθήκες , η ακολουθία $\{z_k\}$ μπορεί να συγκλίνει σε ένα σταθερό διάνυσμα z^* που ικανοποιεί την εξίσωση Ισορροπίας [42]

$$z^* = f_\theta(z^*; x).$$

Αυτό το διάνυσμα z^* αντιστοιχεί θεωρητικά στο όριο ενός δικτύου απέριου βάθους όπου το ίδιο block επαναλαμβάνεται επ' άπειρον. Η οπτική της μεθόδου Ισορροπίας(equilibrium επομένως είναι ότι: μπορούμε να δούμε το πρόβλημα ως αντί να κατασκευάσουμε και να εκπαιδεύσουμε ρητά ένα πολύ βαθύ δίκτυο με πολλαπλά διακριτά στρώματα, επιλύουμε απευθείας για το σημείο Ισορροπίας z^* της επαναληπτικής απεικόνισης και το χρησιμοποιούμε ως εσωτερική αναπαράσταση του συστήματος[41].



Σχήμα 3.14 Αναπαράσταση σχεδίασης στην ιδέα σταθερού σημείου [42]

Πρακτική σημασία. (α) Μειώνεται το μνημονικό αποτύπωμα κατά την εκπαίδευση, επειδή δεν απαιτείται αποθήκευση όλων των ενδιάμεσων ενεργοποιήσεων $\{z_k\}$: οι ενημερώσεις των παραμέτρων μπορούν να εξαχθούν απευθείας από το σταθερό σημείο z^* .

(β) Αναδιαμορφώνεται η διαδικασία εκπαίδευσης: αντί για τη συμβατική οπισθοδιάδοση με ξετύλιγμα ολόκληρου του δικτύου, χρησιμοποιείται έμμεση παραγώγιση για τον υπολογισμό των κλίσεων ως προς τις παραμέτρους θ του δομικού block.

Συνοπτικά, ορίζουμε τη συνάρτηση f_θ (περιλαμβάνοντας W, U, b και τυχόν παραμέτρους της σ) και θεωρούμε ένα κόστος $\mathcal{L}(z^*)$ που εξαρτάται από την τελική κατάσταση z^* . Τότε οι παράγωγοι ως προς θ δίνονται μέσω της εξίσωσης που προκύπτει από το θεώρημα της έμμεσης παραγώγισης. Έτσι ο βαθμωτός υπολογίζεται ως

$$\frac{\partial \mathcal{L}}{\partial \theta} = -v^\top \frac{\partial f_\theta}{\partial \theta},$$

όπου το διάνυσμα v είναι η λύση του γραμμικού συστήματος

$$(I - \frac{\partial f_\theta}{\partial z})^\top v = (\frac{\partial \mathcal{L}}{\partial z})^\top.$$

Έτσι η ιδέα του DEQ είναι: αντί να κατασκευάσουμε και να εκπαιδεύσουμε ένα τεράστιο, ντετερμινιστικό δίκτυο βάθους K και να κάνουμε οπισθοδιάδοση σε όλα τα ενδιάμεσα στάδια, βρίσκουμε κατευθείαν ένα σταθερό σημείο και χρησιμοποιούμε έμμεση παραγώγιση[41]. Έτσι εξικονομούμε μνήμη και μαθαίνουμε το τελικό σταθερό σημείο.

3.4 Οι λειτουργίες του μοντέλου Ισορροπίας :

3.4.1 Σταθερό σημείο μοντέλου Ισορροπίας με επαναληπτικό χάρτη

Ξεκινάμε την ανάλυση κρατώντας ως κεντρική εικόνα ότι διαθέτουμε ένα παραμετρικό block R_θ και έναν επαναληπτικό χάρτη f_θ που εφαρμόζεται στην ίδια κατάσταση (state) με βασική μορφή ως

$$x^* = f_\theta(x^*; y)[41],$$

όπου το y συμβολίζει τις μετρήσεις και η f_θ από τις παραμέτρους θ , περιλαμβάνει τόσο τον όρο μαθησης όσο και το βήμα.

Διάσπαση σε ορο συνέπειας δεδομένων και όρο εκπαίδευσης

Σε επίπεδο βήματος, κάθε εφαρμογή του χάρτη $x \mapsto f_\theta(x; y)$ περιλαμβάνει δύο σαφώς διαχωρίσιμες ενέργειες. Πρώτον, τη διόρθωση συμβατότητας με τα δεδομένα: υπολογίζουμε το $r = y - Ax$, το επαναφέρουμε στο χώρο εικόνας μέσω του $A^\top r$ και σχηματίζουμε προσωρινή εκτίμηση

$$v = x + \eta A^\top (y - Ax).$$

Το βήμα αυτό ενσωματώνει τις παρατηρήσεις στις μετρήσιμες συνιστώσες της εικόνας και δεν επηρεάζει άμεσα το nullspace του A . Δεύτερον, ο όρος εκπαίδευσης R_θ εφαρμόζεται στην προσωρινή διόρθωση v και παράγει νέα, εξομαλυμένη εκτίμηση

$$x_{\text{next}} = R_\theta(v).$$

Η επανάληψη αυτής της δυάδας οδηγεί την εκτίμηση σε λύση που εξισορροπεί τον ορο συνέπειας δεδομένων [41] με στατιστικά ρεαλιστικές συμπληρώσεις.

Εμπρόσθιος και οπίσθιος επιλυτής

Ο ρόλος του επιλυτή είναι διττός και καινοριστικός για απόδοση και ευστάθεια.

Εμπρόσθιος επιλυτής (forward solver). Δοθέντων (y, θ) αναζητούμε σταθερό σημείο x^* της απεικόνισης $f_\theta(\cdot; y)$,

$$x^* = f_\theta(x^*; y).$$

Η απλούστερη επαναληπτική μέθοδος είναι

$$x^{k+1} = f_\theta(x^k; y),$$

με κριτήριο τερματισμού, π.χ. $\|x^{k+1} - x^k\| \leq \varepsilon$ ή $\|r^k\| \leq \varepsilon$. Όταν η σύγκλιση είναι αργή ή ασταθής, εφαρμόζονται επιταχύνσεις σταθερού σημείου όπως η Anderson acceleration ή η μέθοδος Broyden, οι οποίες αξιοποιούν ιστορικές επαναλήψεις και υπολείμματα για ταχύτερη και πιο σταθερή σύγκλιση[40].

Οπίσθιος επιλυτής (backward solver) Επειδή η απώλεια ℓ εξαρτάται από τις παραμέτρους μόνο μέσω του $x^*(\theta; y)$, χρησιμοποιούμε έμεση παραγώγιση στη συνθήκη σταθερού σημείου. Θέτουμε

$$F(x, \theta; y) := x - f_\theta(x; y) \quad \text{και} \quad F(x^*, \theta; y) = 0.$$

Παραγωγίζοντας και αναδιατάσσοντας, προκύπτει ότι το διάνυσμα v λύει το γραμμικό σύστημα

$$(I - \partial_x f_\theta(x^*; y))^\top v = \left(\frac{\partial \ell}{\partial x^*} \right)^\top,$$

και στη συνέχεια η κλίση ως προς τις παραμέτρους δίνεται από

$$\left[\frac{\partial \ell}{\partial \theta} = (\partial_\theta f_\theta(x^*; y))^\top v \right].$$

Στην πράξη δεν σχηματίζουμε ρητά τον Jacobian. Αντ' αυτού, χρησιμοποιούνται επαναληπτικοί επιλυτές γραμμικών συστημάτων (π.χ. conjugate gradients όταν ο τελεστής είναι συμμετρικός θετικά ορισμένος), οι οποίοι βασίζονται αποκλειστικά σε γινόμενα τύπου

$$w \mapsto (\partial_x f_\theta(x^*; y))w, \quad w \mapsto (\partial_\theta f_\theta(x^*; y))^\top w,$$

τα οποία υπολογίζονται αποδοτικά μέσω automatic differentiation (π.χ. JVP/VJP).

Ο επιλυτής, σε κάθε εσωτερικό βήμα, ζητά τα παραπάνω γινόμενα, λαμβάνει το αποτέλεσμα από την αυτόματη διαφοροποίηση και προχωρά μέχρι να βρεθεί το v με την απαιτούμενη ακρίβεια.

Ρύθμιση ακρίβειας Η προσεκτική επιλογή ανεκτικοτήτων (απόλυτης/σχετικής) και μέγιστου αριθμού επαναλήψεων, τόσο για τον εμπρόσθιο όσο και για τον οπίσθιο επιλυτή, είναι ουσιώδης. Η αριθμητική ακρίβεια της επίλυσης επηρεάζει άμεσα την ποιότητα του x^* και την ορθότητα των κλίσεων.

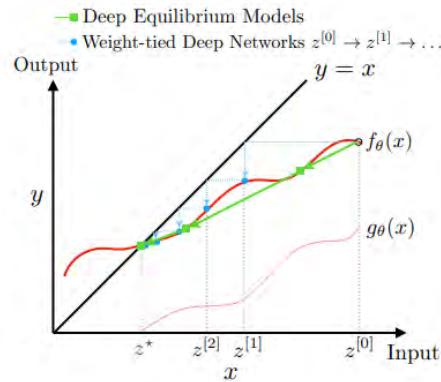
Διαχείριση μνήμης και ροή εκπαίδευσης Κατά την εκτέλεση ενός βήματος εκπαίδευσης διατηρούμε το τελικό x^* για τον υπολογισμό της απώλειας $\ell(x^*; \theta)$ και το διάνυσμα παραγώγου $\partial \ell / \partial x^*$, το οποίο τροφοδοτεί τον οπίσθιο επιλυτή (για την έμμεση διαφοροποίηση).

Στη μέθοδο επιτάχυνσης Anderson απαιτείται μία μικρή ενδιάμεση μνήμη με τα τελευταία m υπολείμματα $r^k = f_\theta(x^k; y) - x^k$ ή/και τις τελευταίες επαναληπτικές προσεγγίσεις x^k [43]. Το αποτύπωμα μνήμης είναι έτσι πολύ μικρότερο από την αποθήκευση ολόκληρης της ακολουθίας (x^1, \dots, x^K) που θα απαιτούσε εκπαίδευση με της επαναληπτικής διαδικασίας.

Μόνιμα αποθηκεύονται οι ενημερωμένες παράμετροι θ και η κατάσταση του αλγορίθμου βελτιστοποίησης (π.χ. όροι ορμής, εκτυπήσεις ροπών), καθώς και, προαιρετικά, στιγμιότυπα καταγραφής για παρακολούθηση της πορείας.

Στη φάση αξιολόγησης (χωρίς εκπαίδευση) επιστρέφεται μόνο η λύση x^* , δεν πραγματοποιούνται ενημερώσεις παραμέτρων ούτε υπολογισμοί κλίσεων.

Πού και πότε μαθαίνει το δίκτυο Η προσαρμογή των παραμέτρων θ πραγματοποιείται μετά την επίλυση της συνθήκης Ισορροπίας (σταθερού σημείου) για κάθε δείγμα ή παρτίδα. Συγκεκριμένα: ο εμπρόσθιος επιλυτής υπολογίζει τη λύση $x^*(\theta; y)$, η συνάρτηση απώλειας μετρά την απόκλιση από την εικόνα αναφοράς, ο οπίσθιος επιλυτής υπολογίζει την ευαισθησία μέσω έμμεσης διαφοροποίησης της συνθήκης Ισορροπίας, και ο αλγόριθμος βελτιστοποίησης εφαρμόζει την ενημέρωση στις θ . Η διαδικασία επαναλαμβάνεται σε όλο το σύνολο δεδομένων. Ο παραμετρικός τελεστής R_θ «συμπληρώνει» τα μη παρατηρήσιμα συστατικά στον πυρήνα του τελεστή μέτρησης, ενώ ο όρος συνέπειας δεδομένων επιβάλλει συμμόρφωση προς τις μετρήσεις. Ο συνδυασμός τους οδηγεί σε λύσεις Ισορροπίας με φθίνουσα απώλεια έως τη σύγκλιση σε σταθερό σημείο [43].



Σχήμα 3.15 Απλή δυσδιάστατη αναπαράσταση επίλυσης για ένα σημείο Ισορροπίας[43]

Τι παραμένει μετά από κάθε βήμα και παράγοντες ποιότητας Μετά από κάθε βήμα εκπαίδευσης παραμένουν οι επικαιροποιημένες παράμετροι θ και η κατάσταση του αλγορίθμου βελτιστοποίησης [43]. Η ποιότητα της ανακατασκευής εξαρτάται από: (i) την αντιπροσωπευτικότητα του συνόλου δεδομένων, (ii) τη χωρητικότητα και την αρχιτεκτονική του R_θ , και (iii) την ακρίβεια με την οποία ο εμπρόσθιος επιλυτής προσεγγίζει το x^* και ο οπίσθιος επιλυτής υπολογίζει το v . Πρακτικές επιλογές περιλαμβάνουν θερμή εκκίνηση του εμπρόσθιου επιλυτή, προ-εκπαίδευση του R_θ ως αποθορυβοποιητή, προσεκτική ρύθμιση των ανοχών στους εμπρόσθιους/οπίσθιους επιλυτές, πιθανή προρρύθμιση (preconditioning) στον οπίσθιο επιλυτή, και χρήση κανονικοποιητών όπως η φασματική κανονικοποίηση και η αποκοπή κλίσεων για ευστάθεια και βελτιωμένη γενίκευση.

3.4.2 DEQ οικογένειες επιλυτών και θεωρητικά κριτήρια

(1) Επιλογές λυτών και υπολογιστικά trade-offs Εμείς εδώ, αναφέραμε την εφαρμογή Anderson στο forward και συγκεκριμένων λυτών στο backward, αλλά στο πλαίσιο equilibrium υπάρχουν πολλαπλές εναλλακτικές με διακριτά μαθηματικά trade-offs. Κάθε επιλογή έχει σαφές μαθηματικό κίνητρο και διαφορετικές απαιτήσεις σε ταχύτητα, μνήμη και preconditioning.

Για παράδειγμα ενδεικτικά μια άλλη ενναλλακτική quasi-Newton μέθοδος είναι το Broyden [43], το οποίο κατασκευάζει επαναληπτικά μια προσέγγιση B_k του Jacobian J ή του αντιστρόφου J^{-1} με rank-one ενημερώσεις βάσει των διαφορών $s_k = x_{k+1} - x_k$ και $y_k = f(x_{k+1}) - f(x_k)$. Μια απλή μορφή ενημέρωσης είναι

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)s_k^\top}{s_k^\top s_k},$$

οπότε αποκτούμε Newton-λικε βήματα χωρίς ποτέ να σχηματίζουμε ολόκληρο τον J . Στην πράξη, το Broyden ταιριάζει όταν ο J μεταβάλλεται ομαλά και επιδιώκουμε ισχυρότερη τοπική σύγκλιση από την Anderson, με μικρό αποτύπωμα μνήμης για λίγες ενημερώσεις ιστορικού[43].

(2) Συνθήκες ύπαρξης/μοναδικότητας και τοπικής σύγκλισης. Τέλος, αξίζει να αναφέρουμε ότι το σταθερό σημείο μπορεί να μην υπάρχει ή να μην είναι μοναδικό, και γι' αυτό στη θεωρία εισάγονται ρητές συνθήκες και θεώρηματα που εξασφαλίζουν την ύπαρξη, την μοναδικότητα και την τοπική σύγκλιση του συστήματος Ισορροπίας .Συγκεκριμένα, το επαναληπτικό $x^{(k+1)} = f_\theta(x^{(k)}; y)$ συγκλίνει σε μοναδικό x^* εάν ο χάρτης $f_\theta(\cdot; y)$ είναι συσταλτικός (contractive), δηλαδή υπάρχει $0 \leq c < 1$ τέτοιο ώστε

$$\|f_\theta(x; y) - f_\theta(\tilde{x}; y)\| \leq c \|x - \tilde{x}\| \quad \text{για όλα } x, \tilde{x}. [41]$$

Ουσιαστικά σημαίνει ότι η συνάρτηση συρρικνώνει τις αποστάσεις μεταξύ δύο οποιωνδήποτε σημείων: αν πάρουμε δύο διαφορετικά σημεία x και y , μετά την εφαρμογή της συνάρτησης f το νέο τους διάστημα θα είναι μικρότερο κατά τουλάχιστον έναν σταθερό παράγοντα c με $0 \leq c < 1$.

Συσταλτικότητα και ρυθμίσεις στα μοντέλα βαθιάς Ισορροπίας Στα μοντέλα βαθιάς Ισορροπίας η συσταλτικότητα εξαρτάται τόσο από τη συμπεριφορά του εκπαιδευόμενου τελεστή R_θ όσο και από το φάσμα του γραμμικού μέρους, συνήθως του $A^\top A$ στη Μαγνητική Τομογραφία. Μια φυσική και τεχνικά διαχειρίσιμη υπόθεση είναι ότι ο μετασχηματισμός $R_\theta - I$ είναι Lipschitz με σταθερά $\alpha > 0$, δηλαδή

$$\|(R_\theta - I)(x) - (R_\theta - I)(y)\| \leq \alpha \|x - y\|, \quad \alpha > 0, [41].$$

Πρακτικά, αυτό εκφράζει την απαίτηση ο εκπαιδευόμενος τελεστής να μην επιφέρει ανεξέλεγκτες μεταβολές.

Ποσοτικές συνθήκες συσταλτικότητας Για πιο συγκεκριμένα αποτελέσματα, οι συνθήκες εκφράζονται μέσω του ακραίου φάσματος $L = \lambda_{\max}(A^\top A)$ και της ελάχιστης ιδιοτυπής $\lambda_{\min}(A^\top A) > 0$ (θετική οριστικότητα), σε συνδυασμό με την τιμή της α . Για τη βαθμιδική εκδοχή (DE-Grad) επαρκεί ενδεικτικά το μέγεθος βήματος η να ικανοποιεί

$$\eta < \frac{1}{L+1},$$

ώστε ο επαγόμενος χάρτης να είναι συσταλτικός [41]. Στην εκδοχή χάρτης Βαθιάς Ισορροπίας με Μέθοδο Εγγύς Κλίσης (DE-Prox), η ανάλυση παρέχει μη κενό διάστημα επιτρεπτών τιμών για το η που εξαρτάται ρητά από L και α : ενδεικτικά, υπό τις προϋποθέσεις της ανάλυσης, εμφανίζεται το διάστημα

$$\frac{1}{L(1+1/\alpha)} < \eta < \frac{2}{L},$$

υπό την προϋπόθεση ότι η α είναι αρκετά μικρή ώστε το διάστημα να μην είναι κενό [43]. Τέτοιες ανισότητες είναι επαρκείς για συσταλτικότητα (άρα και για ύπαρξη/μοναδικότητα σταθερού σημείου), όχι όμως αναγκαίες.

Κεφάλαιο 4

Το Σύνολο Δεδομένων

4.1 Ο χώρος k (k-space)

Ο χώρος k είναι ο μαθηματικός χώρος στον οποίο το MR σύστημα αποθηκεύει τις μετρήσεις πριν την μετατροπή τους σε εικόνα, και στην ουσία είναι ο Fourier-μετασχηματισμός της εικόνας. Αν η διανυσματική (σύνθετη) κατανομή της εγκάρσιας μαγνήτισης στο αντικείμενο είναι $x(\mathbf{r})$ (όπου $\mathbf{r} = (x, y)$ στον 2D απλό περιστατικό), τότε ο k -χώρος είναι η συνάρτηση

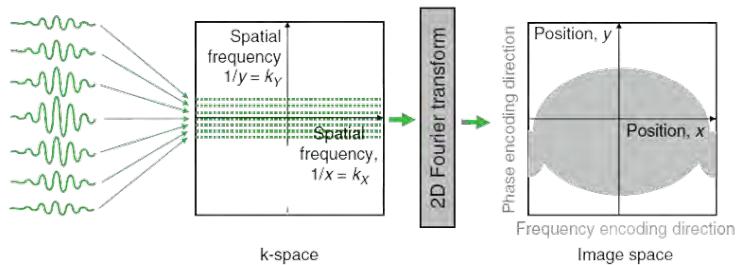
$$K(\mathbf{k}) = \int x(\mathbf{r}) e^{-i2\pi\mathbf{k}\cdot\mathbf{r}} d\mathbf{r},$$

και η εικόνα ανακτάται με τον αντίστροφο Fourier

$$x(\mathbf{r}) = \int K(\mathbf{k}) e^{i2\pi\mathbf{k}\cdot\mathbf{r}} d\mathbf{k}.$$

Αυτό το ένα-προς-ένα ζευγάρι [46] (εικόνας \leftrightarrow χώρου k) είναι η μαθηματική βάση της MRI: ο σαρωτής μετράει σημεία στο K , και ένα πλήρες σύνολο κατάλληλων σημείων επιτρέπει την ακριβή ανακατασκευή x με την IFFT.

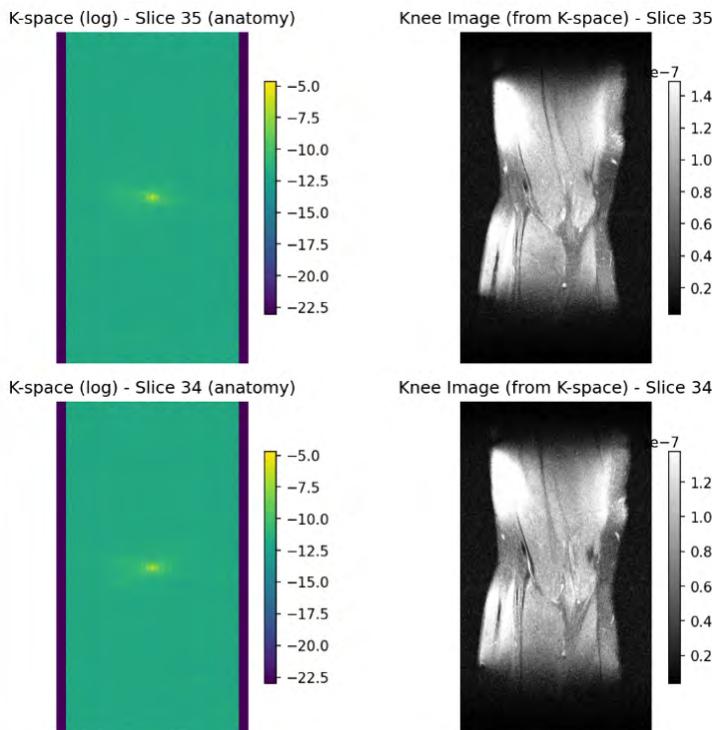
Επομένως, ο χώρος k ορίζεται ως το διάνυσμα των χωρικών συχνοτήτων (k_x, k_y) που αντιστοιχούν στις αρμονικές συνιστώσες της χωρικής κατανομής της έντασης[45]. Η εικόνα στο χώρο $x - y$ μπορεί να θεωρηθεί ως το αποτέλεσμα της υπέρθεσης όλων αυτών των αρμονικών συνιστωσών με συγκεκριμένα πλάτη και φάσεις, και η μαθηματική σχέση μεταξύ τους δίνεται από τον μετασχηματισμό Fourier και τον αντίστροφό του.



Σχήμα 4.1 Καθε σημείο του k -χώρου αντιστοιχησμένο στον χώρο εικόνας [45]

Τι αποθηκεύει πρακτικά το χώρου k των .h5 και γιατί είναι complex: Κάθε θέση (k_x, k_y) στον πίνακα του χώρου k περιέχει πληροφορία για μία χωρική συχνότητα και τη σχετική φάση αυτής της συνιστώσας[47]. Τα MR σήματα είναι σύνθετα (πραγματικό + φανταστικό ή magnitude + phase) επειδή το καταγεγραμμένο σήμα έχει και πλάτος και φάση. Η φάση μεταφέρει κρίσιμη πληροφορία χωρικής τοποθέτησης και χρονικής μετατόπισης. Έτσι τα στοιχεία του χώρου k αποθηκεύεται ως πίνακας σύνθετων τιμών (π.χ. complex64). Στα πρακτικά datasets (π.χ. fastMRI single-coil knee) κάθε τομή αντιπροσωπεύεται ως 2D complex matrix με διαστάσεις περίπου $H \times W$ (phase-encode \times frequency-encode), και τα αρχεία .h5 κρατούν αυτές τις φέτες ως [slices, H, W] arrays δηλαδή ένα 3D tensor από 2D k-spaces[47].

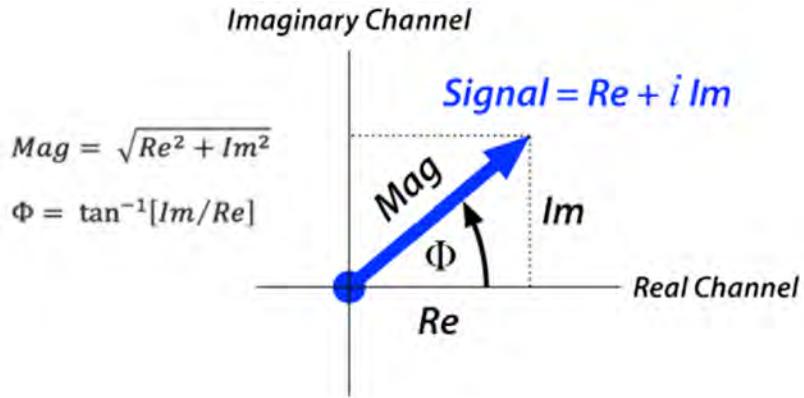
Κέντρο vs άκρα του χώρου k: τι σημαίνει κάθε περιοχή. Το κέντρο του χώρου k (σημεία κοντά στο $k = 0$) περιέχει τις χαμηλές χωρικές συχνότητες. Αυτές ορίζουν τη γενική ένταση (contrast) [46] και την μεταβολή φωτεινότητας στην εικόνα, έχουν μεγάλη επίδραση στο SNR και στο συνολικό contrast. Τα άκρα του χώρου k περιέχουν υψηλές χωρικές συχνότητες: αυτές κωδικοποιούν λεπτομέρειες, όρια και αιχμές.[46] Επομένως, κόβοντας ή υποδειγματοληπτώντας τις εξωτερικές περιοχές η εικόνα χάνει λεπτομέρεια (θολώνει), ενώ απώλεια κεντρικών δειγμάτων αλλάζει το contrast και μειώνει SNR.



Σχήμα 4.2 Παράδειγμα από το σύνολο δεδομένων fastMRI [47]: τομές του χώρου k σε λογαριθμική κλίμακα (αριστερά) και οι αντίστοιχες ανακατασκευές με IFFT στον χώρο εικόνας (δεξιά)

Επίδραση ενός σημείου χώρου k στην εικόνα: Κάθε σημείο του χώρου k επηρεάζει ολόκληρη την εικόνα, ένα συντελεστής δεν αντιστοιχεί σε ένα pixel. Ένα σφάλμα ή θόρυβος σε ένα χώρου k στοιχείο θα διαχυθεί σε όλη την εικόνα μετά την IFFT.[46] Γι' αυτό τα σχήματα θορύβου/artefacts εξαρτώνται έντονα από το ποια σημεία λείπουν ή είναι αδύναμα.

Φάση (phase) vs. magnitude: γιατί η εικόνα είναι συχνά complex. Η IFFT του χώρου k παράγει σύνθετη εικόνα $x(\mathbf{r})$ με πραγματικό και φανταστικό μέρος. Το τελικό αποτέλεσμα που συνήθως δείχνουμε είναι το magnitude $|x|$ επειδή είναι ανεξάρτητο από ομοιόμορφη φασική μετατόπιση και πιο εύκολα ερμηνεύσιμο. Όμως η φάση φέρει πληροφορία (π.χ. phase-contrast MRI, διορθώσεις coil, φασικά artefacts) και σύγχρονες μέθοδοι ανακατασκευής πολλές φορές δουλεύουν απευθείας με το σύνθετο σήμα[53].



Σχήμα 4.3 Αναπαράσταση του σήματος MR στο μηδικό επίπεδο: πραγματικό και φανταστικό μέρος (ή μέτρο και φάση) [45].

Συνοπτικά χρήσιμα συμπεράσματα για reconstruction / unrolling. Οι ανακατασκευές καθιστίζονται από το πώς γεμίζεται ο χώρος, ποια σημεία λείπουν (mask), τις ευαισθησίες των coils (αν multi-coil) και το θόρυβο. Για unrolled networks / learned reconstructions αυτό σημαίνει πρακτικά δύο πράγματα:

1. Το ορος συνέπειας (δεδομένων). ζονσιστενςψ βήμα πρέπει να δουλεύει με το σωστό forward model A που περιλαμβάνει FFT/IFFT, μάσκα M και (αν multi-coil) τους coil maps — αυτές οι πράξεις είναι φυσικές και δεν χρειάζεται να τις μάθει το δίκτυο.
2. Τα priors/denoisers που μαθαίνει το δίκτυο πρέπει να χειρίζονται σωστά τη σύνθετη φύση του σήματος (image-domain, k -domain ή συνδυασμό).

4.2 Μάσκες υποδειγματοληψίας και ρυθμίσεις επιτάχυνσης (π.χ. 4 \times / 8 \times)

Η μάσκα στον k χώρο είναι ένας δυαδικός (ή δυαδικού τύπου) δείκτης ο οποίος λέει σε ποιες θέσεις του διακριτού k -χώρου έχουμε πραγματικά μετρήσεις και ποιες θέσεις ψευδούμε μη-μετρημένες (συνήθως μηδενισμένες). Στο πρακτικό επίπεδο αυτό αντιστοιχεί σε ένα πολλαπλασιασμό του πλήρους k χώρου με έναν πίνακα M (mask), και η μη-ολοκλήρωση αυτού του πίνακα (δηλαδή το να λείπουν γραμμές/σημεία) μειώνει το χρόνο σάρωσης, άρα επιταχύνει το πρωτόκολλο. Ένα τέτοιο undersampling κάνει το forward operator A (στο μοντέλο $y = Ax + \varepsilon$) μη-πλήρες και στο αντίστροφο πρόβλημα : υπάρχουν άπειρες εικόνες που συμφωνούν με τα μειωμένα δεδομένα, εκτός αν εισάγουμε επιπλέον πληροφορία.

Τα πειράματα στην εργασία μας βασίζονται στο "Deep Equilibrium Architectures for Inverse Problems in Imaging" που χρησιμοποιούν τυπικά 4 \times και 8 \times επιταχύνσεις ώστε να ποσοτικοποιήσουν την επίδραση της υποδειγματοληψίας στην ποιότητα ανακατασκευής [41]. Ειδικότερα, οι μάσκες που χρησιμοποιούνται στα πειράματα ανακατασκευής MRI βασίζονται σε ένα καρτεσιανό πρότυπο δειγματοληψίας, όπως στην τυπική ρύθμιση fastMRI[47]. Όταν λέμε 4 \times ή 8 \times , αυτό σημαίνει ότι ο αριθμός των μετρήσεων είναι αντίστοιχα περίπου 1/4 ή 1/8 του πλήρους. Τόσο για το 4 \times όσο και για το 8 \times , το κεντρικό 4% των συχνοτήτων δειγματοληπτείται πλήρως, ενώ οι υπόλοιπες συχνότητες δειγματοληπτούνται σύμφωνα με μια Gaussian κατανομή με κέντρο τη συχνότητα 0 και $\sigma = 1$.

Τι κάνει η μάσκα στο image domain Μαθηματικά, η zero-filled επανασύνθεση γράφεται

$$\hat{x}_{\text{zf}} = \mathcal{F}^{-1}\{M \cdot \mathcal{F}\{x\}\} = x * \text{PSF}, \quad \text{PSF} = \mathcal{F}^{-1}\{M\}.$$

Δηλαδή το αποτέλεσμα ισοδυναμεί με συνέλιξη της αληθινής εικόνας με το point-spread-function που προκύπτει από την IFFT της μάσκας. Η μορφή του PSF καθορίζει αν το αποτέλεσμα θα εμφανίσει συστηματικό aliasing/folding (όταν η υποδειγματοληψία είναι συστηματική, π.χ. κάθε R -στή γραμμή) ή διασκορπισμένο aliasing (όταν η υποδειγματοληψία είναι τυχαιοποιημένη/variable-density).



Σχήμα 4.4 Εφαρμογή μάσκας (*masking*) για επιτάχυνση $8\times$, βάσει των πειραμάτων μας [41]. Το σχήμα δείχνει ενδεικτικές Καρτεσιανές μάσκες δειγματοληψίας μεταβλητής πυκνότητας, που χρησιμοποιήθηκαν για υποδειγματοληψία $4\times$ και $8\times$ *undersampling*.

Περιοχές του k -space και γιατί διατηρούμε το κέντρο (center fraction / ACS). Οπως αναφέραμε κέντρο του k -space συγχρατεί τις χαμηλές χωρικές συχνότητες, δηλαδή το γενικό contrast και μεγάλο μέρος του SNR ενώ τα άκρα κωδικοποιούν τις λεπτομέρειες/όρια[46]. Συχνά διατηρείται μια «fully sampled central region» (calibration region ή ACS lines) η οποία εξασφαλίζει ότι το contrast και η βασική πληροφορία παραμένουν άθικτα και παρέχει τα απαραίτητα δεδομένα για εκτίμηση (estimation) των coil sensitivities ή για την εκπαίδευση/κανονικοποίηση κάποιων αλγορίθμων.

Πώς φτιάχνονται μάσκες στον κώδικα Στον κώδικα (π.χ. `singlecoil-mri.py`) υπάρχει συνήθως μια συνάρτηση `create_mask(shape, center_fraction, acceleration, seed=0, flipaxis=False)` που υλοποιεί την παραγωγή 1D Cartesian μάσκας με variable-density επιλογή, με τα εξής βασικά βήματα:

1. Η παράμετρος `center_fraction` περιγράφει πόσο μεγάλο είναι το κεντρικό κομμάτι (π.χ. 0.08 σημαίνει 8% των γραμμών του phase-encode ως πλήρως δειγματοληπτημένες). Η επιλογή του `center_fraction` είναι trade-off: μεγαλύτερο κέντρο δίνει καλύτερο SNR και πιο αξιόπιστα ACS δεδομένα, αλλά μειώνει την πραγματική επιτάχυνση.
2. Υπολογίζεται το πλήθος `num_low_freqs = ρουνδ(num_cols * center_fraction)` και αυτές οι κεντρικές θέσεις (η fully sampled ζώνη) θέτονται σε `True`.

Συγκεκριμένα, ο αλγόριθμος στο αρχείο:

1. Υπολογίζει `num_cols` (π.χ. πλήθος γραμμών phase-encode).
2. Δημιουργεί τον πίνακα `sample_grid` και τοποθετεί την κεντρική fully sampled ζώνη (`pad:pad+num_low_freqs = True`).

3. Επαναληπτικά παίρνει τυχαίες τιμές από $\mathcal{N}(0, 1)$ και αν αυτές βρίσκονται εντός ενός προκαθορισμένου interval (π.χ. $|sample_point| < 3$), τις αντιστοιχεί στην κοντινότερη θέση του grid και θέτει εκεί True.
4. Ελέγχει την τρέχουσα αναλογία δειγμάτων και συνεχίζει μέχρι η πραγματική αναλογία να πλησιάσει ή να ξεπεράσει τη ζητούμενη επιτάχυνση.
5. Μορφοποιεί σε σωστό σχήμα (reshape) και επιστρέφει tensor μάσκας.

Όλο αυτό το μοτίβο είναι υλοποιημένο στη exhaustive_sample και την create_mask που βρίσκονται στο singlecoil_mri.py.

Επιπλέον, στο forward_models_mri.py οι helpers toKspace/fromKspace και οι fftshift/ifftshift wrappers εφαρμόζουν τη μάσκα με απλό πολλαπλασιασμό, π.χ.

```
kspace_data = kspace_data * self.mask
```

Αυτό σημαίνει ότι η μάσκα είναι πράγματι συστατικό του forward model και δεν απαιτείται να εκπαιδευτεί από το δίκτυο.

To dataloader (π.χ. fastmri_dataloader.py) σε αυτή την έκδοση δεν φορτώνει έτοιμες μάσκες από το dataset. Αντίθετα, προετοιμάζει εικόνες/targets από τα .h5 k-σπασες και αφήνει να δημιουργηθεί η μάσκα κατά τη διάρκεια του training/testing (ή να χρησιμοποιηθούν εξωτερικά δημιουργημένες μάσκες). Το αρχείο δείχνει πώς γίνεται το center crop, η IFFT στο raw k-σπασες και η προετοιμασία του input/target για το δίκτυο. Αν θέλουμε να εφαρμόσουμε τις μάσκες που φτιάχνει η create_mask, η κατάλληλη θέση είναι πριν την IFFT (δηλαδή kspace * mask και μετά IFFT για το zero-filled input).

4.3 Το σύνολο δεδομένων (fastMRI)

Τι παίρνουμε από τα .h5 αρχεία. Στο κεφάλαιο αυτό περιγράφουμε με λεπτομέρεια τι ακριβώς εξάγουμε από τα διαθέσιμα .h5 αρχεία του συνόλου δεδομένων (fastMRI single-coil / multi-coil) και πώς το αξιοποιούμε στην αλυσίδα ανακατασκευής μας.

singlecoil_test	6/4/2025 4:16 PM	File folder
singlecoil_train	12/2/2022 9:55 PM	File folder
singlecoil_val	6/4/2025 4:14 PM	File folder

Fig. 4.5 Οι φάκελοι του συνόλου δεδομένων Facebook(META) FastMRI Dataset

Κάθε .h5 αρχείο είναι ένας container που αντιστοιχεί σε ένα scan (συνήθως μία εξέταση / ένα session) και περιέχει έναν αριθμό slices μαζί με μεταδεδομένα.[47] Στα τυπικά αρχεία που χρησιμοποιούμε βρίσκουμε τουλάχιστον το dataset kspace και (όπου παρέχονται) προϋπολογισμένες ανακατασκευές όπως reconstruction_esc και reconstruction_rss.

file1000022.h5	4/19/2023 5:07 PM	H5 File	68,095 KB
file1000036.h5	4/19/2023 5:07 PM	H5 File	61,395 KB
file1000055.h5	4/19/2023 5:07 PM	H5 File	68,095 KB
file1000056.h5	4/19/2023 5:07 PM	H5 File	60,735 KB
file1000063.h5	4/19/2023 5:07 PM	H5 File	65,115 KB
file1000093.h5	4/19/2023 5:07 PM	H5 File	62,575 KB
file1000104.h5	4/19/2023 5:07 PM	H5 File	61,395 KB

```
File attributes:
acquisition: CORPDFS_FBK
max: 0.000851878253624366
norm: 0.0596983310320022
patient_id: 0beb8905d9b7fad304389b9d4263c57d5b069257ea0fdc5bf7f2675608a47406
```

Σχήμα 4.6 Παράδειγμα δεδομένων ενός υποσυνόλου στοιχείων .h5 που χρησιμοποιήθηκαν

Στοιχεία ανά αρχείο (train/val). Στο train και στο val, κάθε αρχείο, με ονόματα όπως file1000001.h5, περιέχει τρία βασικά στοιχεία:

- το kspace, ένας τρισδιάστατος πίνακας με διαστάσεις [slices, height, width] (π.χ. $36 \times 640 \times 372$) σε τύπο complex64.
- το reconstruction_rss, δηλαδή το εικόνισμα που προκύπτει με μέθοδο root-sum-of-squares, σε [slices, 320, 320] float32.
- το reconstruction_esc πάλι σε [slices, 320, 320] float32.

Σε attributes βρίσκουμε πληροφορίες όπως ο κωδικός patient_id, μια ονομασία της ακολουθίας (π.χ. CORPDFS_FBK), καθώς και παράμετροι κανονικοποίησης (norm) και ορίων (max).[46]

```
Dataset 'kspace':
  Shape: (36, 640, 372)
  Data type: complex64
  Dimensions: 3D
  Size: 65.4 MB
  K-space format: Complex64 (slices, height, width)
  Number of slices: 36
  Spatial dimensions: 640 x 372
```

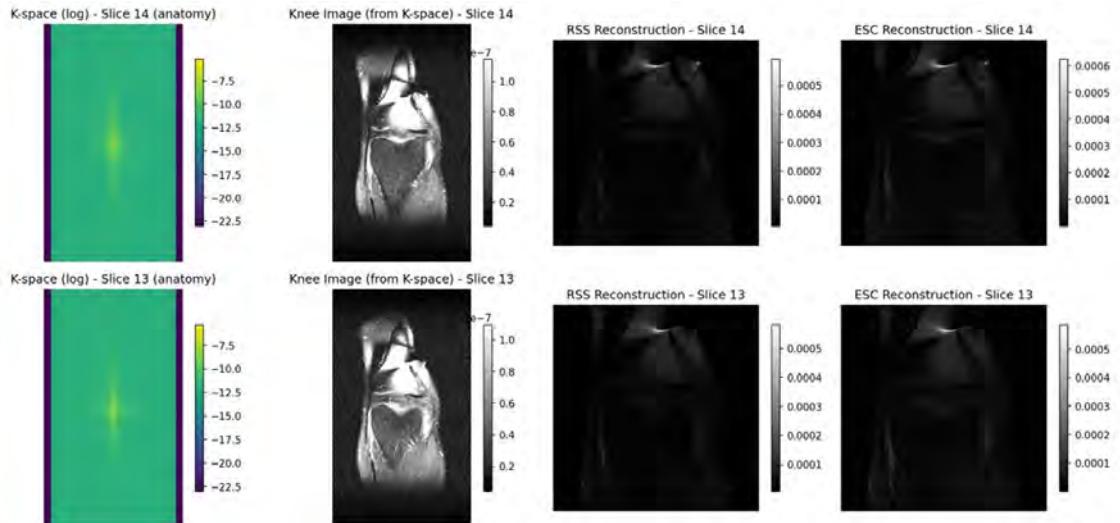
Σχήμα 4.7 Παράδειγμα περιεχομένου στοιχείων στον χώρο-k[47]

Παράδειγμα σχημάτων/μεγεθών. Στο παράδειγμα που δείχνετε στα screenshots, ένα τυπικό αρχείο έχει kspace με shape (36, 640, 372) (δηλαδή 36 slices, spatial dims 640×372) και reconstruction_esc / reconstruction_rss με shape (36, 320, 320) (κεντρικό crop). Τα μεγέθη των αρχείων και ο αριθμός slices ποικίλλουν μεταξύ scans — σε exploratory runs βρίσκουμε αρχεία με $\sim 30\text{--}45$ slices, μέση τιμή ~ 36.4 slices και μέγεθος αρχείου συχνά $\sim 80\text{--}120\text{ MB}$ ανά αρχείο. Οι διαστάσεις 640 × 372 (περίπου) αφορούν το μέγεθος του χώρου k για κάθε slice

(640 γραμμές phase-encode \times 372 στήλες frequency-encode). Στο επίπεδο του .h5 αρχείου, το χώρου k αποθηκεύεται ως ένα dataset μεγέθους [N_slices, H, W] (συχνά complex64), όπου H είναι ο αριθμός των γραμμών phase-encode και W ο αριθμός των frequency-encode. Για κάθε slice, λοιπόν, το χώρου k είναι ένας 2Δ πίνακας[47]: στον άξονα των γραμμών (π.χ. 640) αλλάζουμε την κίνηση του μαγνητικού πεδίου κατά τον προσανατολισμό phase-encode, κι έτσι συλλέγεται μια φάση του σήματος: στον άξονα των στηλών (π.χ. 368 ή 372), αυτό που καταγράφεται είναι η ένταση του σήματος κατά τον άξονα frequency-encode.[47]

```
TRAIN Summary:
Common keys (in all files): ['kspace', 'reconstruction_rss', 'ismrmrd_header', 'reconstruction_esc']
```

Τι εννοούμε με slice. Με slice εννοούμε μία δισδιάστατη τομή του όγκου του ασθενούς (π.χ. μία εγκάρσια τομή γόνατος). Στην αναπαράσταση που δίνει το fastMRI, κάθε slice αποθηκεύεται ως ένας πίνακας στο χώρο των χωρικών συχνοτήτων, δηλαδή ως το χώρου k αυτής της τομής. Ένα .h5 αρχείο περιέχει πολλαπλά τέτοια slices[47]. Στη συνέχεια, κατά την προεπεξεργασία, παίρνουμε κάθε slice ξεχωριστά ως ανεξάρτητο training example. Η μονάδα που τροφοδοτούμε στο δίκτυο συνήθως είναι το slice μετά από εφαρμογή μάσκας / IFFT / crop / normalization και όχι ολόκληρος ο 3-Δ όγκος.



Σχήμα 4.8 Πλήρης οπτικοποημένο παράδειγμα στοιχείων από το σύνολο *singlecoil train*

Σενάρια μονού πηγίου έναντι πολλαπλών, πρακτική σημασία. Στο single-coil σενάριο το κάθε slice αντιστοιχεί σε ένα μοναδικό χώρου k και ο forward operator απλοποιείται σε

$$y = M \mathcal{F}\{x\} + \varepsilon,$$

όπου M είναι η μάσκα δειγματοληψίας, \mathcal{F} ο διαχριτός Fourier (DFT/FFT), x η (σύνθετη) εικόνα στο πεδίο χώρου και ε ο θόρυβος. Στο multi-coil σενάριο κάθε coil c καταγράφει

$$y_c = M \mathcal{F}\{C_c \cdot x\} + \varepsilon_c,$$

όπου $C_c(\mathbf{r})$ είναι ο χάρτης ευαισθησίας του coil c . Οι μετρήσεις από όλα τα coils στοιχίζονται στον συνολικό forward operator A . Στα multi-coil αρχεία υπάρχει επομένως η πληροφορία πολλαπλών k-spaces ανά slice, και συχνά παρέχονται και coil combination reconstructions (RSS/ESC).[47]

ESC & RSS Η βασική τους διαφορά αφορά το αν είχαμε multi-coil δεδομένα. Εδώ, επειδή έχουμε single-coil, δίνουν το ίδιο (ίδια τιμή). Στο multi-coil σενάριο, κάθε coil καταγράφει το δικό του κωδικοποιημένο σήμα: κάθε έχει ελαφρώς διαφορετική θέση και ευαισθησία γύρω από τον ασθενή και άρα παράγονται ξεχωριστά χώρου k δεδομένα για κάθε coil. Η RSS ανακατασκευή τότε υπολογίζεται παίρνοντας, για κάθε εικονοστοιχείο \mathbf{r} ,

$$x_{\text{rss}}(\mathbf{r}) = \sqrt{\sum_{c=1}^C \left| \mathcal{F}^{-1}\{K_c(\mathbf{k})\}(\mathbf{r}) \right|^2} [47].$$

Ενώνουμε έτσι την πληροφορία από όλα τα coils, πετυχαίνοντας καλύτερο SNR. Όμως, στο single-coil dataset που επεξεργαζόμαστε, υπάρχει μόνο ένα coil: άρα δεν υπάρχει τίποτα να συνδυάσουμε μεταξύ διαφορετικών πηγών σήματος. Το αποτέλεσμα είναι ότι και για την ESC και για την RSS, η διαδικασία καταλήγει στο ίδιο ακριβώς magnitude image: παίρνουμε τον σύνθετο αριθμό σε κάθε pixel μετά την inverse FFT, υπολογίζουμε το $|z| = \sqrt{(\Re z)^2 + (\Im z)^2}$ και το αποθηκεύουμε.

Αναπαράσταση χώρου k και τύποι δεδομένων. Το χώρου k αποθηκεύεται ως 3-Δ tensor στο single-coil σενάριο με διάταξη

$$\text{kspcace.shape} = (\text{num_slices}, H, W),$$

και κάθε στοιχείο είναι σύνθετο (τυπικά dtype complex64). Στο multi-coil format υπάρχει επιπλέον διάσταση για τα coils, οπότε ένα κοινό shape είναι

$$\text{kspcace.shape} = (\text{num_slices}, C, H, W)$$

ή, σε κάποιες παραλλαγές, (C, H, W) ανά slice[47]. Τα προϋπολογισμένα πεδία `reconstruction_rss / reconstruction_esc` είναι float32 2-Δ εικόνες (συνήθως κεντρικό crop, π.χ. 320×320).

Τι παίρνουμε εμείς μη η πρακτική ροη μας. Στην υλοποίησή μας δεν βασιζόμαστε σε προϋπολογισμένα reconstructions που μπορεί να υπάρχουν στο .h5.[47] Αντίθετα, το μόνο που εξάγουμε απευθείας από τα αρχεία είναι το raw k-space για κάθε slice και από αυτό παράγουμε προγραμματικά όλα τα υπόλοιπα. Αρχικά, συμβολίζουμε το πλήρες (unmasked) χώρου k του slice με $K(\mathbf{k})$. Εάν επιλέξουμε μία μάσκα δειγματοληψίας $M(\mathbf{k})$ (διάνυσμα/πίνακας με μηδενικά/1), τότε το masked k-space που «βλέπει» ο scanner (ή που θα δώσουμε στο δίκτυο ως input) είναι

$$K_{\text{masked}}(\mathbf{k}) = M(\mathbf{k}) K(\mathbf{k}).$$

H zero-filled εικόνα (warm-start) προκύπτει με inverse Fourier transform:

$$\tilde{x}_{\text{zf}}(\mathbf{r}) = \mathcal{F}^{-1}\{K_{\text{masked}}(\mathbf{k})\}.$$

To ground truth (reference image) λαμβάνεται από την inverse FFT του πλήρους χώρου k:

$$x_{\text{gt}}(\mathbf{r}) = \mathcal{F}^{-1}\{K(\mathbf{k})\},$$

και προαιρετικά εφαρμόζουμε center-crop σε σταθερό μέγεθος (π.χ. 320×320) ώστε όλα τα training examples να έχουν ομοιόμορφο μέγεθος. Σε multi-coil περίπτωση, αν θέλουμε ένα single magnitude reference, κάνουμε coil-combination με root-sum-of-squares:

$$x_{\text{rss}}(\mathbf{r}) = \sqrt{\sum_{c=1}^C \left| \mathcal{F}^{-1}\{K_c(\mathbf{k})\}(\mathbf{r}) \right|^2}.$$

Στην πράξη, εφαρμόζουμε normalization / clipping (π.χ. instance mean/std) πριν την εκπαίδευση, ώστε η κλίμακα να είναι σταθερή μεταξύ δειγμάτων.

Τεχνικές παρατηρήσεις — γιατί δεν χρησιμοποιούμε απλά τις ανακατασκευές. Παρότι τα .h5 περιλαμβάνουν reconstruction_rss ή reconstruction_esc, εμείς δεν τα χρησιμοποιούμε ως ground truth στον βασικό μας training flow.[41] Υποστηρίζουμε στη διπλωματική ότι υπάρχει αξία στο να παράγουμε τα reference images εμείς από το raw k-space με μια ξεκάθαρη και επαναλήψιμη pipeline (IFFT → crop → normalize). Με αυτόν τον τρόπο ελέγχουμε πλήρως τη μονάδα ανακατασκευής, αποφεύγουμε πιθανές παραλλαγές στην παραγωγή ground truth και διασφαλίζουμε reproducibility. Επιπλέον, η δημιουργία των zero-filled inputs και η εφαρμογή των μασκών γίνονται πάντα στην pipeline και όχι έξω από αυτήν — αυτό είναι κρίσιμο όταν θέλουμε να τρέξουμε πειράματα με διαφορετικά sampling patterns όπως στο paper

Πρακτικά: τι υποσύνολα χρησιμοποιούμε. Στη μελέτη μας ακολουθούμε το μοτίβο του paper/repo που υλοποιούμε: χρησιμοποιούμε το singlecoil-train split ως training set και το singlecoil_val ως validation/development

set. Δεν βασιζόμαστε στο official test split του fastMRI για benchmarking, επειδή το πείραμά μας δημιουργεί δικές του μάσκες[41]. Τέλος τα δεδομένα μας αφορούν αποκλειστικά singlecoil data.

Εμείς εξάγουμε από τα αρχεία του fastMRI αποκλειστικά το raw k-space για κάθε slice και, με μία επαναλήψιμη pipeline (εφαρμογή μάσκας, inverse FFT, center-crop, normalization), παράγουμε τόσο τα inputs (zero-filled warm-starts) όσο και τα targets (full IFFT ground truth). Η τυπική αναπαράσταση χώρου k είναι complex64, σε shape (num_slices, H, W) για single-coil και με επιπλέον διάσταση coils για multi-coil αρχεία. Στο πείραμά μας ακολουθούμε το train/val split του fastMRI (**single-coil train** → training, **single-coil val** → validation) και δεν χρησιμοποιούμε τα official test αρχεία για τα εσωτερικά benchmarks, καθώς δημιουργούμε δικά μας undersampling patterns και μάσκες όπως στο paper.

4.4 Μηχανισμός φόρτωσης και Προεπεξεργασία Δεδομένων

Η πρώτη ενέργεια μας είναι να επιλέξουμε ποιες φέτες (slices) από το αρχείο θα επεξεργαστούμε. Στην υλοποίησή μας υλοποιούμε μια multi-slice προσέγγιση, δηλαδή κρατάμε ένα block τομών (συγκεκριμένα τις φέτες 11–27) αντί να χρησιμοποιούμε ένα slice ανά αρχείο. Αυτό γίνεται σκόπιμα επειδή οι ενδιάμεσες φέτες περιέχουν περισσότερο χρήσιμο ανατομικό περιεχόμενο του γόνατου και βελτιώνουν την εκπαίδευτική στατιστική χωρίς να αλλάζουν τη θεμελιώδη λογική της ανακατασκευής.

Αφού έχουμε επιλέξει τα slices, μετατρέπουμε τα δεδομένα σε μορφή που ο κώδικας FFT/IFFT περιμένει. Η εσωτερική μας αναπαράσταση για τις πολύπλοκες τιμές είναι συχνά ένα πραγματικό διάνυσμα με τελευταίο άξονα δύο στοιχείων (real, imag), γιατί έτσι είναι σχεδιασμένα τα helpers fft2/ifft2 και οι wrappers toKspace / fromKspace που εξετάσαμε.

Ο dataloader μας (MultiSliceFastMRIDataloader) είναι σχεδιασμένος να υλοποιεί όλα τα παραπάνω με έναν βήμα-προς-βήμα τρόπο. Ανοίγει κάθε .h5, διαβάζει kspace, επιλέγει slices 11–27, δημιουργεί (ή φορτώνει) τη μάσκα, εφαρμόζει τη μάσκα στο χώρου k, εκτελεί IFFT για να φτιάξει zero-filled input, εκτελεί IFFT για το πλήρες χώρου k (αυτό δίνει το ground truth image) και επιστρέφει για κάθε δείγμα ένα σύνολο αντικειμένων που χρησιμοποιούμε κατά την εκπαίδευση: το (masked_kspace, zero_filled_image, ground_truth_image). Ο Loader υποστηρίζει batching σε διάφορα επίπεδα: είτε επιστρέφει stacks πολλαπλών slices ως ένα sample (multi-slice batch), είτε ξεχωριστά slices ανά δείγμα, ανάλογα με τη ρύθμιση training. Στην περίπτωσή μας αποφασίσαμε να κρατήσουμε την ίδια λογική εισόδου/target με το paper που εφαρμόζουμε[41], αλλά να δουλεύουμε με block slices (11–27) ώστε να εκμεταλλευτούμε σταθερότερο training και πλουσιότερο

περιεχόμενο σε κάθε αρχείο.

Στο παράρτημα A παρατίθεται ο πλήρης κώδικας κατασκευής του multislice approach.

Το επόμενο βήμα είναι η κατασκευή της μάσκας δειγματοληψίας.

Η μάσκα παράγεται στη συνάρτηση `create_mask` στο `singlecoil_mri.py` (ή η αντίστοιχη κλάση/μέθοδος στο `dataloader`): πρώτα ορίζουμε μία fully-sampled κεντρική ζώνη μεγέθους `center_fraction` (οι κεντρικές γραμμές που κρατάμε πάντα), και στη συνέχεια συμπληρώνουμε γραμμές/σημεία τυχαία (με `seed` για reproducibility) ώστε να εκπεφρασθεί ο επιθυμητός παράγων επιτάχυνσης acceleration. Η μάσκα επιστρέφεται ως δυαδικός πίνακας (True/False) που έχει σχήμα συμβατό με το χώρου k .

Πριν τροφοδοτήσουμε το δίκτυο κάνουμε επιπλέον επεξεργασία: πρώτον, εφαρμόζουμε `center_crop` στα spatial dimensions ώστε όλα τα δείγματα να έχουν το ίδιο μέγεθος που απαιτεί το μοντέλο (συνηθισμένα 320×320 σε εργασίες με knee). Δεύτερον, κάνουμε normalization per-instance: αφαιρούμε το μέσο και διαιρούμε με την τυπική απόκλιση του δείγματος (ή εφαρμόζουμε άλλους κανόνες scaling όπως division by $\max(\text{abs}(\text{image}))$). Η normalization step είναι χρήσιμη για σταθερή συμπεριφορά του optimizer. στην practice pipeline την υλοποιούμε με τον τύπο

$$x \leftarrow \frac{x - \mu}{\sigma},$$

όπου μ, σ υπολογίζονται per example. Τέλος, εφαρμόζουμε clipping (π.χ. clamp σε $[-6, 6]$ σε normalized units) αν το training scheme το απαιτεί, ώστε να αφύγουμε ακραία outliers που θα διαταράξουν το batch statistics.

Ο `cartesianSingleCoilMRI` (όπως εμφανίζεται στα helpers / forward operators) παρέχει την υλοποίηση του operator A και του adjoint A^T για single-coil Cartesian λήψη. Στην πράξη ο `cartesianSingleCoilMRI` εφαρμόζει τις βασικές πράξεις που άμεσα αντιστοιχούν σε: πολλαπλασιασμό με `map` ευαισθησίας (σε multi-coil περίπτωση), εφαρμογή FFT/IFFT με σωστή οργάνωση shift, και εφαρμογή μάσκας M .

Στο παράρτημα A `SinglecoilMRI.py` παρατίθεται ο κώδικας.

Επιπρόσθετες τεχνικές λεπτομέρειες που ενσωματώσαμε: κρατάμε το forward model ακριβές και μη-εκπαιδευτούμενο και επιτρέπουμε στο learned μέρος να μαθαίνει το υπόλοιπο/ denoising/prox behaviour.

Συνοψίζοντας, στο στάδιο `dataloader + preprocessing` παίρνουμε το raw `.h5` χώρου k , επιλέγουμε slice blocks, δημιουργούμε deterministic variable-density μάσκες, εφαρμόζουμε την μάσκα στο χώρου k , παράγουμε zero-filled inputs με

IFFT, παράγουμε ground truth images με full IFFT, κάνουμε center crop και instance normalization/clipping, και τροφοδοτούμε το δίκτυο με τα (input, target) ζεύγη. Οι helpers FFT/IFFT/shift που χρησιμοποιούμε είναι οι wrappers του `forward_models_mri.py` (κλάσεις `toKspace/fromKspace` και οι συναρτήσεις `fft2/ifft2`) οι οποίες εξασφαλίζουν συνεπή, centered, orthonormal μετασχηματισμό κατά μήκος των διαστάσεων εικόνας/συχνότητας.

Κεφάλαιο 5

Μεθοδολογία: ανακατασκευή MRI με μέθοδο Βαθιάς Ισορροπίας με Μέθοδο Εγγύς Κλίσης

Τώρα που αναλύσαμε τις βασικές έννοιες των μασκών , της προεπεξεργασίας και του k-space μπορούμε να προχωρήσουμε στην ανάλυση επεξεργασίας της ανακατασκευής και της μεθόδου που ακολουθείται.Περιγράφεται η υλοποίηση του συστήματος, με βάση τη μελέτη που παρουσιάστηκε στο προηγούμενο κεφάλαιο. Δίνονται οι λεπτομέρειες υλοποίησης με την συγκεκριμένη προσέγγιση για τους βασικούς αλγορίθμους του συστήματος.

5.1 Βασικές διατυπώσεις μοντέλου Ισορροπίας:

Πρώτα απ' όλα οπως αναλύθηκε, ποιο είναι το πρόβλημα: έχουμε μια άγνωστη εικόνα x^* και λαμβάνουμε μετρήσεις y που είναι μια μετασχηματισμένη, ημιτελής και ύφορυβώδης, εκδοχή αυτής της εικόνας. Το μοντέλο, όπως είδαμε, γράφεται απλά ως

$$y = Ax^* + \varepsilon.$$

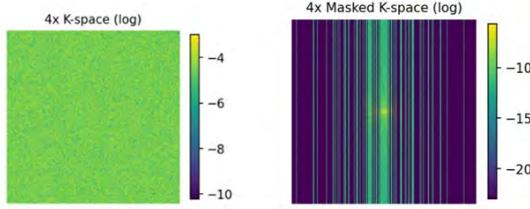
Αυτό σημαίνει: το A παίρνει την εικόνα και επιστρέφει τις μετρήσεις που βλέπει το μηχάνημα, και το ε είναι ο ύφορυβος που προστίθεται στις μετρήσεις. Η βασική δουλειά του αλγορίθμου είναι να βρει μια εκτίμηση \hat{x} που να είναι αντιπροσωπευτική ως προς αυτές τις μετρήσεις και παράλληλα να μοιάζει ρεαλιστική ως εικόνα.

Ποσοτικές συνθήκες συστατικότητας Στη συγκεκριμένη διατύπωση προσεγγίζουμε το πρόβλημα στον χώρο του MRI, ως ο y να είναι είναι ένα πλέγμα συχνοτήτων όπου κρατάμε τις έγκυρες μετρήσεις και οι υπόλοιπες είναι μηδενικές.

Πρακτικά,

$$y = full - kspace \odot mask$$

Αυτή χρειάζεται για να κρατάει πάντα στον χώρο k-space τις τιμές της πραγματικής μέτρησης στα σημεια που δεν υπάρχει μάσκα.[41]



Σχήμα 5.1 Περιεχόμενο του χώρου k στις y τιμές μας

Τι κάνει ο operator A στην MRI περίπτωση και τι είναι ο A^\top : Εδώ, ο A να είναι ο συνδυασμός του Fourier μετασχηματισμού και της μάσκας. Πρακτικά,

$$Ax = FFT(x) \odot mask$$

δηλαδή έχουμε εικόνα, την μετασχηματίζουμε στον χώρο Fourier κρατώντας μόνο τις συχνότητες που μετρήθηκαν (οι άλλες μηδενίζονται)[22]. Ο A^\top (ο adjoint) κάνει το αντίστροφο[22]: παίρνει τις υποδειγματισμένες συχνότητες, εφαρμόζει ξανά τη μάσκα και επιστρέφει στο χώρο εικόνας με IFFT, επομένως

$$A^\top z = IFFT(z \odot mask).$$

Η πρώτη ανακατασκευή που φτιάχνουμε είναι η αρχικοποίηση με μηδενική πλήρωση $x^{(0)} = A^\top y$, δηλαδή κάνουμε τον αντίστροφο μετασχηματισμό Fourier πάνω στο masked k-χώρο. Αυτή η εικόνα έχει σωστές τιμές στις μετρημένες συχνότητες αλλά εμφανή aliasing/artefacts εκεί που δεν μετρήσαμε. Η μηδενική πλήρωση, είναι η πρώτη μας ανακατασκευη η οποία προκύπτει από το τη διαδικασία προεπεξεργασίας μας ως εικόνας εισόδου έχοντας ορίσει τις μάσκες και τους μετασχηματισμούς με τον παραπάνω τρόπο.

Τι είναι πλεον ο όρος συνέπειας δεδομένων» :Ο ορος συνέπειας δεδομένων είναι ο μηχανισμός που επιστρέφει την εκτίμησή μας έτσι ώστε στις θέσεις που έχουμε μετρήσεις στον χώρο k να συμφωνεί με το y. Ένα απλό μέτρο αυτής της απόκλισης είναι $y - Ax$ (πόσο απέχει το k-space της πρόβλεψής μου από τις αληθινές μετρήσεις)[22]. Επιστρέφοντας αυτό στο πεδίο εικόνας με A^\top έχουμε την εικόνα των σφαλμάτων

$$A^\top(y - Ax)$$

για κάθε pixel, πόσο πρέπει να μετατοπιστεί το εκάστοτε βήμα. Αυτό το σήμα προστίθεται στην εκτίμηση ώστε να την φέρει πιο κοντά στα δεδομένα

$$x + \eta A^\top (y - Ax)$$

Ο κανονικοποιητής(regularizer) συνοπτικά: επειδή λοιπόν υπάρχει έλλειψη πληροφορίας από τιμές που δεν μετρήθηκαν, χρειαζόμαστε ένα στοιχείο που θα συμπληρώνει τα κενά με ρεαλιστικό τρόπο. Αυτό το ρόλο παίζει ο κανονικοποιητής ή το εκπαιδευμένο δίκτυο R_θ και έχουμε: Η ιδέα είναι να εκπαιδεύσουμε ένα δίκτυο που γνωρίζει από εικόνες τι μοιάζει ρεαλιστικό και να το χρησιμοποιούμε για να καθαρίζει την εκτίμηση μας. Στις μεθόδους DE-Prox[22], η οποία εφαρμόστηκε, αυτή η λειτουργία αντικαθιστά το prox-βήμα της κλασικής proximal gradient με ένα trainable module [41]

$$R_\theta(x + \eta A^\top (y - Ax))$$

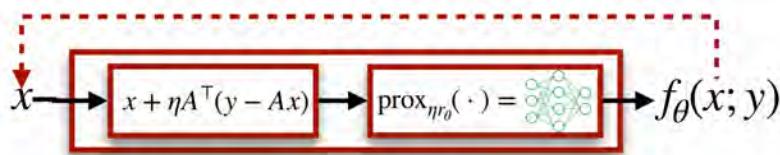


Figure 4: Deep Equilibrium Proximal Gradient Descent (DE-PROX)

[41]

5.2 Προσεγγιση και Αλγοριθμική Λογική μεθόδου:

Το DE-Prox [22] είναι μια συγκεκριμένη επιλογή του f_θ . Φρτιάχνουμε το f_θ ώστε να περιλαμβάνει ορος συνέπειας δεδομένων και learned prox. Δεύτερον, προτιμούμε να κάνουμε pretraining του R_θ (π.χ. ως denoiser) για καλύτερη αρχικοποίηση. Συνδέοντας την ιδέα του μοντέλου ισορροπίας, για το forward επιλέγουμε Anderson όταν θέλουμε γρήγορη σύγκλιση και ρυθμίζουμε ένα μέγιστο πλήθος επαναλήψεων. Για το backward χρησιμοποιούμε επαναληπτικό λύτη [41] για να βρούμε v και φροντίζουμε η ακρίβεια του λύτη να είναι επαρκής ώστε τα gradients να είναι χρήσιμα. Χρησιμοποιούμε τεχνικές σταθεροποίησης στο R_θ (spectral normalization, residual connections) ώστε ο χάρτης να ικανοποιεί υπόθεση Lipschitz που δίνει θεωρητική σύγκλιση.

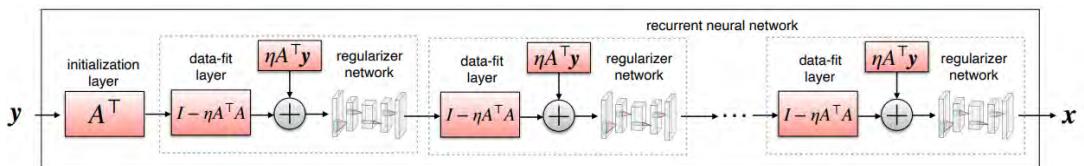
Πώς συνδέονται όλα αυτά με το MR πρόβλημά μας: Έχουμε αναφέρει ως x είναι η εκάστοτε εικασία εικόνων, y είναι οι μετρήσεις στο k-space, A είναι ο masked FFT (ή γενικότερα ο forward operator), και R_θ είναι το learned denoiser ή proximal. Η επαναληπτική σχέση στο DE-Prox είναι

$$x^{(k+1)} \leftarrow R_\theta(x^{(k)} + \eta A^\top (y - Ax^{(k)})).$$

Ο στόχος της εκπαίδευσης είναι να μάθουμε θ ώστε το τελικό x^* που προκύπτει από αυτή τη σχέση να είναι κοντά στην αληθινή εικόνα x^{true} ως προς ένα σφάλμα (μέσο τετραγωνικό σφάλμα ή άλλο). Συγκεκριμένα, στο πλαίσιο αυτό παίρνουμε

$$f_\theta(x; y) = R_\theta(x + \eta A^\top(y - Ax)),$$

όπου A είναι ο τελεστής μέτρησης, y οι παρατηρήσεις, η ένα μικρό βήμα, και R_θ το learned proximal ή denoiser.



Σχήμα 5.3 Ενσωμάτωση των A και A^\top μέσω ξετυλίγματος (unrolling) με proximal gradient. Εδώ η προσέγγιση απεικονίζεται ως ξετύλιγμα του αλγορίθμου καθόδου εγγύτητας κλίσης (proximal gradient descent), όπου ο proximal τελεστής αντικαθίσταται από ένα βαθύ CNN [22].

To DE-Prox συνδυάζει την ιδέα του learned prox από τα unrolled proximal-gradient δίκτυα με τη φιλοσοφία DEQ: δεν εκπαίδευούμε για προκαθορισμένο πλήθος βημάτων K , αλλά για το τελικό fixed point x^* όπου $x^* = f_\theta(x^*; y)$ [41]. Αυτή είναι και η βασική διαφορά: ίδιο δομικό στοιχείο με τα unrolled prox-grad, αλλά απεριόριστο βάθος και implicit training.

Εμπρόσθιος λύτηςforward solver Στην απλούστερη μορφή, κάνουμε απλή επανάληψη

$$x_{k+1} = f_\theta(x_k; y)$$

ξεκινώντας από x_0 (συνήθως το zero-filled inverse ή άλλη αρχικοποίηση). Αυτό λειτουργεί καλά όταν ο χάρτης f_θ είναι contractive όπως αναλύθηκε (έχει Lipschitz σταθερά μικρότερη του 1), οπότε οι επαναλήψεις προσεγγίζουν σταθερά ένα μοναδικό σημείο. Στην πράξη, οι χάρτες μπορεί να μην είναι τόσο απλοί, οπότε υπάρχουν επιταχυντές ή root-finding μέθοδοι που χρησιμοποιούμε για να φτάσουμε πιο γρήγορα ή πιο σταθερά στο fixed point. Οι δύο πιο διαδεδομένες είναι η Anderson acceleration και η μέθοδος του Broyden. Στο DE-Prox paper συγχρίνονται και τα δύο και καταλήγουμε ότι η Anderson ήταν η πιο πρακτική επιλογή για τα πειράματά μας, ενώ το Broyden ήταν επίσης χρήσιμο αλλά πιο ακριβό ανά βήμα[41].

Οπίσθιος επιλυτής (backward solver) Στο ξετυλιγμένο μοντέλο, αν έχουμε K βήματα, η οπισθιοδίδοση περνά μέσα από κάθε βήμα και χρειάζεται να κρατήσουμε όλα τα x_1, \dots, x_K , άρα η μνήμη αυξάνεται γραμμικά με το K . Στο DEQ, στην εκπαίδευση βρίσκουμε πρώτα το σταθερό σημείο x^* και έπειτα υπολογίζουμε

το gradient της απώλειας $\ell(x^*)$ ως προς τις παραμέτρους θ με βάση το θεώρημα της έμεσης παραγώγισης . Αυτό οδηγεί

$$\frac{\partial \ell}{\partial \theta} = \frac{\partial \ell}{\partial x^*} (I - \partial_x f_\theta(x^*; y))^{-1} \partial_\theta f_\theta(x^*; y).$$

Η πρακτική υλοποίηση δεν αντιστρέφει τον μεγάλο πίνακα. Αντίθετα, λύνουμε για το διάνυσμα v που ικανοποιεί

$$(I - \partial_x f_\theta(x^*; y))^\top v = \left(\frac{\partial \ell}{\partial x^*} \right)^\top,$$

και στη συνέχεια

$$\frac{\partial \ell}{\partial \theta} = v^\top \partial_\theta f_\theta(x^*; y).$$

Αυτό σημαίνει ότι χρειαζόμαστε έναν αριθμητικό λύτη για να βρούμε το v . Ο λύτης αυτός δεν σχηματίζει ποτέ ολόκληρο τον Jacobian· αντίθετα απαιτεί μόνο JVP ή VJP (Jacobain–vector και vector–Jacobain products), που τα σύγχρονα autograd frameworks (PyTorch ή JAX) υπολογίζουν αποτελεσματικά. Το autograd είναι το σύστημα που υπολογίζει αυτόματα παραγώγους. Αυτή είναι η ουσία του implicit differentiation: υπολογίζουμε gradients χωρίς να ξεδιπλώσουμε όλες τις επαναλήψεις και συνεπώς η μνήμη δεν κλιμακώνεται με το K . Αυτό είναι το κύριο πρακτικό κέρδος των DEQ έναντι του deep unrolling.

Στο DEQ χρειάζεται να καλέσουμε πολλές φορές προϊόντα Jacobian–vector κατά τη λύση του γραμμικού συστήματος για το v . Τα frameworks παρέχουν API για τον υπολογισμό των προϊόντων αυτών χωρίς ρητή δημιουργία του Jacobian. Ο wrapper (π.χ. ένα DEQModule σε PyTorch) λειτουργεί ως εξής: στο forward καλεί τον λύτη σταθερού σημείο και επιστρέφει x^* . Κατά το backward, ο wrapper έχει hook που τρέχει τη ρουτίνα έμεσης παραγώγισης: καλεί επαναληπτικό λύτη, κάνει πολλαπλές κλήσεις σε JVP ή VJP και τελικά επιστρέφει τα gradients για τα βάρη. Έτσι, στην εκπαίδευση δεν χρειάζεται να χειριστούμε ξεχωριστά forward και backward, καθώς ο wrapper τα ενσωματώνει και αποκρύπτει την πολυπλοκότητα.

Μνήμη και υπολογιστικό trade-off. Η κύρια εξοικονόμηση μνήμης προέρχεται από το ότι δεν αποθηκεύουμε όλα τα ενδιάμεσα x_k για να κάνουμε backprop μέσω χρόνου[42]. Απαιτείται μνήμη για το τελικό x^* και λίγα βοηθητικά αντικείμενα. Αυτό καθιστά πρακτικά εφικτή την εκπαίδευση στο fixed point, κάτι που τα unrolled δίκτυα δυσκολεύονται να επιτύχουν χωρίς μεγάλη χρονική επιβάρυνση. Το τίμημα είναι ότι το backward απαιτεί επαναληπτικό λύτη, δηλαδή πολλαπλές εφαρμογές JVP/VJP με υπολογιστικό κόστος[42]. Το trade-off είναι μνήμη έναντι υπολογισμού: το μοντέλο IσορροπίαςDEQ μεταφέρει το κόστος από μνήμη σε περισσότερες επαναληπτικές πράξεις στο backward.

Οι μετρικές PSNR και SSIM :

Στη μετρική PSNR(Peak Signal to Noise Ratio)[50] μετράμε την πιστότητα ανακατασκευής μέσω μιας λογαριθμικής κανονικοποίησης του μέσου τετραγωνικού σφάλματος ως προς το δυναμικό εύρος της έντασης. Θεωρούμε δύο εικόνες $x, y \in \mathbb{R}^{M \times N}$ και ορίζουμε το σφάλμα $e_{ij} = x_{ij} - y_{ij}$. Το μέσο τετραγωνικό σφάλμα είναι $\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N e_{ij}^2$. Αν L δηλώνει το δυναμικό εύρος της έντασης, για B -bit απεικόνιση ισχύει $L = 2^B - 1$ ενώ σε εικόνες κανονικοποιημένες στο $[0, 1]$ ισχύει $L = 1$. Ως ορισμό του λόγου έχουμε :

$$\text{PSNR [dB]} = 10 \log_{10} \left(\frac{L^2}{\text{MSE}} \right) = 20 \log_{10} \left(\frac{L}{\sqrt{\text{MSE}}} \right)$$

Η μονάδα είναι dB και μεγαλύτερες τιμές υποδηλώνουν μικρότερο σφάλμα ως προς το μέγιστο επιτρεπτό πλάτος σήματος. Η μετρική είναι υπολογιστικά οικονομική, κλιμακώνεται με το δυναμικό εύρος και επιτρέπει συγκρίσεις μεταξύ αλγορίθμων κωδικοποίησης ή αποθήραυβοποίησης, αλλά δεν ενσωματώνει αντιληπτικά χαρακτηριστικά και μπορεί να μην αντανακλά τοπικές δομικές αλλοιώσεις ή οπτικά σημαντικές παραμορφώσεις.

Για την αντιληπτική αξιολόγηση υιοθετούμε τον Δείκτη Δομικής Ομοιότητας SSIM (Structural Similarity Index Measure)[51], ο οποίος αποτυπώνει συστηματικά φωτεινότητα, αντίθεση και δομή σε τοπικά παράθυρα και έπειτα τα συνδυάζει. Με αυτό το τρόπο αποτελεί μια φυσική μετρική ομοιότητας εικόνων. Με τοπικούς μέσους μ_x, μ_y , τυπικές αποκλίσεις σ_x, σ_y και συνδιακύμανση σ_{xy} , ο γενικός ορισμός είναι

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma, \quad l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1},$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \quad s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}.$$

Συνήθως τίθενται $\alpha = \beta = \gamma = 1$ και $c_3 = c_2/2$, που οδηγούν στη συμπαγή μορφή

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

όπου οι σταθερές σταθεροποίησης κλιμακώνονται με το δυναμικό εύρος L ως $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ με τυπικές τιμές $k_1 = 0.01$, $k_2 = 0.03$. Ο δείκτης λαμβάνει τιμές κοντά στη μονάδα για υψηλή ομοιότητα, υπολογίζεται συνήθως κατά παράθυρο και έπειτα εξάγεται ο μέσος όρος επί του πεδίου, παρουσιάζει καλύτερη συσχέτιση με την ανθρώπινη αντίληψη διότι είναι ευαίσθητος σε δομικές αλλαγές, και απαιτεί ζεύγος εικόνας αναφοράς(ground truth) και υπό εξέταση για να είναι καλά ορισμένος.

5.3 Χάρτης βαθιάς Ισορροπίας με μέθοδο εγγύς κλίσης

Στόχευση. Θέλουμε να υλοποιήσουμε έναν επαναληπτικό χάρτη που ενσωματώνει ρητά τη συνέπεια δεδομένων και έναν εκμαθημένο κανονικοποιητή τύπου αποθρούβοποιητή, ώστε το σταθερό σημείο να Ισορροπεί ανάμεσα στο προωθητικό μοντέλο μέτρησης και στις στατιστικές των εικόνων.

Ορισμός χάρτη και σταθερού σημείου. Θέτουμε

$$f_\theta(x; y) = R_\theta\left(x + \eta A^\top(y - Ax)\right), \quad x^* = f_\theta(x^*; y).$$

Ο όρος $x + \eta A^\top(y - Ax)$ είναι βήμα κλίσης για την τετραγωνική συνέπεια δεδομένων $\frac{1}{2}\|Ax - y\|_2^2$, ενώ ο R_θ λειτουργεί ως (προσεγγιστικός) εγγύς τελεστής ενός εκμαθημένου προτέρου/κανονικοποιητή.

Διφασική υλοποίηση (α) Προεκπαίδευση του R_θ ως αποθρούβοποιητή. Εκπαιδεύουμε τον R_θ ως DnCNN αποθρούβοποιητή σε δεδομένα εικόνας με συνθετικό λευκό Γκαουσιανό θόρυβο, σε υπολειμματική μορφή

$$R_\theta(v) = v + N_\theta(v),$$

και με φασματική κανονικοποίηση στα συνελικτικά φίλτρα, ώστε να ελέγχεται η τελεστηριακή νόρμα και να σταθεροποιείται ο Ιακωβιανός (μικρή $\text{Lip}(R_\theta - I)$). ([43])

(β) Ένταξη στο επαναληπτικό μπλοκ DE-Prox (MRI). Ο προεκπαϊδευμένος R_θ εντάσσεται στον χάρτη f_θ ανωτέρω και η ανακατασκευή ορίζεται ως αναζήτηση σταθερού σημείου $x^* = f_\theta(x^*; y)$.

Μετάβαση από μέθοδο ξεδιπλώματος σε Ισορροπία Υλοποιούμε εύρεση σταθερού σημείου με επιτάχυνση Άντερσον, ενώ το οπίσθιο πέρασμα βασίζεται σε έμμεση διαφοροποίηση της συνθήκης Ισορροπίας, χρησιμοποιώντας γινόμενα Ιακωβιανού-διανύσματος και συζυγούς Ιακωβιανού-διανύσματος (χωρίς αποθήκευση όλων των ενδιάμεσων επαναλήψεων). Οι βασικές υπερπαράμετροι είναι: το μέγεθος μνήμης m (Άντερσον), ο συντελεστής τακτικοποίησης λ , η ανοχή tol και ο συντελεστής μίξης β . ([41],[43])

Ισορροπία ανάμεσα σε αποτύλιξη και Ισορροπία Με τη συγκεκριμένη σχεδίαση:

- Ενσωματώνουμε ρητά τους τελεστές A και A^\top στον επαναληπτικό κύκλο, ώστε κάθε βήμα να έχει σαφή ερμηνεία ως βήμα συνέπειας δεδομένων.

- Δεν δεσμευόμαστε σε σταθερό βάθος K : εκπαιδεύουμε/εκτελούμε απευθείας στο σταθερό σημείο x^* , εξοικονομώντας μνήμη στο οπίσθιο πέρασμα με έμμεση διαφοροποίηση και αποκτώντας ανθεκτικότητα ως προς τον αριθμό επαναλήψεων.
- Η παρουσία του αποθορυβοποιητή R_θ γεφυρώνει το κενό πληροφορίας από τον υποδειγματοληπτημένο χώρο k , ενώ ο έλεγχος ομαλότητας μέσω φασματικής κανονικοποίησης και υπολειμματικής παραμετροποίησης καθιστά ρεαλιστική την τήρηση των θεωρητικών υποθέσεων σύγκλισης. ([43])

Κεφάλαιο 6

Αποθορυβοποιητής (Denoiser) και Αξιολόγηση Μοντέλου Ισορροπίας

6.1 Προεκπαιδευμένος Αποθορυβοποιητής

Όπως έχουμε επισημάνει στο equilibrium, ο regularizer διαδραματίζει τον ρόλο denoiser. Εξετάζουμε γιατί προτιμούμε αυτόν τον denoiser και τι απαιτείται ώστε να λειτουργεί σωστά μέσα στο σχήμα DE-Prox, με έμφαση στην επιλογή R_θ ως DnCNN[52], στο pretraining, στη spectral normalization και στο warm start του forward, τα οποία από κοινού ενισχύουν τη σταθερότητα και βελτιώνουν τη γενίκευση.

Ξεκινάμε από τον ρόλο του R_θ μέσα στον επαναληπτικό χάρτη. Στο DE-Prox ορίζουμε

$$f_\theta(x; y) = R_\theta(x + \eta A^\top(y - Ax)),$$

όπου το R_θ λειτουργεί ως learned proximal που καταστέλλει θόρυβο και artefacts και συμπληρώνει το nullspace του A . Με αυτόν τον τρόπο διασφαλίζουμε ότι ο denoiser αναλαμβάνει τις αβέβαιες συνιστώσες και τις μικρές λανθασμένες διακυμάνσεις που απομένουν μετά το βήμα ορος συνέπειας δεδομένων. Η σύνδεση με τα θεωρήματα σύγχλισης είναι άμεση και καθοδηγεί τις επιλογές μας. Τα αποτελέσματα τύπου DE-Grad και DE-Prox προϋποθέτουν έλεγχο του $\text{Lip}(R_\theta - I)$ [41] και επιλογή του η σε διάστημα που εξαρτάται από τις ιδιοτιμές του $A^\top A$. Η residual παραμετροποίηση και η spectral normalization είναι τα εργαλεία για τον πρώτο όρο, ενώ το pretraining μειώνει το λειτουργικό εύρος του N_θ άρα και την πρακτική Lipschitz συμπεριφορά στην κατανομή εισόδου. . Με αυτά τα στοιχεία στη θέση τους, ο χάρτης f_θ γίνεται σταθερός και το fixed-point training αποκτά προβλέψιμη συμπεριφορά.

Από πλευράς υλοποίησης του backward με implicit differentiation, ο προεκπαι-

δευμένος DnCNN διευκολύνει τη λύση της γραμμικής εξίσωσης

$$(I - \partial_x f_\theta(x^*; y))^\top v = \left(\frac{\partial \ell}{\partial x^*} \right)^\top v,$$

ενώ η spectral normalization περιορίζει αιχμές στο φάσμα του Jacobian. Έτσι οι iterative solvers συγκλίνουν ταχύτερα σε χρήσιμη ακρίβεια, κάτι που μεταφράζεται σε σαφές υπολογιστικό όφελος στο backward. Επιπλέον, στο βασικό paper έχουμε πειραματιστεί και με U-Nets, αλλά διαπιστώσθηκε ότι ο DnCNN απέδωσε ανώτερα τόσο για τις προτεινόμενες deep equilibrium μεθόδους όσο και για τις μεθόδους σύγκρισης[41].

Η spectral normalization αποτελεί δεύτερο στοιχείο σταθερότητας. Τη θεωρούμε αναγκαία επειδή οι θεωρητικές συνθήκες σύγκλισης εκφράζονται με φράγματα τύπου $\text{Lip}(R_\theta - I) \leq \alpha$ με $\alpha < 1$, και το επιτρεπτό εύρος τιμών του βήματος η επηρεάζεται από το φάσμα του $A^\top A$. Πρακτικά, η spectral normalization αποτρέπει μεγάλες ενισχύσεις συγκεκριμένων χωρικών μοτίβων που αλλιώς θα μπορούσαν να οδηγήσουν είτε σε ασταθές fixed point είτε σε προβληματικά gradients στο backward[41].

6.1.1 Προεκπαίδευση με λευκό Γκαουσιανό θόρυβο

Στο pretraining του R_θ ως denoiser κατασκευάζουμε ζεύγη εκπαίδευσης από καθαρές εικόνες x και προσθέτουμε λευκό Gaussian θόρυβο, δηλαδή δημιουργούμε

$$\tilde{x} = x + \sigma \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, I),$$

με σ δειγματοληπτούμενο από ένα εύρος τιμών. Ο denoiser εκπαιδεύεται αποκλειστικά σε λευκό Gaussian θόρυβο με απώλεια τύπου MSE, είτε στην καθαρή εικόνα

$$\|R_\theta(\tilde{x}) - x\|_2^2,$$

είτε στο residual

$$\|N_\theta(\tilde{x}) - (x - \tilde{x})\|_2^2.$$

Αυτή η φάση προσφέρει δύο ουσιώδη οφέλη. Από πλευράς βελτιστοποίησης, το δίκτυο μαθαίνει από νωρίς να πραγματοποιεί μικρές και ρεαλιστικές διορθώσεις αντί για ασταθείς μεγάλες μεταβολές, γεγονός που μειώνει αστοχίες στον forward fixed-point solver, οδηγεί σε λιγότερες επαναλήψεις και σταθεροποιεί την εκπαίδευση με implicit differentiation. Από πλευράς γενίκευσης, το δίκτυο αποκτά τάση να διατηρεί τις δομές των εικόνων και να απορρίπτει υψηλής συχνότητας θόρυβο, κάτι που μεταφέρεται όταν αργότερα το σύστημα εκτίθεται στην πιο σύνθετη επαναληπτική ροή με τον τελεστή A . Επιπλέον το γεγονός ότι εχουμε προσθέσει αντίστοιχο θόρυβο στις μασκαρισμένες εκδοχές εικόνων δημιουργεί την απαίτηση δημιουργίας του DnCNN[52].

6.1.2 Αρχιτεκτονική DnCNN

Σε επίπεδο αρχιτεκτονικής, υλοποιούμε τον DnCNN[52] με διαδοχικές συνελικτικές στρώσεις και ενεργοποιήσεις τύπου ReLU, υλοποιώντας τον N_θ ως DnCNN 17 layer με μικρούς πυρήνες 3×3 και 64 κανάλια σε κάθε ενδιάμεση συνελικτική, με διάταξη

$$\text{Conv(SN)} \rightarrow \text{ReLU} \rightarrow [\text{Conv(SN)} \rightarrow \text{GroupNorm(affine = False)} \rightarrow \text{ReLU}]^{ \times 15} \rightarrow \text{Conv(SN)}.$$

To spectral norm σε κάθε Conv κρατά τη μέγιστη ιδιοτιμή των φίλτρων υπό έλεγχο[41]. Το τελικό Conv επιστρέφει residual με ίδιο πλήθος καναλιών με την είσοδο (δύο κανάλια για το σύνθετο image-space), ώστε

$$R_\theta(v) = v + N_\theta(v)$$

να εφαρμόζεται άμεσα μέσα στο βήμα

$$f_\theta(x; y) = R_\theta(x + \eta A^\top(y - Ax)).$$

Στην πράξη, αυτό το pretraining μειώνει τον αριθμό επαναλήψεων που χρειάζεται ο fixed-point solver, σταθεροποιεί το φάσμα του Ιακωβιανού $\partial_x f_\theta$ και επιταχύνει την (implicit) backpropagation, ενώ το fine-tuning στο DE-Prox μετατρέπει τον γενικό denoiser σε ειδικό learned proximal προσαρμοσμένο στον A , στη mask υποδειγματοληψίας και στα στατιστικά του dataset.

6.1.3 Αξιολόγηση του αποθρυβοποιητή

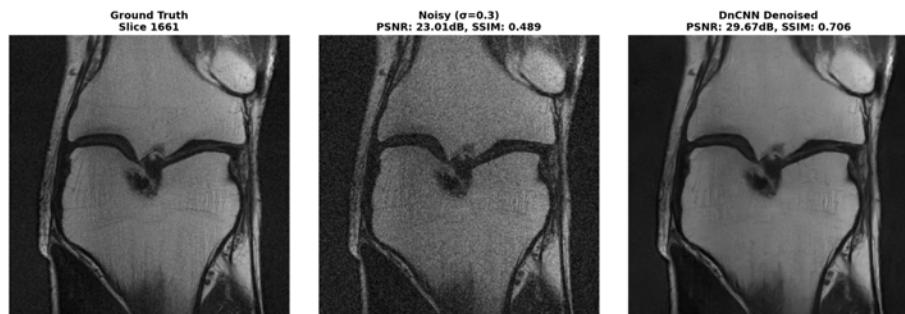
Για να αξιολογήσουμε συστηματικά τη συμπεριφορά του R_θ εκτός του βρόχου Ισορροπίας, εκπαιδεύουμε τρεις DnCNN denoisers σε τρία διακριτά επίπεδα λευκού Gaussian θορύβου στο πεδίο εικόνας, με $\sigma \in \{0.30, 0.10, 0.01\}$. Για την εκπαίδευση, χρησιμοποιήσαμε συνολικά training size 2000 διαφορετικών slices από τα διάφορα single coil knee .h5 αρχεία του singlecoil train fastMRI Dataset[47] ακολουθούμενοι το preprocessing του MultiSlice Approach που εξηγήθηκε στο Κεφάλαιο 2. Τα στοιχεία αυτά αποτελούν διαφορετικές τομές γονάτων στα οποία προσθέθηκε ο θόρυβος.

Για $\sigma = 0.30$ επιβαρύνουμε έντονα τις τομές $\tilde{x} = x + \sigma \varepsilon$ ώστε να ελέγξουμε αν ο denoiser αφαιρεί μαζικά υψηλές συχνότητες χωρίς να υστιάζει ανατομικές δομές. Στη συνέχεια επαναλαμβάνουμε με $\sigma = 0.10$, που προσεγγίζει το SNR του pretraining που θα χρησιμοποιηθεί ως warm start, διατηρώντας τις ίδιες υπερπαραμέτρους οι οποίες ορίζονται ως 80 epochs, batch size = 8, Adam με $lr = 10^{-3}$, Schedule Step=20 και lr gamma=0.6. Τέλος, εκπαιδεύουμε αντίστοιχα στο $\sigma = 0.01$, τιμή που ταυτίζεται με το επίπεδο πολύ χαμηλού θορύβου, το οποίο ωστόσο αντιστοιχεί στον θόρυβο που προσθέθηκε στο επόμενο βήμα του DE-Prox.

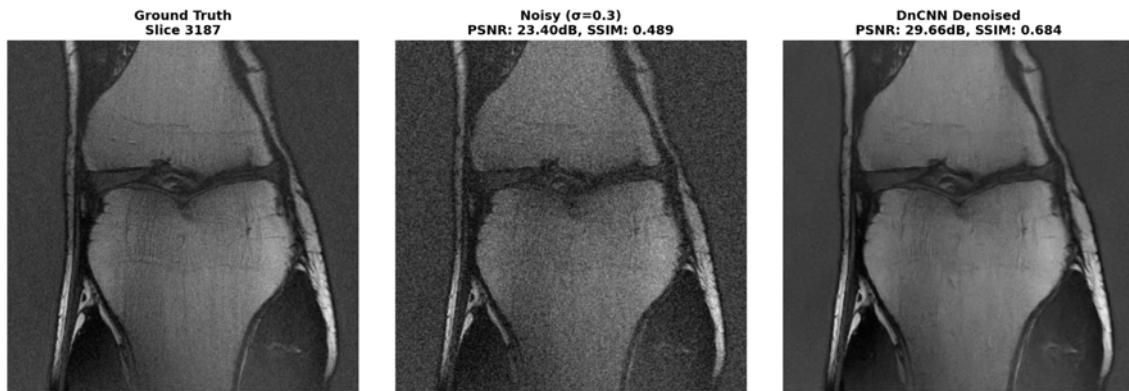
Μεθοδολογικά κρατάμε αυστηρά σταθερό το πειραματικό πρωτόκολλο ανά σ: ίδια δειγματοληψία από το fastMRI[48], ίδια κανονικοποίηση εικόνων, ίδιο πρόγραμμα εκπαίδευσης και το ίδιο DnCNN (17 στρώσεις με πυρήνες 3×3 , spectral normalization σε όλα τα Conv). Τα αποτελέσματα συνοψίζονται οπτικά στις συνοδευτικές απεικονίσεις: στο $\sigma = 0.30$ ο denoiser απομακρύνει εμφανή κόκκο και αυξάνει το PSNR κατά αρκετά dB, στο $\sigma = 0.10$ η βελτίωση παραμένει σταθερή αλλά η οπτική διαφορά είναι ηπιότερη, ενώ στο $\sigma = 0.01$ ο denoiser παραμένει ουσιαστικά near-identity όπως καταδεικνύουν οι μετρικές. Σε συνδυασμό με τη χαμηλή απώλεια, αυτό μας παρέχει βάρη έτοιμα για fine-tuning στο DE-Prox με το ίδιο στατιστικό καθεστώς θορύβου.

Παράμετρος	Τύπος	Προεπιλογή / Σχόλια
n_epochs		80
batch_size	int	8
and_maxiters		100
and_beta	float	1.0
and_m	int	5
lr	float	0.001
etainit	float	0.5
lr_gamma	float	0.6 (Less aggressive decay)
sched_step	int	20
acceleration	float	8.0
noise_sigma	float	0.01,0.1,0.3 (Learnable noise level)

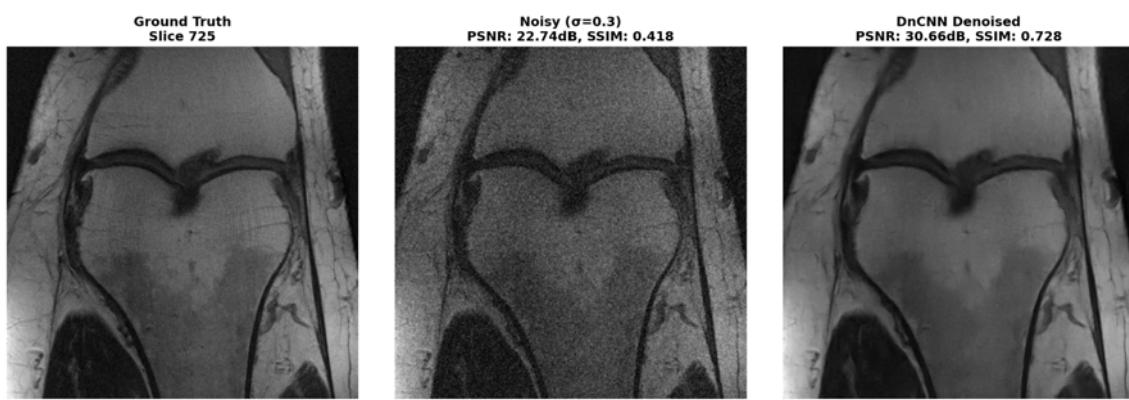
Τα παρακάτω αποτελέσματα προέρχονται από 200 random slices του φακέλου singlecoil val, δηλαδή από δεδομένα εκτός εκπαίδευσης που παρέχει το dataset έχοντας αντλήσει όσα χρειαζόμαστε. Σε κάθε slice προσθέτουμε λευκό Gaussian θόρυβο με στάθμη $\sigma \in \{0.30, 0.10, 0.01\}$ και εφαρμόζουμε τον προεκπαίδευμένο DnCNN χωρίς περαιτέρω fine tuning, ώστε η αξιολόγηση (PSNR/SSIM) να αντανακλά out of sample επίδοση και να αποκλείει διαρροή πληροφορίας από το σύνολο εκπαίδευσης.



Σχήμα 6.1 Παράδειγμα 1 προσθετικού θορύβου τιμής 0.3.



Σχήμα 6.2 Παράδειγμα 2 προσθετικού θορύβου τιμής 0.3

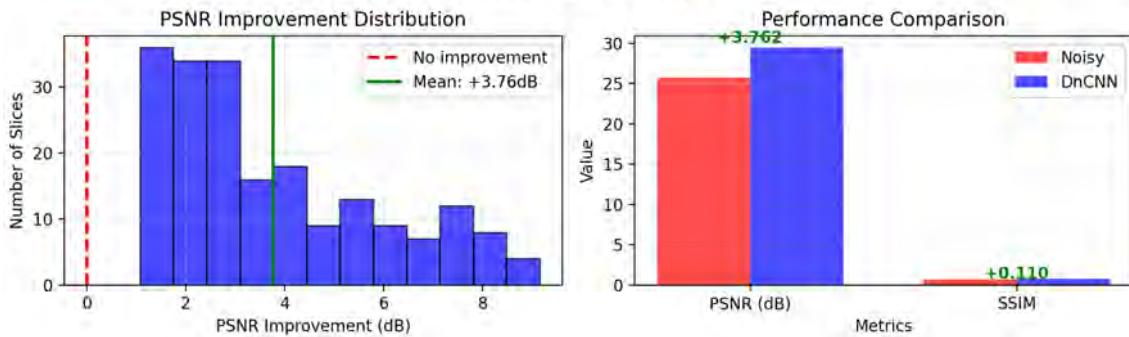


Σχήμα 6.3 Παράδειγμα 3 προσθετικού θορύβου τιμής 0.3

DnCNN DENOISER ASSESSMENT:

- ✓ Mean PSNR Improvement: +3.76 dB
- ✓ Mean SSIM Improvement: +0.110
- ✓ Success Rate: 200/200 slices (100.0%)
- ✓ Correlation (PSNR-SSIM): 0.126
- ✓ Noise Level Tested: $\sigma=0.3$

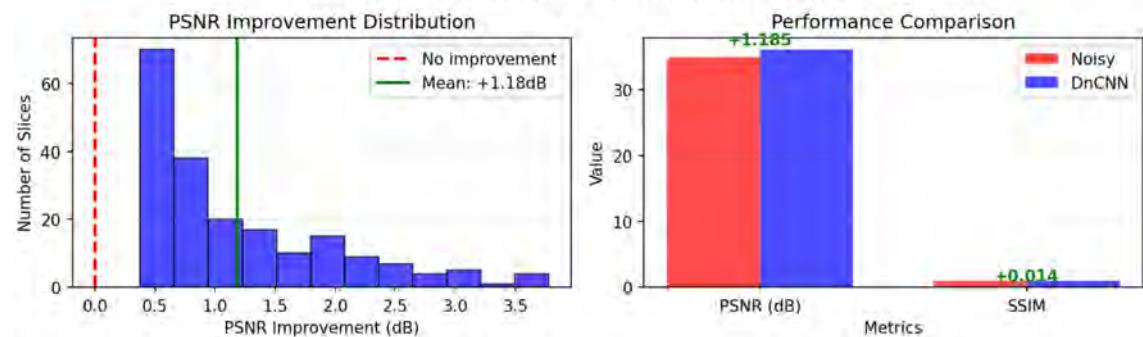
DnCNN Denoiser Validation Summary



DNCNN DENOISER ASSESSMENT:

- ✓ Mean PSNR Improvement: +1.18 dB
- ✓ Mean SSIM Improvement: +0.014
- ✓ Success Rate: 200/200 slices (100.0%)
- ✓ Correlation (PSNR-SSIM): 0.430
- ✓ Noise Level Tested: $\sigma=0.1$

DnCNN Denoiser Validation Summary



Σχήμα 6.8 Παράδειγμα 1 προσθετικού θορύβου τιμής 0.1



Σχήμα 6.9 Παράδειγμα 2 προσθετικού θορύβου τιμής 0.1



Σχήμα 6.10 Παράδειγμα προσθετικού θορύβου τιμής 0.01

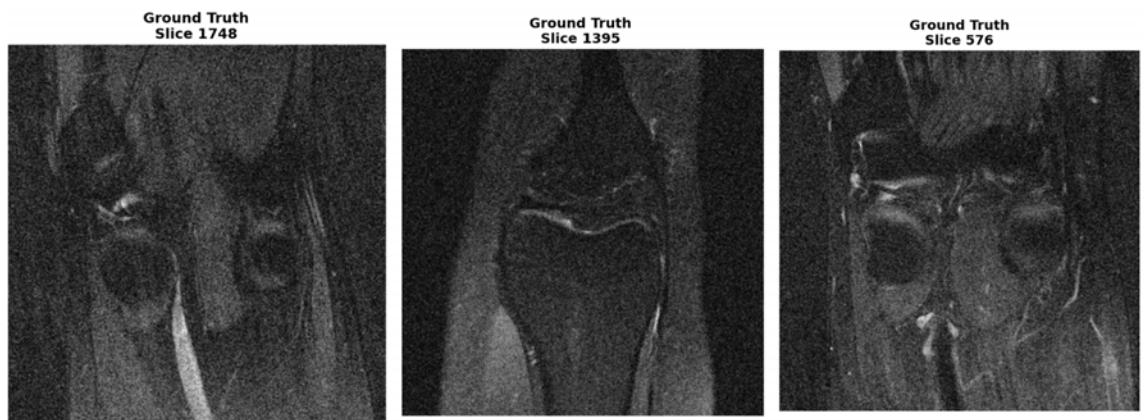
Στα σενάρια θορύβου $\sigma = 0.3, 0.1$ παρατηρούμε την επιτυχία αποθρούβοποίησης του Dncnn μέσω των οπτικών αποτελεσμάτων και των μετρικών μας PSNR, SSIM. Αξίζει πλήρως να σημειωθεί ότι η κατά μεσο όρο μικρή αυξηση του πσυρ και ΣΣΙΜ σε σχέση με μεμονωμένες ακτινογραφίες που παρατηρείται πολύ σημαντικότερη βελτιωση οφειλεται στο γεγονος ότι το dataset παρέχει ακτινογραφίες(slices) οι οποίες οι ιδιες ειναι κουνημένες και με θόρυβο ακομα και μεγαλυτερο των παραπάνω σενάριων. Στο σενάριο με $\sigma = 0.3$ η αποθρούβοποίηση είναι ακόμα πιο αισθητη, λόγος που δημιουργήθηκε και το συγκεκριμένο σενάριο . Στο τελευταίο σενάριο όπου $\sigma = 0.01$, η θορυβώδης είσοδος είναι οπτικά αδιάκριτη από το ground truth και το PSNR δεν μεταβάλλεται αισθητά και ο χάρτης R_θ μαθαίνει συμπεριφορά κοντά στην ταυτότητα σε υψηλό SNR, κάτι που είναι επιμυμητό όταν θα δράσει πάνω στο $v = x + \eta A^\top(y - Ax)$ μέσα στο DE-Prox (μικρές, ελεγχόμενες διορθώσεις αντί για επιθετικές μεταβολές).

6.2 Υλοποίηση του μοντέλου Ισορροπίας

Για το επόμενο στάδιο ακολουθώντας το στήσιμο του paper για το πρόβλημα της ανακατασκευής υιοθετούμε έναν denoiser που εκπαιδεύτηκε στο $\sigma = 0.1$ για το στάδιο Ισορροπίας, διότι ευθυγραμμίζεται με το noise model του επόμενου βήματος (complex στο k -space με τυπική απόκλιση 10^{-2}), σέβεται τη φιλοσοφία του learned proximal (μικρό $\text{Lip}(R_\theta - I)$, ελεγχόμενα Jacobians λόγω spectral normalization) και επιτυγχάνει Ισορροπία μεταξύ διατήρησης λεπτομερειών και ήπιων διορθώσεων που θα δέσουν με το βήμα ορος συνέπειας δεδομένων μέσα στο DE-Prox. Ετσι έχουμε έναν προεκπαιδευμένο με βάρη denoiser ο οποίος λειτουργεί ως θεωρητικό starting point για τους αναγκαίους λόγους ενός equilibrium οπου εξηγήθηκαν στο Κεφάλαιο 5.

Πριν από τους αριθμούς, θέτουμε το πλαίσιο των διαφορών έναντι της βιβλιογραφίας από την οποία αντλήθηκε η μεθοδολογική ιδέα[41]. Πρώτον, εκπαιδεύσαμε με περίπου 1000 slices, δηλαδή στο μισό μέγεθος δείγματος από αυτό που χρησιμοποιη-

ήθηκε στο αρχικό έργο που εισήγαγε τα βασικά αποτελέσματα λόγω χαμηλότερης υπολογιστικής ισχύς. Δεύτερον, το pre-selection των τομών διαφέρει: ο δικός μας dataloader δεν φιλτράρει επιθετικά ύφορυβάδεις ή ελαφρώς κουνημένες τομές ούτε επιβάλλει αυστηρή ομοιομορφία στο σχήμα του γόνατος. Σε αρκετές περιπτώσεις το ίδιο το “ground truth” φέρει αισθητό ύφορυβο ή μικρές κινήσεις. Αυτό δημιουργεί εκ προοιμίου μια ασυμφωνία ανάμεσα στην αποθρυβοποιημένη ανακατασκευή και στη ύφορυβή αναφορά, με συνέπεια το PSNR/SSIM να είναι συντηρητικά ακόμη κι όταν η οπτική ποιότητα βελτιώνεται ουσιαστικά. Υπό αυτές τις συνθήκες, το να διατηρούμε συστηματικά κέρδος έναντι της zero-filled βάσης είναι πιο δύσκολο και, ως εκ τούτου, πιο ενδεικτικό της ισχύος της μεθόδου.

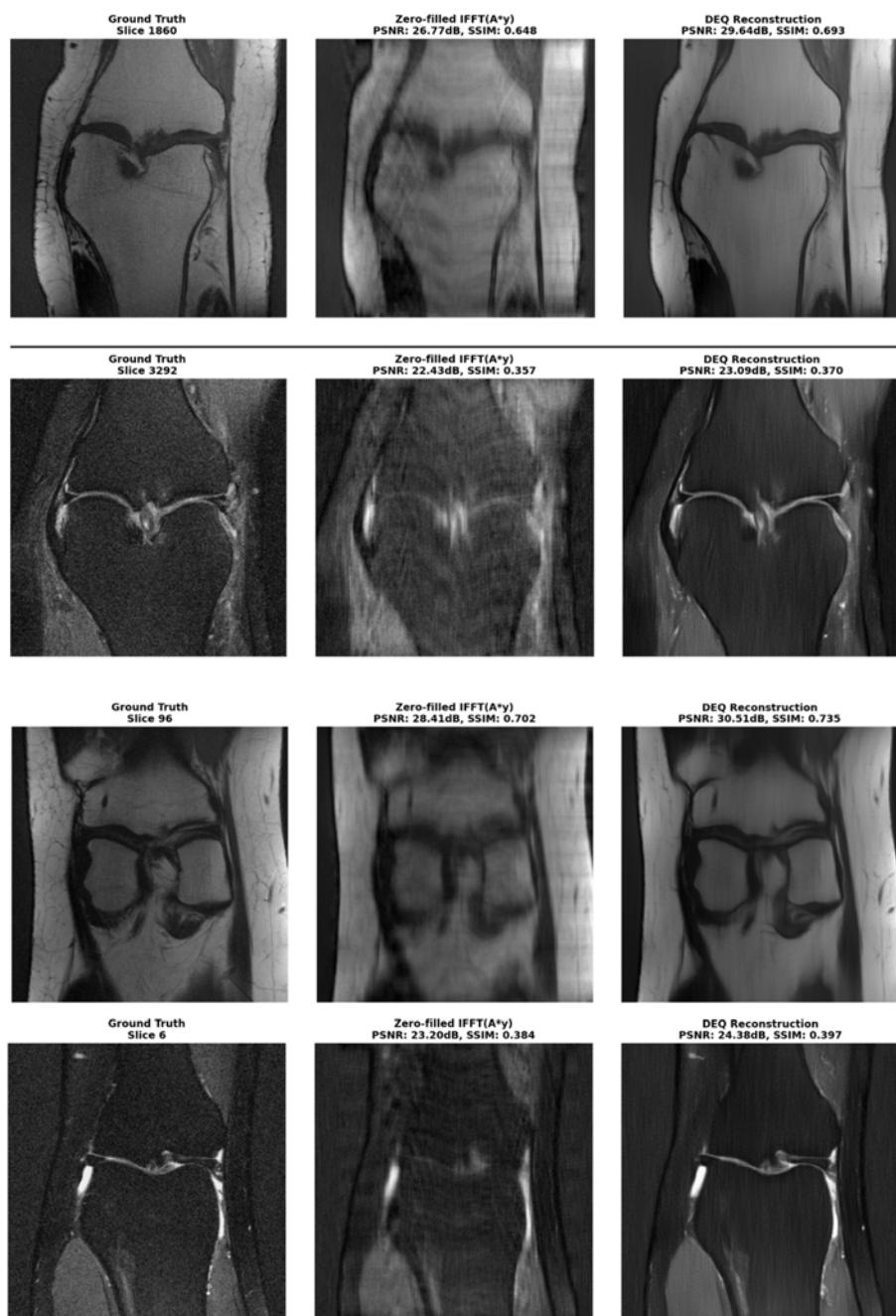


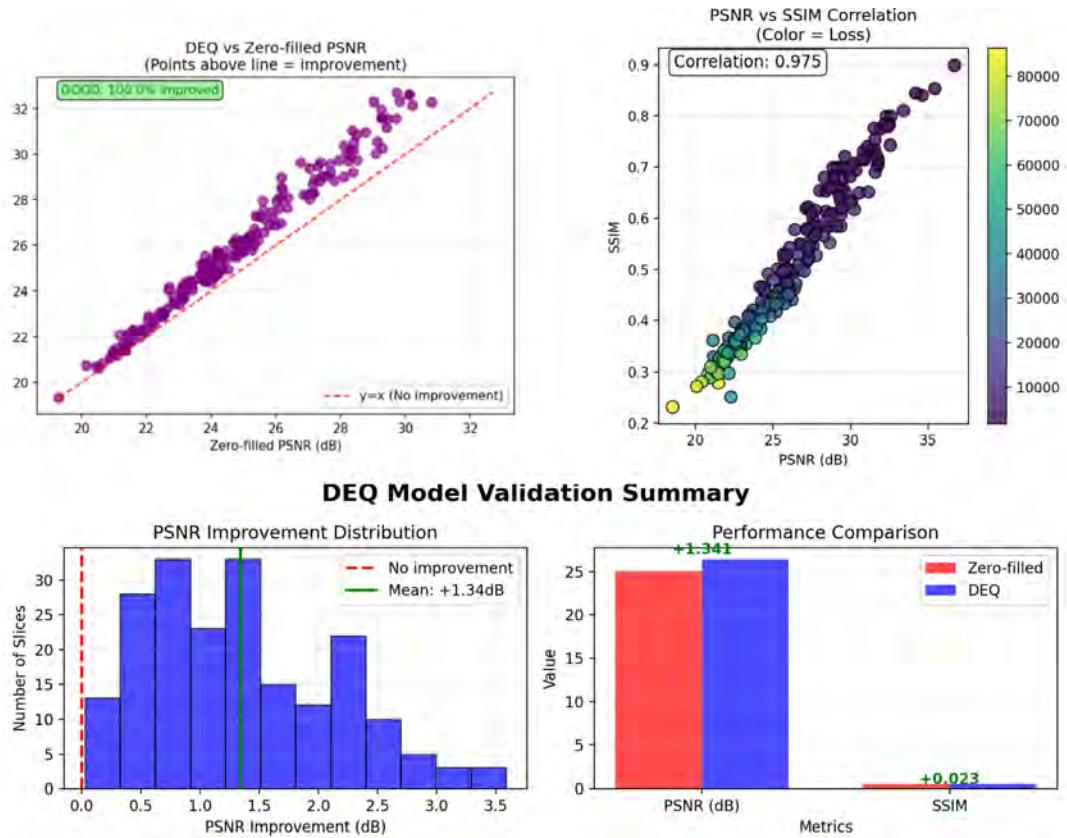
Σχήμα 6.11 Παραδείγματα τομών προσέγγισης μονού πηνίου που χρησιμοποιήθηκαν στην εκπαίδευση(οι ίδιες οι ground truths παρέχονται με εκτενή θόρυβο)

Αξιολογήσαμε την πλήρη υλοποίηση DE Prox σε fastMRI single coil knee με χαρτεσιανές μάσκες υποδειγματοληψίας $8 \times$ (fully-sampled κέντρο και μεταβλητή πυκνότητα στις υπόλοιπες γραμμές). Το σύστημα λειτουργεί αξιόπιστα και σταθερά, επιτυγχάνοντας συνεπή βελτίωση έναντι της zero-filled ανακατασκευής τόσο με βάση τις μετρικές όσο και οπτικά. Προχωρώντας στα ποσοτικά, το σύνολο επικύρωσης δείχνει συνεπή υπεροχή του DEQ σε κάθε τομή. Η κατανομή των κερδών σε PSNR είναι μονοχρωφική με ουρά έως 3 dB και μέση βελτίωση περίπου +1.13 dB, ενώ ο μέσος όρος στο SSIM αυξάνεται κατά 0.015. Σε διάγραμμα διασποράς DEQ-έναντι-zero-filled όλα τα σημεία βρίσκονται πάνω από τη $y = x$ ευθεία zero filled. Η συσχέτιση PSNR-SSIM είναι ψευδής και υψηλή (0.97), ενώ η απώλεια του training συσχετίζεται αρνητικά με το PSNR (0.81), ένα υγιές sanity check. Το εύρος PSNR των τομών (περίπου 19–33 dB) είναι πλήρως ρεαλιστικό για knee MRI με $8 \times$ υποδειγματοληψία και επιβεβαιώνει ότι δεν αξιολογούμε σε εύκολο σενάριο.

Τα ενδεικτικά παραδείγματα ενισχύουν την εικόνα αυτή. Στην τομή 6, το PSNR αυξάνεται από 23.2 dB σε 24.38 dB, με ορατή μείωση aliasing και καθαρότερη

οριοθέτηση, συνοδευόμενη όμως από λεπτές κατακόρυφες ραβδώσεις που παραπέμπουν σε υπολειμματικό streaking αυτό συμβαίνει συχνά με καρτεσιανές μάσκες και μπορεί να εντείνεται όταν το ορος συνέπειας δεδομένων βήμα είναι «επιθετικό». Βλέπουμε μεμονωμένες, οι οποίες παρέχουν ως PSNR 30.51 dB και η βελτίωση αντανακλάται και στην ίδια την εικόνα. Σε πλήθος άλλων τομών η οπτική βελτίωση είναι εμφανής, ακόμη και όταν η αύξηση στο SSIM είναι μετριοπαθής, κάτι αναμενόμενο όταν το ground truth φέρει ύφορυβο και η δική μας ανακατασκευή είναι πιο καθαρή από την ground truth. Αυτό αντιστοιχεί και στο γεγονός ότι έχουμε κάποιες υψηλές τιμές σε μεμονωμένες ανακατασκευές ενώ σε μερικές λόγω ground truth οι μέσες τιμές των γραφημάτων είναι χαμηλότερες.





Ο ορος συνέπειας δεδομένων επαναφέρει ότι μετρήθηκε, ενώ ο προεχπαιδευμένος DnCNN ως R_θ λειτουργεί ως learned proximal που εξαλείφει όσα υπάρχουν στο nullspace με στατιστικά εύλογες δομές. Η χρήση ενός διαφορετικού denoiser σε διαφορετικό είδος ύστορυβου θα ήταν ίσως καταλληλότερη λόγω του ήδη υπάρχων μεγαλύτερων ύστορυβων σε αρχικές ακτινογραφίες. Σημειώνουμε επίσης ότι σε περιβάλλον όπου η αναφορά είναι ύστορυβώδης, οι μετρικές πιστότητας τύπου PSNR/SSIM τείνουν να υποτιμούν μεθόδους που αποθρύβουν, επειδή απομακρύνουν την έξοδο από τη ύστορυβώδη ground truth. Παρατηρούμε καθαρό μέσο κέρδος και βελτίωση ανά τομή.

Συνοψίζοντας, το σύστημα Ισορροπίας με DE-Prox και DnCNN επιδεικνύει λειτουργικότητα και αναβαθμίσεις ποιότητας έναντι της zero-filled βάσης σε ένα απαιτητικό καθεστώς δεδομένων (ιδιαίτερα στο δύσκολο σενάριο του 8x acceleration). Οι περιορισμοί στο μέγεθος εκπαίδευσης, στο μικρότερο learning rate λόγω υπολογιστικής ισχύς και στη καθαρότητα του συνόλου εκπαίδευσης αποτελούν τους πρακτικούς μας περιορισμούς. Υπό πιο ομοιογενή και λιγότερο ύστορυβώδη ground-truth, μεγαλύτερο σύνολο εκπαίδευσης και learning rate τα κέρδη τείνουν να είναι μεγαλύτερα, κάτι που ενισχύει τη θέση μας ότι τα παρόντα νούμερα αποτελούν ισχυρή, αλλά ακόμη και συγκρατημένη, εκτίμηση της πραγματικής ωφέλειας της μεθόδου.

Κεφάλαιο 7

Συνάρτηση Απώλειας \perp loss

7.1 Εισαγωγή

Στην ανακατασκευή MRI τα pixel είναι μιγαδικές ποσότητες και επομένως αντιστοιχούν σε διανύσματα στο επίπεδο με δύο θεμελιώδεις ιδιότητες: μέτρο, που αποτυπώνει την ένταση του σήματος, και φάση, που αποτυπώνει τον προσανατολισμό του. Ένα βασικό ζητούμενο είναι η ταυτόχρονη ανάκτηση σωστού μήκους και σωστής κατεύθυνσης για κάθε μιγαδικό διάνυσμα. Αν γράψουμε την ground truth και την πρόβλεψη ως $y = c + i d$ και $\hat{y} = a + i b$, τότε το μέτρο και η φάση δίνονται από $|y| = \sqrt{c^2 + d^2}$, $|\hat{y}| = \sqrt{a^2 + b^2}$, $\varphi = \arg(y)$, $\hat{\varphi} = \arg(\hat{y})$. Η απόκλιση φάσης $\Delta\varphi = \hat{\varphi} - \varphi$ και το σφάλμα στο μήκος $|\hat{y}| - |y|$ αποτελούν δύο διακριτές πηγές λάθους που πρέπει να αντιμετωπιστούν ρητά.

Η κλασική συνάρτηση κόστους MSE, όταν εφαρμόζεται στο μιγαδικό πεδίο ως ευκλείδεια απόσταση ανάμεσα σε πραγματικό και φανταστικό μέρος, εισάγει μεριληφία υπέρ της συρρίκνωσης του μέτρου όταν υπάρχει φασική απόκλιση[53]. Η σχέση

$$\text{MSE}(\hat{y}, y) = |\hat{y}|^2 + |y|^2 - 2|\hat{y}||y| \cos(\Delta\varphi)$$

δείχνει ότι, αν παγώσουμε τη γωνία $\Delta\varphi$, το ελάχιστο ως προς το μήκος της πρόβλεψης προκύπτει στο $|\hat{y}|^* = |y| \cos(\Delta\varphi)$, που είναι μικρότερο από $|y|$ για κάθε $\Delta\varphi \neq 0$. Άρα ο MSE προτιμά να μειώνει την ένταση αντί να διορθώνει τη φάση, με αποτέλεσμα συστηματική υποεκτίμηση του μέτρου και απώλεια λεπτομέρειας.

Η προτεινόμενη συνάρτηση κόστους perp-loss αντιμετωπίζει αυτή την ασυμμετρία εισάγοντας ένα γωνιακό σκέλος[53] που μετρά καθαρά την απόκλιση κατεύθυνσης, ανεξάρτητα από την κλίμακα της πρόβλεψης, και ένα μετρικό σκέλος που μετρά συμμετρικά το σφάλμα στο μήκος. Για το γωνιακό μέρος αξιοποιούμε τον εμβαδικό όρο $|a d - b c| = |\hat{y}||y||\sin(\Delta\varphi)|$, ο οποίος είναι ανάλογος με την κάθετη απόσταση μεταξύ των δύο διανυσμάτων. Κανονικοποιώντας ως προς το μήκος της πρόβλεψης, ορίζουμε

$$\ell_{\perp}(\hat{y}, y) = \frac{|a d - b c|}{|\hat{y}| + \varepsilon},$$

όπου $\varepsilon > 0$ διασφαλίζει αριθμητική σταθερότητα όταν $|\hat{y}|$ είναι πολύ μικρό. Με αυτόν τον ορισμό, για σταθερή $\Delta\varphi$ το κόστος δεν αλλάζει αν αλλάζει η κλίμακα της πρόβλεψης και συνεπώς δεν υπάρχει κίνητρο συρρίκνωσης.

Για να διατηρήσουμε ομαλή συμπεριφορά σε μεγάλες γωνίες, χρησιμοποιούμε λεία συνέχεια με βάση το συνημίτονο που προκύπτει από το εσωτερικό γινόμενο

$$\cos(\Delta\varphi) = \frac{a c + b d}{|\hat{y}| |y| + \varepsilon} [53], \quad L_{\perp}(\hat{y}, y) = \begin{cases} \ell_{\perp}(\hat{y}, y), & \cos(\Delta\varphi) \geq 0, \\ 2|y| - \ell_{\perp}(\hat{y}, y), & \cos(\Delta\varphi) < 0. \end{cases} [53]$$

Με αυτόν τον ορισμό η απώλεια είναι συνεχής στο 90° και αυξάνει ομαλά έως 180° , τιμωρώντας έντονα την αναστροφή κατεύθυνσης χωρίς να δημιουργεί ασυνέχειες στα παράγωγα.

Για το μήκος υιοθετούμε τον συμμετρικό όρο

$$L_{\text{mag}}(\hat{y}, y) = (|\hat{y}| - |y|)^2,$$

ο οποίος ελαχιστοποιείται στο $|\hat{y}| = |y|$ ανεξάρτητα από τη γωνία και αποκαθιστά την ορθή κλίμακα. Ο πλήρης όρος ανά pixel είναι

$$L_{\text{pixel}}(\hat{y}, y) = L_{\perp}(\hat{y}, y) + \alpha L_{\mu\gamma}(\hat{y}, y), [53]$$

με $\alpha > 0$ να ελέγχει την Ισορροπία ανάμεσα σε φάση και μέτρο. Σε επίπεδο εικόνας αυθοίζουμε ή μεσοποιούμε πάνω στα pixels, διατηρώντας τον ίδιο διαχωρισμό ανάμεσα σε γωνιακή και μετρική ποινή. Η συνδυαστική αυτή διατύπωση ενσωματώνει τη σωστή γεωμετρία των μιγαδικών σημάτων, είναι ομερόληπτη ως προς το μέτρο, ευαίσθητη σε αποκλίσεις φάσης, συνεχής και αριθμητικά σταθερή. Στην πράξη οδηγεί σε ανακατασκευές που ευθυγραφίζουν τη φάση και στη συνέχεια κλειδώνουν το σωστό μήκος, αποφεύγοντας την τάση του MSE.

Με αυτή τη θεωρητική θεμελίωση η συνάρτηση που προτείνουμε ορίζει ρητά τι σημαίνει σωστή κατεύθυνση και σωστό μήκος για κάθε σύνθετο pixel και διαμορφώνει ένα κριτήριο εκπαίδευσης που είναι εναρμονισμένο με τη φυσική του MRI. Στις επόμενες ενότητες αναπτύσσουμε τη μορφή των όρων, αιτιολογούμε τις επιλογές ομαλής συνέχειας και σταθεροποίησης και δείχνουμε πώς ο συνδυασμός τους αποφέρει πρακτικά οφέλη στην ποιότητα ανακατασκευής.

7.2 Η συνάρτηση μέσου τετραγωνικού σφάλματος και η μεροληψία του

Το μέσο τετραγωνικό σφάλμα MSE ανά pixel γράφεται

$$\text{MSE}(\hat{y}, y) = (a - c)^2 + (b - d)^2 = |\hat{y}|^2 + |y|^2 - 2|\hat{y}||y| \cos(\Delta\varphi).$$

Αν παγώσουμε τη γωνιακή απόκλιση και ζητήσουμε το μήκος $|\hat{y}|$ που ελαχιστοποιεί τον MSE, προκύπτει

$$\frac{\partial}{\partial|\hat{y}|} \text{MSE} = 2|\hat{y}| - 2|y| \cos(\Delta\varphi) = 0 \Rightarrow |\hat{y}|^* = |y| \cos(\Delta\varphi).$$

Για οποιαδήποτε μη μηδενική $\Delta\varphi$ ισχύει $\cos(\Delta\varphi) < 1$, άρα $|\hat{y}|^* < |y|$. Αυτό είναι η συρρίκνωση: όταν οι φάσεις δεν ταιριάζουν, ο MSE προτιμά να κονταίνει την ένταση αντί να στρίψει τη φάση στη σωστή κατεύθυνση.

7.3 Περιγραφή της συνάρτηση απώλειας \perp loss

7.3.1 Η γωνιακή συνιστώσα της προτεινόμενης απώλειας

Θέλουμε ένας όρος να μετρά τη γωνιακή συνιστώσα. Παίρνουμε τον εμβαδικό όρο $|ad - bc| = |\hat{y}||y||\sin(\Delta\varphi)|$ και τον κανονικοποιούμε με το μήκος της πρόβλεψης. Ορίζουμε

$$\ell_{\perp}(\hat{y}, y) = \frac{|ad - bc|}{|\hat{y}| + \varepsilon} = \frac{|\hat{y}||y||\sin(\Delta\varphi)|}{|\hat{y}| + \varepsilon} \approx |y||\sin(\Delta\varphi)|, [53]$$

όπου $\varepsilon > 0$ είναι πολύ μικρός αριθμός για αριθμητική σταθερότητα όταν $|\hat{y}|$ είναι σχεδόν μηδενικό. Για σταθερή γωνία ο παραπάνω όρος δεν εξαρτάται ουσιαστικά από $|\hat{y}|$. Έτσι καταργείται το κίνητρο της συρρίκνωσης: το να μικρύνει κανείς το $|\hat{y}|$ δεν βίχνει τον γωνιακό όρο.

Το γωνιακό σκέλος είναι ανεξάρτητο από κλίμακα της πρόβλεψης. Για δεδομένη $\Delta\varphi$ η ℓ_{\perp} δεν αλλάζει όταν αλλάζει το $|\hat{y}|$. Έτσι δεν υπάρχει κίνητρο να μικρύνει κανείς την ένταση για να κρύψει το φασικό σφάλμα. Το μετρικό σκέλος είναι τέλεια συμμετρικό ως προς υπερεκτίμηση και υποεκτίμηση. Για μικρές γωνίες ισχύει $\sin(\Delta\varphi) \approx \Delta\varphi$, άρα το γωνιακό σκέλος είναι περίπου γραμμικό σε $|\Delta\varphi|$ και επομένως πιο ευαίσθητο σε μικρές αποκλίσεις φάσης από ό,τι ένας καθαρός MSE που τιμωρεί τετραγωνικά την απόκλιση

7.3.2 Η λεία συνέχεια για μεγάλες γωνίες

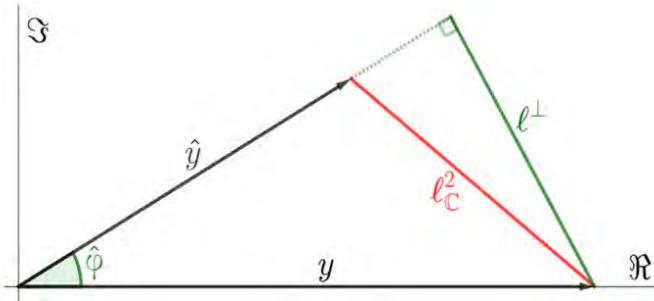
Χρειαζόμαστε ομαλή συμπεριφορά όταν η γωνία περνά τις 90° . Υπολογίζουμε

$$\cos(\Delta\varphi) = \frac{a c + b d}{|\hat{y}| |y| + \varepsilon},$$

και ορίζουμε ένα συνεχές γωνιακό λοσς με δύο κλάδους

$$L_{\perp}(\hat{y}, y) = \begin{cases} \ell_{\perp}(\hat{y}, y), & \cos(\Delta\varphi) \geq 0, \\ 2|y| - \ell_{\perp}(\hat{y}, y), & \cos(\Delta\varphi) < 0. \end{cases}$$

Στο ακριβές 90° οι δύο κλάδοι συμπίπτουν σε $L_{\perp} = |y|$. Καθώς η γωνία πλησιάζει 180° , ο δεύτερος κλάδος τείνει στο $2|y|$, δηλαδή δίνει υψηλό κόστος για πλήρη αναστροφή φάσης. Η χρήση του συνημιτόνου, σε συνδυασμό με αποκοπή στην περιοχή $[-1, 1]$, κρατά την αριθμητική σταθερή.



Σχήμα 7.1 Γεωμετρική απεικόνιση της παλινδρόμησης στο \mathbb{R}^2 . Το εκτιμώμενο διάνυσμα \hat{Y} συγκρίνεται με το διάνυσμα-στόχο Y , με λόγο μετρών $\lambda = |\hat{Y}|/|Y|$ και φασική (γωνιακή) διαφορά $\hat{\varphi}$. Η απώλεια ℓ_2 είναι το μέτρο του κόκκινου διανύσματος. Η προτεινόμενη συνάρτηση κόστους \perp -loss αποτιμά το σφάλμα φάσης $\hat{\varphi}$ ως τη scalar rejection \perp , δηλαδή το μήκος της κάθετης απόστασης από το \hat{Y} προς το Y [53]

7.3.3 Μετρική συνιστώσα που αποκαθιστά το σωστό μήκος

Το δεύτερο κομμάτι μετρά το σφάλμα στο μήκος με συμμετρικό τρόπο ως προς υπερεκτίμηση και υποεκτίμηση

$$L_{\mu\alpha\gamma}(\hat{y}, y) = (|\hat{y}| - |y|)^2.$$

Αυτός ο όρος δεν επηρεάζεται από τη γωνία και εξασφαλίζει ότι, μόλις η κατεύθυνση ευθυγραμμιστεί, η πρόβλεψη θα οδηγηθεί να πάρει το σωστό μέτρο.

7.3.4 Η πλήρης απώλεια και η αμεροληφία ως προς το μέτρο

Σε επίπεδο πιξελ συνδυάζουμε τα δύο σκέλη

$$L_{\pi\xi\epsilon\lambda}(\hat{y}, y) = L_{\perp}(\hat{y}, y) + \alpha L_{\mu\alpha\gamma}(\hat{y}, y),$$

με βάρος $\alpha > 0$ που ρυθμίζει την Ισορροπία. Για μια εικόνα ή ένα batch η συνολική απώλεια είναι άθροισμα ή μέσος όρος πάνω στα πιξελς

$$\mathcal{L} = \sum_{n,p} \left(L_{\perp}(\hat{y}_{n,p}, y_{n,p}) + \alpha (|\hat{y}_{n,p}| - |y_{n,p}|)^2 \right).$$

Αν εξετάσουμε για σταθερή γωνία την ελαχιστοποίηση του $L_{\text{πιξελ}}$ ως προς $|\hat{y}|$, ο γωνιακός όρος είναι ουσιαστικά σταθερός και το μοναδικό εξαρτώμενο μέρος είναι $\alpha(|\hat{y}| - |y|)^2$. Το ελάχιστο είναι στο $|\hat{y}| = |y|$. Άρα η απώλεια είναι αμερόληπτη ως προς το μέτρο, σε αντίθεση με τον MSE όπου το ελάχιστο μετατοπίζεται σε $|y| \cos(\Delta\varphi)$.

7.3.5 Αριθμητική σταθερότητα και ειδικές περιπτώσεις

Ο πολύ μικρός όρος ε μέσα σε ρίζες και παρονομαιστές κρατά τα πάντα ορισμένα όταν $|\hat{y}|$ ή $|y|$ είναι πολύ μικρά, όπως συμβαίνει στο υπόβαθρο.[53]

7.4 Ενσωματωση στην πράξη

Τλοποιούμε την απώλεια πάνω στα δύο κανάλια πραγματικού και φανταστικού της εξόδου και του στόχου. Τπολογίζουμε τα μέτρα $|\hat{y}| = \sqrt{a^2 + b^2 + \varepsilon}$ και $|y| = \sqrt{c^2 + d^2 + \varepsilon}$. Τπολογίζουμε τον εμβαδικό αριθμητή $f = |ad - bc|$ και τον γωνιακό όρο $\ell_{\perp} = f / (|\hat{y}| + \varepsilon)$. Παράλληλα υπολογίζουμε το συνημίτονο της διαφοράς φάσης ως $\cos(\Delta\varphi) = (ac + bd) / (|\hat{y}| |y| + \varepsilon)$, και παίρνουμε τον συνεχές γωνιακό όρο L_{\perp} με τον κατάλληλο κλάδο. Ο μετρικός όρος είναι $(|\hat{y}| - |y|)^2$. Αθροίζουμε σε όλα τα πιξελς, βάζουμε βάρος α στο μετρικό σκέλος και σχηματίζουμε την τελική τιμή της απώλειας. Στον κώδικα μας χρησιμοποιήσαμε μικρό βάρος $\alpha = 0.05$ ώστε να προτεραιοποιείται η διόρθωση της φάσης και να ακολουθεί η εξομάλυνση του μέτρου. Για σταθερότητα κλιμακώσαμε το αποτέλεσμα σε κλίμακα αντίστοιχη με MSE ώστε ο ρυθμός μάθησης να παραμένει χρήσιμος.

Στον forward κάνουμε το εξής: φτιάχνουμε τις μετρήσεις y από το καθαρό target με τον measurement_process, παίρνουμε warm-start $x^{(0)} = A^{\top} y$ και καλούμε το deep_eq_module.forward(y, initial_point). Μέσα εκεί, με torch.no_grad(), τρέχει ο solver (π.χ. Anderson) μέχρι να βρεθεί το σταθερό σημείο z^* του χάρτη $f_{\theta}(z; y) = R_{\theta}(z + \eta A^{\top}(y - Az))$. Έπειτα γίνεται μια επιπλέον εφαρμογή του f_{θ} με ενεργό autograd (δηλαδή ξαναϋπολογίζεται $z = f_{\theta}(z^*, y)$) και ο wrapper στήνει ένα backward_hook: κρατά ένα παγωμένο ζεύγος (z_0, f_0) και δηλώνει ότι, όταν ζητηθούν παράγωγοι προς z , θα λύσει εσωτερικά το implicit γραμμικό σύστημα $(I - \partial_x f_{\theta})^{\top} v = \partial \mathcal{L} / \partial z^*$ με τον ίδιο solver, και θα επιστρέψει το v στο autograd. Μέχρι εδώ δεν έχει χρησιμοποιηθεί καμία loss: ο forward solver είναι αμιγώς γεωμετρία του f_{θ} .

H loss μπαίνει στο επόμενο κομμάτι: υπολογίζουμε έξω από τον solver το $\mathcal{L}(z^*, x^{\text{true}})$ (στο baseline ήταν MSELoss(reduction=sum), τώρα είναι η δική μας perp_loss_complex). Όταν καλέσουμε loss.backward(), το autograd υπολογίζει πρώτα το διάνυσμα $\partial\mathcal{L}/\partial z^*$ και το στέλνει στο backward_hook. Εκεί, ο wrapper λύνει (με Anderson ή ό,τι του δώσαμε) το $(I - \partial_x f_\theta)^\top v = (\partial\mathcal{L}/\partial z^*)^\top$ και επιστρέφει το v . Από εκεί και πέρα το autograd ξέρει πώς να πολλαπλασιάσει το v με τις τοπικές παραγώγους $\partial_\theta f_\theta$ και να παραγάγει τα gradients προς τα βάρη του denoiser R_θ (και προς οποιοδήποτε learnable στον linear operator, αν υπάρχει). Με άλλα λόγια: η loss τροφοδοτεί το backward μέσω του $\partial\mathcal{L}/\partial z^*$, και το implicit solve μετατρέπει αυτό το συνολικό σήμα σε σωστά gradients για τα βάρη χωρίς να χρειάζεται να ξεδιπλώσουμε όλες τις επαναλήψεις του forward.

Άρα αυρό αλλάζει όταν πάμε από MSE σε \perp -loss: Τίποτα στην πορεία του forward solve: συνεχίζουμε να βρίσκουμε z^* με τον ίδιο Anderson/το ίδιο f_θ /το ίδιο η . Αυτό που αλλάζει είναι το διάνυσμα $\partial\mathcal{L}/\partial z^*$ που παράγεται από το χριτήριο: με MSE είναι απλώς διαφορά $(z^* - x^{\text{true}})$ στα δύο κανάλια (real/imag), ενώ με \perp -loss είναι γεωμετρικά ευθυγραμμισμένο με σφάλμα φάσης και μέτρου (δηλαδή το hook λαμβάνει διαφορετικό, πιο καθοδηγητικό gradient). Ο implicit λύτης και το $(I - \partial_x f_\theta)^\top v = \partial\mathcal{L}/\partial z^*$ μένουν ακριβώς τα ίδια· απλώς η δεξιά πλευρά αλλάζει μορφή. Στο training loop αυτό αποτυπώνεται στο ένα σημείο που άλλαξε: παλαιότερα είχαμε lossfunction = torch.nn.MSELoss(reduction=sum), πλεον lossfunction = perp_loss_complex. Ο forward solver δεν βλέπει τη loss, ο backward τη χρησιμοποιεί μία φορά ανά batch για να φτιάξει τα σωστά gradients, και το optimizer.step() ενημερώνει τα βάρη του denoiser.

Στο validation ισχύει το ίδιο: η loss που τυπώνουμε/συνυπολογίζουμε δεν επηρεάζει τη λύση· είναι μετρική αποτίμηση, ενώ η ανακατασκευή προκύπτει από το fixed point. Συνοψίζοντας: στο equilibrium η loss υλοποιείται έξω από τον forward solver, χρησιμοποιείται μόνο για να φτιάξει το σήμα $\partial\mathcal{L}/\partial z^*$ στο backward, και έπειτα μέσω του implicit βήματος επιστρέφει gradients στα βάρη. Παλαιότερα αυτό το \mathcal{L} ήταν MSE (ανά κανάλι), τώρα είναι \perp -loss (φάση+μέτρο). Ο solver, το f_θ , το η , ο Anderson, όλα ίδια· αυτό που αλλάζει είναι η κατεύθυνση με την οποία ωθούμε τον denoiser να βελτιωθεί.

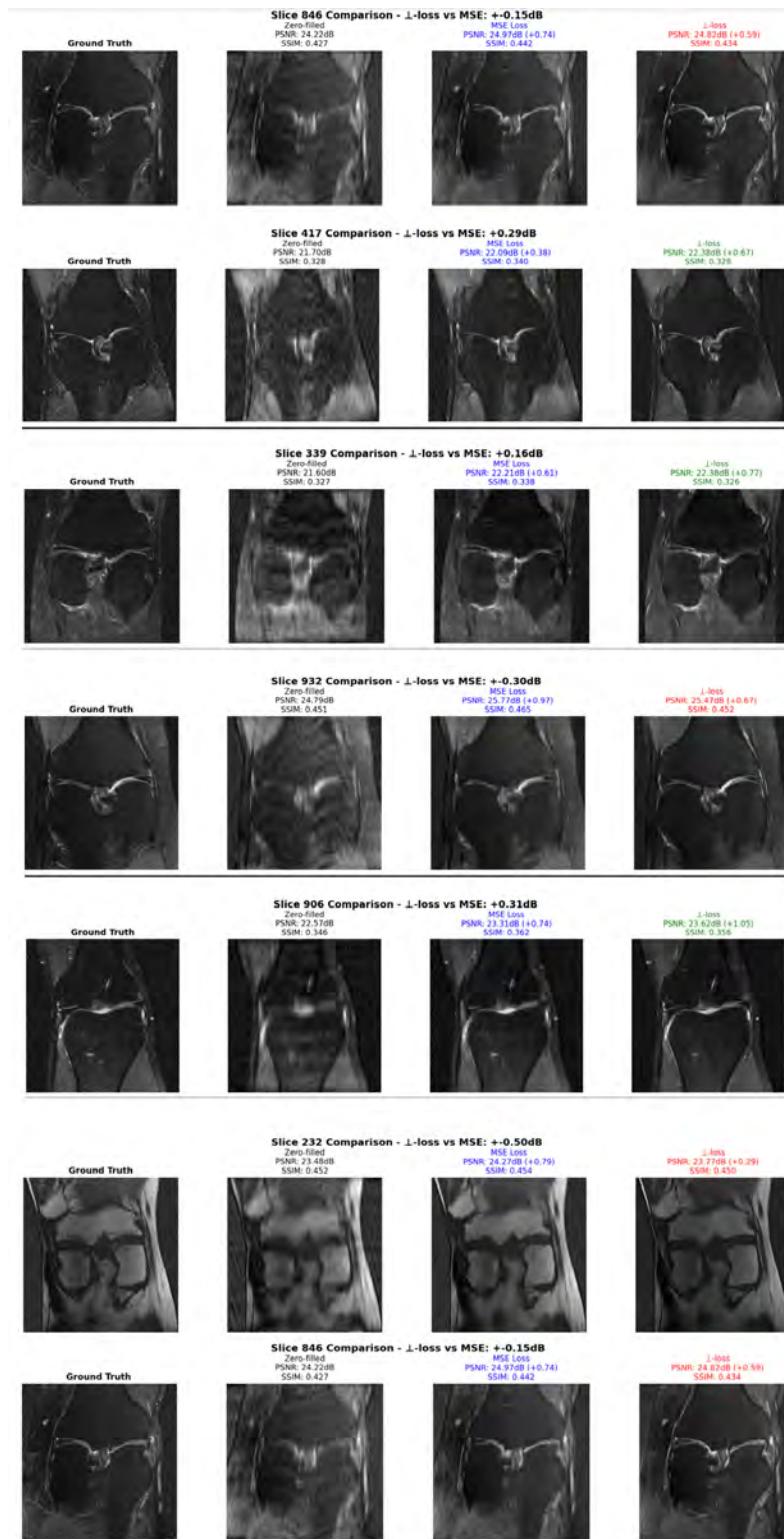
7.4.1 Σχολιασμός αποτελεσμάτων

Πείραμα – ίδιες ρυθμίσεις Η σύγκριση έγινε στο ίδιο ακριβώς πειραματικό καθεστώς για τις δύο συναρτήσεις κόστους (MSE και \perp loss): ίδιος προεκπαίδευμένος denoiser (DnCNN), ίδιο learning rate, batch size, scheduler (γ , step), αριθμός epochs, ίδιος DEQ solver (Anderson) και ίδιο data consistency βήμα. Οπότε, ο forward solve είναι ταυτόσημος· η διαφορά προκύπτει μόνο στο διάνυσμα κλίσεων που επιστρέφεται στο backward μέσω του implicit συστήματος. Η \perp loss που χρησιμοποιήθηκε ήταν ο συνδυασμός

$$\mathcal{L}(\hat{y}, y) = \mathcal{L}_{\perp}(\hat{y}, y) + \alpha, \mathcal{L}_{\text{mag}}(\hat{y}, y) = \mathcal{L}_{\perp}(\hat{y}, y) + \alpha(|\hat{y}| - |y|)^2,$$

με ($\alpha = 1.3$), δηλαδή ελαφρώς αυξημένο βάρος στο μέτρο.

Στης συγκεκριμένη υλοποίηση ξανα, εκπαιδεύσαμε με περίπου 1000 slices, δηλαδή σε σχετικά μικρό αριθμό δείγματος για εκτέλεση πειράματος. Επιπλέον όπως και στα προηγουμενα πειράματα, το pre-selection των τομών διαφέρει: ο δικός μας dataloader δεν φιλτράρει επιθετικά θορυβώδεις ή ελαφρώς κουνημένες τομές ούτε επιβάλλει αυστηρή ομοιομορφία στο σχήμα του γόνατος και σε αρκετές περιπτώσεις το ίδιο το ground truth φέρει αισθητό θόρυβο ή μικρές κινήσεις. Αυτό δημιουργεί εκ προοιμίου μια ασυμφωνία ανάμεσα στην αποθορυβοποιημένη ανακατασκευή και στη θορυβώδη αναφορά, με συνέπεια το μέσο PSNR/SSIM να είναι συντηρητικό ακόμη κι όταν υπάρχουν μεμονωμένες μη θορυβόδεις ανακατασκευές με μεγαλύτερες τιμές όσο η οπική ποιότητα βελτιώνεται .



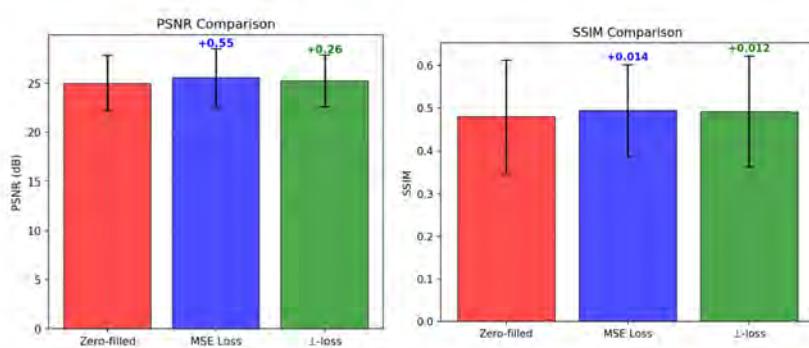
Ποσοτικά αποτελέσματα (validation, $8 \times$ Cartesian)

- Και οι δύο εκδοχές του DEQ υπερτερούν της zero filled βάσης. Από το σύνολο επικύρωσης:

Σε αντιπροσωπευτικά παραδείγματα:

- Slice 339 και Slice 417: (Δ)PSNR (+)0.16 dB και (+)0.29 dB αντίστοιχα υπέρ \perp loss (καλύτερη ευθυγράμμιση λεπτών ορίων).
- Slice 846 και Slice 232: (Δ)PSNR -0.15 dB και -0.50 dB υπέρ MSE (πιο ήπιο φιλτράρισμα/μεγαλύτερη πιστότητα στο θορυβώδες target).

Συμπέρασμα ποσοτικών: Με το συγκεκριμένο ($\alpha = 1.3$) η \perp loss έδωσε σταθερή βελτίωση έναντι zero fill, αλλά λίγο κατώτερη μέση επίδοση από MSE σε PSNR (και πρακτικά ισοδύναμη σε SSIM).



PSNR (μέσος όρος): Zero filled ≈ 25.0 dB, MSE ≈ 25.55 dB (+0.55 dB), \perp loss ≈ 25.26 dB (+0.26 dB).

SSIM (μέσος όρος): Zero filled ≈ 0.48 , MSE $\approx +0.014$, \perp loss $\approx +0.012$ πάνω από το baseline. Τα error bars (τυπικές αποκλίσεις) είναι παρόμοια μεταξύ των δύο losses. Στη σύγκριση MSE vs \perp loss σε PSNR, η κατανομή πλεονεκτήματος δείχνει μέση διαφορά ~ -0.29 dB (υπέρ MSE) και win rate για \perp loss $\approx 29.6\%$ (37/125 τομές). Δηλαδή, η \perp loss είναι συνολικά ελαφρώς ασθενέστερη σε PSNR, παρ' όλα αυτά κερδίζει σε σημαντικό υποσύνολο τομών.

Στο validation σύνολο παρατηρήσαμε μικρή μεν αλλά σταθερή βελτίωση έναντι του zero-filled και για τα δύο κριτήρια, με το MSE να υπερέχει οριακά σε μέση αύξηση PSNR/SSIM, ενώ το L_{\perp} παράγει συγχρίσιμες ή σε μεμονωμένες περιπτώσεις και ανώτερες ανακατασκευές. Συνολικά, τα οπικά αποτελέσματα είναι συνεπή: λιγότερο aliasing, σαφέστερες επιφάνειες χόνδρου και υποχώρηση των έντονων artefacts σε σχέση με το zero-filled, για αμφότερα τα κριτήρια.

Ποιοτική αξιολόγηση

- Σταθερή οπτική βελτίωση έναντι zero filled και για τις δύο απώλειες: καθαρότερα όρια χόνδρου, λιγότερο aliasing, αποθορυβοποίηση υποβάθρου.
- Η \perp loss βελτιώνει συστηματικά την αναγνωσιμότητα (άκρα, μείωση aliasing), ακόμα κι όταν οι μετρικές είναι συγχρατημένες λόγω θορυβωδών GT. Σε μη τεχνητά καθαρά datasets, αυτή η οπτική υπεροχή μπορεί να είναι πιο χρήσιμη κλινικά από μικρές διαφορές στο PSNR.
- Διαφορά μετρικών-οπτικής: όπου το ground truth είναι θορυβώδες ή ελαφρώς κουνημένο, μια αποθορυβοποιημένη ανακατασκευή είναι οπτικά ανώτερη αλλά μπορεί να βαθμολογείται συντηρητικά σε PSNR/SSIM, επειδή απομακρύνεται από τον θορυβώδη στόχο. Αυτό εξηγεί γιατί βλέπουμε προφανείς οπτικές βελτιώσεις με μικρές μόνο αυξήσεις στις μετρικές.

Πηγές απόκλισης & περιορισμοί

- Μικρό training set (~ 1000 slices) για equilibrium εκπαίδευση.
- Διαφορετικό pre selection από το αρχικό έργο: εμείς δεν φιλτράρουμε επιθετικά θορυβώδεις/ελαφρώς κουνημένες τομές ούτε επιβάλλουμε ομοιομορφία στο σχήμα γονάτου. Σε αρκετές περιπτώσεις και το ίδιο το GT είναι θορυβώδες, γεγονός που φέρνει δομική ασυμφωνία με μια καθαρή ανακατασκευή.
- Βάρος μέτρου ($\alpha = 1.3$): η έμφαση στο magnitude ενδέχεται να μετριάζει τα οφέλη της γωνιακής διόρθωσης σε slices όπου η διαφορά φάσης είναι το κυρίαρχο σφάλμα.

Σημείωση: Γιατί μικρά dB μπορεί να φαίνονται οπτικά μεγάλα. Το PSNR είναι λογαριθμικό.

Διαφορά 0.57 dB:

$$\Delta \text{PSNR} = 0.57 = 10 \log_{10} \left(\frac{\text{MSE}_{\text{ZF}}}{\text{MSE}_{\text{MSE}}} \right)$$

$$\Rightarrow \frac{\text{MSE}_{\text{ZF}}}{\text{MSE}_{\text{MSE}}} = 10^{0.057} \approx 1.140$$

$$\Rightarrow \text{MSE}_{\text{MSE}} \approx 0.877 \text{ MSE}_{\text{ZF}} \quad (\sim 12.3\% \text{ μείωση σφάλματος}).$$

Διαφορά 0.89 dB:

$$10^{0.089} \approx 1.228 \Rightarrow \text{MSE}_{\perp} \approx 0.815 \text{ MSE}_{\text{ZF}} \quad (\sim 18.5\% \text{ μείωση})$$

πράγμα που έχει ορατό αποτέλεσμα στα edges και στα aliasing artefacts.

Επίδραση του (α) στον συνδυασμό ($\perp + \text{magnitude}$) Ο πλήρης όρος (εξίσωση πάνω) επιτρέπει να ρυθμίζουμε την Ισορροπία φάσης–μέτρου:

$$\mathcal{L} = \mathcal{L}_{\perp} + \alpha(|\hat{y}| - |y|)^2.$$

Με ($\alpha = 1.3$) δώσαμε παραπάνω βάρος στο μέτρο για σταθερότητα. Τα παραπάνω αποτελέσματα υποδηλώνουν ότι αξίζει σάρωση (α):

- ($\alpha = 1.0$): ισόρροπη στάθμιση, πιθανώς πιο ουδέτερη ως προς PSNR.
- ($\alpha < 1$): έμφαση στη γωνιακή ποινή: αναμένουμε περισσότερη ευθυγράμμιση φάσης και ίσως καλύτερα άκρα, με ρίσκο ελαφρώς μεγαλύτερης διακύμανσης στο μέτρο.

Συμπέρασμα Τα αποτελέσματα επιβεβαιώνουν ότι το equilibrium DEQ δουλεύει και με τις δύο απώλειες. Με ίδιες συνθήκες εκπαίδευσης, ο MSE παραμένει ελαφρώς ανώτερος σε μέσο PSNR, όμως η \perp loss προσφέρει ποιοτικά οφέλη και κερδίζει σε αρκετές τομές, όπως η υποδηλώνει δυναμικό βελτίωσης. Δεδομένων των περιορισμών του dataset και της επιλογής ($\alpha = 1.3$), μια σάρωση διαφορετικών α και επαναπροσδιορισμό απαλλαγμένο από preselection περιορισμούς και συνολικά μεγαλύτερο σύνολο δεδομένων θα αποτυπώσουν πληρέστερα τα αποτελέσματα.

Κεφάλαιο 8

Συμπέρασμα και Μελλοντική Εργασία

8.1 Συμπεράσματα και κατευθύνσεις

Εργαστήκαμε με σύνολο εκπαίδευσης $\sim 10^3$ τομών, με προεπιλογή διαφορετική από τη βασική βιβλιογραφία [41] και με εικόνες αναφοράς που συχνά περιέχουν θόρυβο/μικρές κινήσεις. Μεθοδολογικά, υλοποιήσαμε DE-Prox για μονοπηνιακή MRI με R_θ τύπου DnCNN ως warm start και εκμαθημένο εγγύς τελεστή, Anderson επιτάχυνση, και έμμεση διαφοροποίηση, πιστά στη φιλοσοφία των μοντέλων βαθιάς Ισορροπίας (εκπαίδευση απευθείας στο σταθερό σημείο, έλεγχος προϋπολογισμού επαναλήψεων στη φάση συμπερασμού). Διαμορφώσαμε δίκαιη σύγκριση μέσης τετραγωνικής απόκλισης έναντι απώλειας \perp υπό ταυτόσημες ρυθμίσεις, με αξιολόγηση που παράγει συγκριτικά πλάνα ανά τομή και συνοπτικά στατιστικά PSNR/SSIM με ορθή κανονικοποίηση. Το DE-Prox λύνει αξιόπιστα το σταθερό σημείο με Anderson και έμμεση διαφοροποίηση, εξηγώντας τη σταθερότητα και το μικρό αποτύπωμα μνήμης· οι ταυτόσημες ρυθμίσεις καθιστούν τα αποτελέσματα δίκαια και επαναλήψιμα.

8.2 Περίληψη συνεισφορών και αποτελεσμάτων

Πειραματικά, παρατηρούμε σταθερή υπεροχή έναντι της γραμμής βάσης μηδενικής συμπλήρωσης και με τις δύο συναρτήσεις απώλειας· η μέση τετραγωνική απόκλιση υπερέχει οριακά σε μέσο PSNR έναντι της (\perp), με σχεδόν ισοδύναμο SSIM. Η (\perp) βελτιώνει συστηματικά την ορατότητα ορίων και παράγει καθαρότερες φάσεις, όπως προβλέπεται θεωρητικά (συμμετρία φάσης/μέτρου, αποφυγή μεροληψίας συρρίκνωσης). Τα ευρήματα συνάδουν με τη βιβλιογραφία DEQ/DE-Prox (αξιόπιστη επίλυση σταθερού σημείου, επιτάχυνση Anderson, βελτιώσεις όταν ο κανονικοποιητής ταιριάζει στο πρόβλημα) και εξηγούν γιατί το οπτικό κέρδος μπορεί να υπερβαίνει τα συντηρητικά μεγέθη PSNR/SSIM σε θορυβώδεις εικόνες αναφοράς.

Κατευθύνσεις: πολυπηνιακή επέκταση, προτεραιότητες, εναλλακτικοί στόχοι, κλινική επικύρωση

- (α) Πολυπηνιακή επέκταση: επίλυση με δεδομένα πολλαπλών πηγών ώστε ο εκμαθημένος εγγύς τελεστής να δρά συνεκτικά στο χώρο πηγών· η αρχιτεκτονική DEQ υποστηρίζει φυσικά το σενάριο.
- (β) Δεδομένα και κλινική επικύρωση: καλύτερη προεπιλογή (αφαίρεση υπερθορυβωδών τομών), μεγαλύτερο σύνολο εκπαίδευσης και προοπτική πολυκεντρικής αξιολόγησης· η εκπαίδευση DEQ επιτρέπει αύξηση κόστους χυρίων στο οπίσθιο πέρασμα (έμμεση επίλυση) με μικρό αποτύπωμα μνήμης.
- (γ) Ισχυρότερες προτεραιότητες: για την (\perp) διερεύνηση ισοροπημένης/ενισχυμένης στάθμισης όρων φάσης—μέτρου, π.χ. $\alpha \in \{0.5, 0.8, 1.0\}$, πέρα από την παρούσα υλοποίηση.

Σύνοψη. Τλοποιήσαμε πλήρες, λιτό σύστημα ανακατασκευής μονοπηνιακής MRI βασισμένο στο DE-Prox/μοντέλα Βαθιάς Ισορροπίας (με προεκπαίδευμένο DnCNN, επιτάχυνση Anderson, έμμεση διαφοροποίηση) και το αξιολογήσαμε έναντι μηδενικής συμπλήρωσης και των δύο απωλειών, επιτυγχάνοντας συνεπή—έστω μετριοπαθή—βελτίωση σε PSNR/SSIM και σαφή οπτική ανφιβάθμιση: αναγνωρίζουμε τους περιορισμούς (μικρό σύνολο εκπαίδευσης, ύφορυβώδεις/κουνημένες εικόνες αναφοράς, διαφορετική προεπιλογή και υπολογιστικοί περιορισμοί), αλλά το σχήμα Ισορροπίας λειτουργεί αξιόπιστα και παραμένει επεκτάσιμο (πολυπηνιακή λήψη, ισχυρότερες προτεραιότητες, εναλλακτικοί στόχοι) [41].

Βιβλιογραφία

- [1] Keller, P. *Basic Principles of MR Imaging*. GE Medical Systems, 1988. GE Healthcare primer.
- [2] Radiopaedia.org contributors. “MRI — Radiology Reference Article”. Radiopaedia.org, 2025.
- [3] Jenkinson, M. & Chappell, M. *Introduction to Neuroimaging Analysis*. Oxford Neuroimaging Primers. Oxford University Press, 2018. ISBN: 9780198816300.
- [4] Elster, A. D. “MR vs NMR vs MRI (Q&A)”. MRIquestions.com, 2025.
- [5] Lugauer, F. & Wetzl, J. “Magnetic Resonance Imaging”. A. Maier, S. Steidl, V. Christlein, J. Hornegger (eds.), *Medical Imaging Systems: An Introductory Guide*. Lecture Notes in Computer Science, vol. 11111. Cham: Springer, 2018, pp. 91–118. doi:10.1007/978-3-319-96520-8_6.
- [6] “Magnetic resonance imaging”. Wikipedia, The Free Encyclopedia, 2025.
- [7] Patel, K. K., Khan, M. A., & Kar, A. “Recent developments in applications of MRI techniques for foods and agricultural produce—an overview”. *Journal of Food Science and Technology*, 52 (2015), 1–26. doi:10.1007/s13197-012-0917-3.
- [8] Pradeep, D., Tembhre, M. K., Parihar, A. S., & Rao, C. “Magnetic resonance imaging: Basic principles and advancement in clinical and diagnostics approaches in health care”. *Biomedical Imaging Instrumentation*. Elsevier, 2022, pp. 45–66. doi:10.1016/B978-0-323-85650-8.00005-X.
- [9] N Tubridy , C S McKinstry, ”Neuroradiological history: Sir Joseph Larmor and the basis of MRI physics” doi: 10.1007/s002340000400
- [10] Tryfonopoulos, D. *Brain MRI Reconstruction with Cascade of CNNs and a Learnable Regularization*. M.Sc. Thesis, National Technical University of Athens (NTUA), 2023. doi:10.26240/heal.ntua.25193.
- [11] Gruber, B., Froeling, M., Leiner, T., & Klomp, D. W. J. “RF coils: A practical guide for nonphysicists”. *Journal of Magnetic Resonance Imaging*, 48 (2018), 590–604. doi:10.1002/jmri.26187.

- [12] Puiseux, T., Sewonu, A., Moreno, R., Mendez, S., Nicoud, F. “Numerical simulation of time-resolved 3D phase-contrast magnetic resonance imaging.” DOI: 10.1371/journal.pone.0248816.
- [13] Ding, Z. “Assessing the Impact of Flip Angle and on Image Quality and reliability of Ernst Angle optimization across Varied Conditions in Magnetic Resonance Imaging.”
- [14] Thomen, R. P. “Powerpoint Slides: How MRI Works – Part 1: NMR Basics.” Patreon (thePIRL), 2018..
- [15] Tygert, M. & Zbontar, J. “Simulating single-coil MRI from the responses of multiple coils.”
- [16] Hassibi, A., Babakhani, A., & Hajimiri, A. “A Spectral-Scanning Nuclear Magnetic Resonance Imaging (MRI) Transceiver.”
- [17] Moratal, D., Vallés-Luch, A., Martí-Bonmatí, L., & Brummer, M. E. “k-Space tutorial: an MRI educational tool for a better understanding of k-space.” DOI: 10.2349/bijj.4.1.e15.
- [18] Hargreaves, B. “MR Basics Recap: Signal Encoding & k-Space.” Stanford University Radiology, United States.
- [19] Shen, G., Li, M., Farris, C. W., Anderson, S., & Zhang, X. “Learning to reconstruct accelerated MRI through K-space cold diffusion without noise.” DOI: 10.1038/s41598-024-72820-2.
- [20] Kaipio and Somersalo (2005), Statistical and Computational Inverse Problems, Springer..
- [21] Schönlieb, C.-B. & Shumaylov, Z. “Data-driven approaches to inverse problems.” University of Cambridge. <https://arxiv.org/pdf/2506.11732>.
- [22] Ongie, G., Jalal, A., Metzler, C. A., Baraniuk, R. G., Dimakis, A. G., & Willett, R. “Deep Learning Techniques for Inverse Problems in Imaging.”
- [23] Chen, W., Wipf, D., Wang, Y., Liu, Y., & Wassell, I. J. “Simultaneous Bayesian Sparse Approximation With Structured Sparse Models.”
- [24] Bai, Y., Chen, W., Chen, J., & Guo, W. “Deep learning methods for solving linear inverse problems: Research directions and paradigms.”
- [25] Πετρόπουλος Αναστάσιος. ‘Ανακατασκευή εικόνων υψηλής ανάλυσης με χρήση λεξικών και αραιής αναπαράστασης.’ Διπλωματική Εργασία. Επιβλέπων Καθηγητής: Μπερμπερίδης Κωνσταντίνος.
- [26] Shen, G., Li, M., Farris, C. W., Anderson, S., & Zhang, X. “Learning to reconstruct accelerated MRI through K-space cold diffusion without noise.”
- [27] Cruza, G., Yaman, B., Akçakaya, M., Doneva, M., & Prieto, C. “MRI Reconstruction as an Inverse Problem.”

- [28] Geethanath, S., Reddy, R., Konar, A. S., Imam, S., Sundaresan, R., Ramesh Babu, D. R., & Venkatesan, R. “Compressed Sensing MRI: A Review.”
- [29] CEID / SPC Lab. “Modern approaches for information processing.” SET_–10_11_12_(Sparsity_Aware_Learning_and_Distributed-Federated_–Learning_24–25).pdf.
- [30] Yiasemis, G., Sonke, J.-J., Sánchez, C., & Teuwen, J. “Recurrent Variational Network: A Deep Learning Inverse Problem Solver applied to the task of Accelerated MRI Reconstruction.”
- [31] Arbib, M. A. & Haykin, S. “Similarity between biological and artificial neural networks: illustration comparing biological and artificial neurons.” 2003
- [32] “Machine learning.” Wikipedia, The Free Encyclopedia. Διαθέσιμο: https://en.wikipedia.org/wiki/Machine_learning.
- [33] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. “Convolutional neural networks: an overview and application in radiology.” Insights into Imaging. doi: 10.1007/s13244-018-0639-9.
- [34] Goodfellow, I., Bengio, Y., & Courville, A. *Deep Learning*. MIT Press, 2016. Archived 16 Apr. 2016; retrieved 1 Jun. 2016.
- [35] GeeksforGeeks. “Backpropagation in Neural Network.” Διαθέσιμο: <https://www.geeksforgeeks.org/machine-learning/backpropagation-in-neural-networks/>
- [36] Jadon, S. “Introduction to Different Activation Functions for Deep Learning.” Διαθέσιμο: <https://medium.com/@shrutijadon/survey-on-activation-functions-for-deep-learning-3a2a2a1a2a>
- [37] Kingma, D. P. & Ba, J. “Adam: A Method for Stochastic Optimization.” arXiv preprint, 2014. Διαθέσιμο: <https://arxiv.org/abs/1412.6980>.
- [38] Nielsen, M. *Neural Networks and Deep Learning*, 2015.
- [39] Lamarr Institute. “Deep Learning: How do deep neural networks work?” Διαθέσιμο: <https://lamarr-institute.org/blog/deep-neural-networks/>.
- [40] Monga, V., Li, Y., & Eldar, Y. C. “Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing.” [arXiv:1912.10557](https://arxiv.org/abs/1912.10557).
- [41] Gilton, D., Ongie, G., & Willett, R. “Deep Equilibrium Architectures for Inverse Problems in Imaging.” [arXiv:2102.07944](https://arxiv.org/abs/2102.07944).
- [42] Bai, S., Kolter, J. Z., & Koltun, V. Presented by: Alex Wang, Gary Leung, Sasha Doubov “Deep Equilibrium Models.”. Computer Science University of Toronto.
- [43] Bai, S., Kolter, J. Z., & Koltun, V. “Deep Equilibrium Models.” Διαθέσιμο: <https://arxiv.org/abs/1909.01377>.

- [44] Deep Implicit Layers Tutorial. Διαθέσιμο: <http://implicit-layers-tutorial.org/>.
- [45] Plein, S., Greenwood, J., & Ridgway, J. P. (Εδξ.) *Cardiovascular MR Manual*. Second Edition. Springer, Cham, 2015. ISBN: 978-1-4471-6431-4. doi:10.1007/978-1-4471-6431-4.
- [46] “A Review on Image Reconstruction through MRI k-Space Data.”
- [47] “fastMRI: An Open Dataset and Benchmarks for Accelerated MRI.”
- [48] “fastMRI: A Publicly Available Raw k-Space and DICOM Dataset of Knee Images for Accelerated MR Image Reconstruction Using Machine Learning.”
- [49] Deep Implicit Layers — Neural ODEs, Deep Equilibrium Models, and Beyond. Companion website to the NeurIPS 2020 tutorial created by Zico Kolter, David Duvenaud, and Matt Johnson. Contains notes (via Colab notebooks) Διαθέσιμο: <http://implicit-layers-tutorial.org/>.
- [50] Wikipedia contributors. “Structural similarity index measure.” Διαθέσιμο: https://en.wikipedia.org/wiki/Structural_similarity_index_measure.
- [51] Wikipedia contributors. “Peak signal-to-noise ratio.” Διαθέσιμο: https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio.
- [52] Zhang, K., Zuo, W., Gu, S., & Zhang, L. “Learning Deep CNN Denoiser Prior for Image Restoration.”
- [53] Terpstra, M. L., Maspero, M., Sbrizzi, A., & van den Berg, C. A. T. “ \perp -loss: A symmetric loss function for magnetic resonance imaging reconstruction and image registration with deep learning.”

Παράρτημα A'

Παράρτημα

A'.1 Εξηγηση

Παρατίθενται βασικά σημεία του κώδικα πάνω στην υλοποίηση της διπλωματικής:

A)CartesianSinglecoilMRI.py:

```
class cartesianSinglecoilMRI(LinearOperator):
    def __init__(self, kspace_mask):
        super(cartesianSinglecoilMRI, self).__init__()
        self.register_buffer('mask', tensor=kspace_mask)

    def forward(self, input):
        input = ifftshift(input.permute((0, 2, 3, 1)))
        complex_input = torch.view_as_complex(input)
        kspace = torch.fft.fftn(complex_input, dim=1, norm="ortho")
        kspace = torch.fft.fftn(kspace, dim=2, norm="ortho")
        kspace = fftshift(kspace)
        if self.mask is not None:
            kspace_data = kspace * self.mask + 0.0
            kspace_data = ifftshift(kspace_data)
        return torch.view_as_real(kspace_data)

    def gramian(self, input):
        input = ifftshift(input.permute((0, 2, 3, 1)))
        complex_input = torch.view_as_complex(input)
        kspace = torch.fft.fftn(complex_input, dim=1, norm="ortho")
        kspace = torch.fft.fftn(kspace, dim=2, norm="ortho")
        kspace = fftshift(kspace)
        if self.mask is not None:
            kspace_data = kspace * self.mask + 0.0
            kspace_data = ifftshift(kspace_data)

        kspace_data = torch.fft.ifftn(kspace_data, dim=1, norm="ortho")
        realspace = torch.fft.ifftn(Kspace_data, dim=2, norm="ortho")
        realspace = torch.view_as_real(realspace)

        output = ifftshift(realspace).permute((0,3,1,2))
        return output

    def adjoint(self, input):
        complex_input = torch.view_as_complex(input)
        complex_input = torch.fft.ifftn(complex_input, dim=1, norm="ortho")
        realspace = torch.fft.ifftn(complex_input, dim=2, norm="ortho")

        realspace = torch.view_as_real(realspace)

        output = ifftshift(realspace).permute((0, 3, 1, 2))
        return output
```

B) OUR Multislice approach function:

```

class MultiSliceFastMRIDataloader(Dataset):
    def __init__(self, dataset_location, transform=None, data_indices=None, sketchynormalize=True):
        self.transform = transform
        self.data_directory = dataset_location
        self.fft = forward_models_mri.toKspace()
        self.ifft = forward_models_mri.fromKspace()
        self.sketchynormalize = sketchynormalize

        # Build list of (filename, slice_idx) pairs
        self.slice_pairs = []
        filelist = directory.listdir(dataset_location)
        filelist = [f for f in filelist if f.endswith('.hs')]

        for fname in filelist:
            fpath = os.path.join(dataset_location, fname)
            try:
                with h5py.File(fpath, 'r') as h5f:
                    kspace_data = h5f['kspace']
                    num_slices = kspace_data.shape[0]  # (num slices, H, W)
                    # Restrict to knew-rich middle slices (11/27)
                    start_slice = 11
                    end_slice = min(27, num_slices - 1)  # Don't exceed available slices
                    if end_slice > start_slice:  # Only if we have enough slices
                        for slice_idx in range(start_slice, end_slice + 1):
                            self.slice_pairs.append((fname, slice_idx))
            except Exception as e:
                print(f"Skipping {fname}: {e}")
                continue

        # Apply data_indices if provided (for sampling subset of slices)
        if data_indices is not None:
            self.slice_pairs = [self.slice_pairs[i] for i in data_indices]

    def __len__(self):
        return len(self.slice_pairs)

    def __getitem__(self, item):
        filename, slice_idx = self.slice_pairs[item]
        fpath = os.path.join(self.data_directory, filename)

        with h5py.File(fpath, 'r') as data:
            # Read single slice from multi-slice volume
            kspace_slice = data['kspace'][slice_idx]  # (H, W) complex
            kspace = to_tensor(kspace_slice)  # Convert to (H, W, 2)

        # Apply data_indices if provided (for sampling subset of slices)
        if data_indices is not None:
            self.slice_pairs = [self.slice_pairs[i] for i in data_indices]

    def __len__(self):
        return len(self.slice_pairs)

    def __getitem__(self, item):
        filename, slice_idx = self.slice_pairs[item]
        fpath = os.path.join(self.data_directory, filename)

        with h5py.File(fpath, 'r') as data:
            # Read single slice from multi-slice volume
            kspace_slice = data['kspace'][slice_idx]  # (H, W) complex
            kspace = to_tensor(kspace_slice)  # Convert to (H, W, 2)

        # EXACT same logic as original loader:
        # input_img: crop k-space first, then IFFT
        kspace_cropped = center_crop_slice(kspace, shape=[320, 320])
        input_img = forward_models_mri.ifft2(kspace_cropped).permute((2,0,1))

        # target_img: IFFT full k-space, then crop in image space
        image_space = forward_models_mri.ifft2(kspace)
        target_img = center_crop_slice(image_space, shape=[320, 320]).permute((2,0,1))

        # EXACT same normalization
        target_img, mean, std = normalize_instance(target_img, eps=1e-11)
        target_img = target_img.clamp(-6, 6)

        input_img, mean, std = normalize_instance(input_img, eps=1e-11)
        input_img = input_img.clamp(-6, 6)

        return input_img, target_img

```

C)DnCNN[41]

```

class DnCNN(nn.Module):
    def __init__(self, channels, num_of_layers=17, l1p=1.0):
        super(DnCNN, self).__init__()
        kernel_size = 3
        padding = 1
        features = 64
        layers = []
        layers.append(nn.spectral_norm(nn.Conv2d(in_channels=channels, out_channels=features, kernel_size=kernel_size, padding=padding, bias=False)))
        layers.append(nn.ReLU(inplace=True))
        for _ in range(num_of_layers-2):
            layers.append(nn.spectral_norm(nn.Conv2d(in_channels=features, out_channels=features, kernel_size=kernel_size, padding=padding, bias=False)))
            layers.append(nn.BatchNorm2d(features))
            layers.append(nn.ReLU(inplace=True))
        layers.append(nn.spectral_norm(nn.Conv2d(in_channels=features, out_channels=channels, kernel_size=kernel_size, padding=padding, bias=False)))
        self.dncnn = nn.Sequential(*layers)
    def forward(self, x):
        out = self.dncnn(x)
        return out

```

D) L_{\perp} loss approach[53]:

```

def perp_loss_complex(reconstruction, target, epsilon=1e-8):
    """
    Implements the L-loss from Medical Image Analysis 2022 paper.
    Based on Equations (2), (3), (4), and (5) from the paper.

    Paper: "L-loss: A symmetric loss function for magnetic resonance imaging
    reconstruction and image registration with deep learning"
    """

    # Get real and imaginary parts
    real_recon = reconstruction[:, 0, :, :] # Re(Ŷ)
    imag_recon = reconstruction[:, 1, :, :] # Im(Ŷ)
    real_target = target[:, 0, :, :] # Re(Y)
    imag_target = target[:, 1, :, :] # Im(Y)

    # Calculate magnitude of reconstruction and target
    mag_recon = torch.sqrt(real_recon**2 + imag_recon**2 + epsilon) # |Ŷ|
    mag_target = torch.sqrt(real_target**2 + imag_target**2 + epsilon) # |Y|>

    # Calculate ℓ₁ from Equation (2): scalar rejection
    # This term preserves phase information that MSE completely ignores
    numerator = torch.abs(real_recon * imag_target - imag_recon * real_target)
    l_perp = numerator / (mag_recon + epsilon)

    # Calculate phase difference for smooth continuation (Equation 4)
    # Smart handling of large phase errors (>π/2) that MSE fails on
    dot_product = real_recon * real_target + imag_recon * imag_target
    cos_phase_diff = dot_product / (mag_recon * mag_target + epsilon)
    cos_phase_diff = torch.clamp(cos_phase_diff, -1.0, 1.0) # Numerical robustness
    a = 1.3
    # Smooth continuation from Equation (4) - CRITICAL for stability
    # ℓ₁ = ℓ₁ if |ϕ| < π/2 (cos(ϕ) > 0)
    # ℓ₁ = 2|Y| - ℓ₁ if |ϕ| ≥ π/2 (cos(ϕ) ≤ 0)
    l_perp = torch.where(cos_phase_diff > 0,
                          l_perp,
                          2 * mag_target - l_perp)

    # Calculate unbiased magnitude term (Equation 5)
    # Unlike MSE, this treats over/underestimation symmetrically
    mag_diff = (mag_recon - mag_target)**2

    # Get dimensions for proper averaging (Equation 3)
    batch_size, height, width = real_recon.shape
    total_pixels = batch_size * height * width

    # Apply proper normalization from paper
    perp_loss = torch.sum(l_perp) / total_pixels
    mag_loss = torch.sum(mag_diff) / total_pixels

    # BALANCED: Progressive L-loss weighting for better early learning
    # 0.1 provides substantial L-loss benefits while allowing magnitude learning.
    total_loss = perp_loss + a * mag_loss # 90% L-loss + 10% magnitude for balanced learning

    # Scale to match MSE loss magnitude for training stability with efficiency boost
    # This ensures smooth transition from MSE baseline with accelerated convergence
    return total_loss * total_pixels

```

