

LEARNING-BASED RECONSTRUCTION METHODS FOR MAGNETIC PARTICLE IMAGING

A DISSERTATION SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

By
Alper Güngör
January 2023

Learning-based Reconstruction Methods for Magnetic Particle Imaging

By Alper Güngör

January 2023

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Tolga Çukur(Advisor)

Emine Ülkü Sarıtaş Çukur

İbrahim Aykut Erdem

Ergin Atalar

Abdullah Aydın Alatan

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

LEARNING-BASED RECONSTRUCTION METHODS FOR MAGNETIC PARTICLE IMAGING

Alper Güngör

Ph.D. in Electrical and Electronics Engineering

Advisor: Tolga Çukur

January 2023

Magnetic particle imaging (MPI) is a novel modality for imaging of magnetic nanoparticles (MNP) with high resolution, contrast and frame rate. An inverse problem is usually cased for reconstruction, which requires a time-consuming calibration scan for measuring a system matrix (SM). Previous calibration procedures involve scanning an MNP filled sample with a size that matches desired resolution through field of view. This time-consuming calibration scan which accounts for both system and MNP response imperfections is a critical factor prohibiting its practical use. Moreover, the quality of the reconstructed images heavily depend on the prior information about the MNP distribution as well as the specific reconstruction algorithm, since the inverse problem is highly ill-posed. Previous approaches commonly solve an optimization problem based on the measurement model that iteratively estimates the image while enforcing data consistency in an interleaved fashion. However, while conventional hand-crafted priors do not fully capture the underlying complex features of MPI images, recently proposed learned priors suffer from limited generalization performance. To tackle these issues, we first propose a deep learning based technique for accelerated MPI calibration. The technique utilizes transformers for SM super-resolution (TranSMS) for accelerated calibration of SMs with high signal-to-noise-ratio. For signal-to-noise-ratio efficiency, we propose scanning a low resolution SM with larger MNP sample size. For improved SM estimation, TranSMS leverages the vision transformer to capture global contextual information while utilizing the convolutional module for local high-resolution features. Finally, a novel data-consistency module enforces measurement fidelity. TranSMS is shown to outperform competing methods significantly in terms of both SM recovery and image reconstruction performance. Next, to improve image reconstruction quality, we propose a novel physics-driven deep equilibrium based technique with learned consistency block for MPI (DEQ-MPI). DEQ-MPI embeds deep network operators into iterative optimization procedures for improved modeling of image statistics. Moreover,

DEQ-MPI utilizes learned consistency to better capture the data statistics which helps improve the overall image reconstruction performance. Finally, compared to previous unrolling-based techniques, DEQ-MPI leverages implicit layers which enables training on the converged output. Demonstrations on both simulated and experimental data show that DEQ-MPI significantly improves image quality and reconstruction time over state-of-the-art reconstructions based on hand-crafted or learned priors.



Keywords: Magnetic particle imaging, reconstruction, deep learning, deep equilibrium, implicit, system matrix, calibration, medical imaging.

ÖZET

MANYETİK PARÇACIK GÖRÜNTÜLEME İÇİN ÖĞRENME TABANLI GERİÇATIM TEKNİKLERİ

Alper Güngör

Elektrik ve Elektronik Mühendisliği, Doktora

Tez Danışmanı: Tolga Çukur

Ocak 2023

Manyetik parçacık görüntüleme (MPG), manyetik nanoparçacıkların (MNP) yüksek çözünürlük, kontrast ve kare hızı ile görüntülenmesi için yeni bir yöntemdir. Manyetik parçacık Manyetik parçacık Görüntü geriçatımı için genellikle uzun zaman alan bir kalibrasyon taraması sonucu elde edilen sistem matrisinin (SM) kullanıldığı bir tersine problem söz konusudur. Önceki kalibrasyon yöntemleri, istenen çözünürlükle eşleşen bir boyutta doldurulmuş bir MNP örneğiyle görüş alanının taranmasını içerir. Hem sistem hem de MNP ölçümündeki ideal olmayan durumları içeren bu zaman alıcı kalibrasyon taraması, pratik kullanımını engelleyen kritik bir faktördür. Buna ek olarak, oluşturulan görüntülerin kalitesi, tersine problem kötü konumlanmış olduğu için büyük ölçüde MNP dağılımı hakkındaki önsel bilgilere ve spesifik geriçatım algoritmasına bağlıdır. Önceki yaklaşımlar genellikle yinelemeli olarak görüntüyü tahmin ederken veri tutarlılığını zorlayan ölçüm modeline dayalı bir optimizasyon problemini çözer. Ancak, geleneksel elle tasarlanmış önseller, MPG görüntülerinin altında yatan karmaşık özellikleri tam olarak yakalayamazken, son zamanlarda önerilen öğrenilmiş önseller, sınırlı genelleme performansından muzdariptir. Bu sorunlara çözüm olarak, öncelikle hızlandırılmış MPG kalibrasyonu için derin öğrenmeye dayalı bir teknik öneriyoruz. Teknik, yüksek sinyal gürültü oranına (SGO) sahip SM'lerin hızlandırılmış kalibrasyonu için, dönüştürücüler tabanlı SM süper çözünürlüğüne (TranSMS) dayanır. SGO verimliliği için, daha büyük MNP örnek boyutuna sahip düşük çözünürlüklü bir SM'nin taranması önerilmektedir. İyileştirilmiş SM kestirimi için, TranSMS global bağlamsal bilgileri görüntü dönüştürücüsü yoluyla yakalarken, yerel yüksek çözünürlüklü öznitelikler için evrişimsel modül kullanır. Son olarak, önerilen bir veri tutarlılığı modülü, ölçüm bağımlılığını sağlar. TranSMS'in diğer tekniklere kıyasla hem SM kestirimi hem de müteakip görüntü geriçatımı performansını ciddi seviyede iyileştirdiği gösterilmiştir. Ardından, görüntü geriçatım kalitesini iyileştirmek amacıyla,

MPG için öğrenilmiş veri tutarlılığı bloğu kullanan yeni bir fizik güdümlü derin denge tabanlı teknik önerilmektedir (DEQ-MPI). DEQ-MPI, görüntü istatistiklerinin daha iyi modellenmesi için derin operatörleri yinelemeli optimizasyon prosedürlerine yerleştirir. Ayrıca DEQ-MPI, genel görüntü geriçatım performansını iyileştirmeye yardımcı olan veri istatistiklerini daha iyi yakalamak için öğrenilmiş veri tutarlılığını kullanır. Son olarak, önceki yinelemenin-açılması tabanlı tekniklere kıyasla DEQ-MPI, yakınsanmış çıktı üzerinde eğitimi sağlayan kapalı fonksiyon katmanlardan yararlanır. Hem benzetim hem de deneysel veriler üzerindeki gösterimler, DEQ-MPI'nin, el yapımı veya öğrenilmiş önsellere dayalı tekniklere kıyasla görüntü kalitesini ve geriçatım süresini önemli ölçüde iyileştirdiğini göstermektedir.

Anahtar sözcükler: Manyetik Parçacık Görüntüleme, görüntü oluşturma, derin öğrenme, derin denge, kapalı fonksiyon, sistem matrisi, kalibrasyon, medikal görüntüleme.

Acknowledgement

I am very fortunate to work with my advisor Tolga Çukur. He has been very supportive in every possible way. He has always motivated and encouraged me. He has taught me a lot, and I am still learning from him. I also would like to thank Emine Ulku Saritas. She has helped keep the studies grounded in terms of the magnetic particle imaging domain. I would also like to thank Baris Askin. It was very joyful to work together on these projects. I am looking forward to doing it again.

I would like to thank Aykut Erdem for his insightful comments during thesis tracking committee meetings. His comments helped guide and improve these studies. Moreover, he also helped with all the signatures even though he moved to Istanbul. I also thank the jury members Ergin Atalar and Aydin Alatan for accepting to be on the Ph.D. jury. They both have added new insights to improve the studies through their comments.

I thank Aselsan, my employer for enabling a productive research environment. I also would like to thank my friends and colleagues who made it fun to work and do research. Especially, I would like to thank Mert, Oguzhan, Berkan, Umut, Burak, Alper, Evren, Sefa, Cagri and many others for their support. I thank my friend Oguz for going out of his way to help me get the necessary signatures for completing this thesis. I also would like to thank my supervisor Can Baris Top for motivating, encouraging, and supporting me through these studies. He has been involved in every step although he was very busy.

I would like to first express my deepest gratitude to my beautiful wife, Zülal Özdemir Güngör. She has been by my side from the very beginning of my long journey through academia. I am very lucky to have her full support. It would not be possible to complete this thesis without her. I also want to thank our British Longhair friend, Misty. Although she is new to our lives, she has shown a natural interest in academia and magnetic particle imaging by smelling pictures of magnetic particle imaging scanners. I also would like to thank my mother, father and the rest of the family for supporting me throughout my life.

I acknowledge the support of TÜBİTAK (The Scientific and Technological Research Council of Turkey) within the scope of the 2211/A scholarship.

Contents

1	Introduction	1
1.1	Contributions	3
1.2	Outline	4
2	Magnetic Particle Imaging Basics	5
2.1	Signal Equation	5
2.2	Image Reconstruction Approaches for MPI	8
2.3	Calibration	9
2.3.1	Compressed Sensing-based Acceleration for Calibration . .	10
2.3.2	Super Resolution-based Acceleration for Calibration	11
2.4	Image Reconstruction	12
2.4.1	Algebraic Reconstruction Technique (ART)	12
2.4.2	Alternating Direction Method of Multipliers (ADMM) .	13
2.4.3	Deep Image Prior (DIP)	14
3	Super-Resolution based Accelerated Calibration for Magnetic Particle Imaging	16
3.1	Introduction	17
3.2	Related Work	18
3.3	Theory	20
3.3.1	Magnetic Particle Imaging	20
3.3.2	TranSMS	22
3.4	Methods	26
3.4.1	Datasets	26
3.4.2	TranSMS Implementation	29

3.4.3	Competing Methods	30
3.4.4	Analysis Procedures	31
3.4.5	Quantitative Assessments	33
3.5	Results	34
3.5.1	Ablation Studies	34
3.5.2	Analyses on Simulated Dataset	36
3.5.3	Analyses on Open MPI Dataset	38
3.5.4	Analyses on In-house MPI Dataset	39
3.6	Discussion	40
3.7	Conclusion	44
3.8	Appendix	44
4	Deep Equilibrium based Image Reconstruction for Magnetic Particle Imaging	53
4.1	Introduction	54
4.2	Related Work	56
4.3	Theory	57
4.3.1	MPI Reconstruction	57
4.3.2	DEQ-MPI	60
4.4	Methods	65
4.4.1	Competing Methods	65
4.4.2	MPI Phantoms	66
4.4.3	Experimental Procedures	67
4.4.4	Simulation and Analysis Procedures	68
4.5	Results	69
4.5.1	Ablation Studies	69
4.5.2	Simulated Phantoms	71
4.5.3	Experimental Phantoms	74
4.5.4	Reconstruction Time	76
4.6	Discussion	77
4.7	Conclusion	79
5	Conclusion	80

5.1	Discussions and Future Directions on Accelerated MPI System Matrix Calibration	80
5.2	Discussions and Future Directions on Deep Learning-based Image Reconstruction for MPI	81
A	Code	98



List of Figures

2.1	MPI encoding signal receiving process. (a) A depiction of the received signal for an MNP inside the FFR. The top plot depicts the magnetization of MNPs for a generated H-field. A sinusoidal drive field is applied to an MNP without a field offset. The corresponding magnetization response and the MPI signal, which is proportional to the time derivative of the magnetization, are depicted below. As shown, a high intensity $s(t)$ is received. (b) A depiction of the received signal for an MNP outside the FFR. A sinusoidal drive field is applied to an MNP with a field offset. The corresponding magnetization response and the MPI signal are depicted below. As shown, a negligible $s(t)$ is received.	7
3.1	Overview of prior and proposed techniques for SM calibration. (a) Conventional calibration samples all of the grid points in HR SM. (b) CS calibration randomly subsamples HR SM and uses CS optimization for recovery. (c) Coded calibration scenes samples HR SM using a specially tailored calibration scene and uses CS optimization for recovery. (d) Simulation based SR jointly solves an optimization problem involving both a simulated SM and a sampled LR SM as well as the phantom to be imaged. (e) 2d-SMRnet uniformly subsamples HR SM in a strided fashion and recovers HR SM via a CNN. (f) TranSMS performs LR SM measurements with a larger MNP sample that fills an LR grid, and recovers HR SM via a transformer.	46

- 3.2 **(a)** Schematic of the network architecture in TranSMS. After an initial convolutional layer, encoded feature maps are fed to parallel transformer and convolutional modules. The transformer module extracts contextual representations over longer distances, while the convolutional module captures localized representations. Contextual and localized feature maps are fused and upsampled to the resolution of the HR SM. Finally, a data-consistency module enforces fidelity of recovered HR SM to input LR SM. **(b)** The transformer module contains a cascade of token-embedding and convolutional transformer sub-blocks, and a final upsampler to restore dimensionality. **(c)** The transformer performs convolutional token projection, followed by MHSA and MLP blocks to extract contextual features. 47
- 3.3 **(a)** In-house open-sided FFL scanner. **(b)** The imaging phantom used to acquire in-house MPI data. Cylindrical tubes were prepared with Perimag (Micromod GmbH, Germany) MNPs at a dilution ratio of 1:20 (405 $\mu\text{g}/\text{ml}$ Fe). The length, inner radius and outer radius of the tubes were 20 mm, 2 mm and 4 mm, respectively. 47
- 3.4 Validation performance of TranSMS with (red curve) and without (blue curve) the data-consistency (DC) module. The HR SM recovery performance is reported as nRMSE across training epochs. The upper-right panel displays a zoomed-in portion of nRMSE curves during the initial stages of training. 48
- 3.5 SM recovery performance of TranSMS as a function of the number of SMs included in the training set for SR factors of $2\times$ - $8\times$ 48
- 3.6 HR SM recovery in the simulated dataset for $4\times$ SR using noisy measurements. SF gradient strength and MNP size differed between the training and test sets. SM rows corresponding to different harmonics at 45° are shown. **(a)** Recovered HR SMs are displayed along with the input LR SM, and input strided (str) LR SM. Strided bicubic (BC) interpolation and 2d-SMRnet receive as input the strided LR SM. **(b)** Respective error maps (see colorbar). 49

- 3.7 Phantom images from the simulated dataset at 1 T/m SF gradient strength reconstructed with recovered HR SMs at $4\times$ SR are displayed along with the reference reconstruction using the reference HR SM (top row). Error maps for each method are also displayed (bottom row). The phantom comprised two 22-mm long tubes, the right tube's width fixed at 6 mm, and the left tube's width ranging in 2-6 mm. Inter-tube spacing varied in 8-12 mm. 49
- 3.8 A pre-trained TranSMS model was used to super-resolve a 64×64 HR SM by processing overlapping 16×16 patches in the input 32×32 LR SM. SM rows for different harmonics at 45° are shown. Recovered HR SMs are shown under 30 dB measurement SNR, along with the input LR SM and the reference HR SM. 50
- 3.9 HR SM recovery in Open MPI for $4\times$ SR. Both the SF gradient strength and MNP tracer differed for SMs in training and test sets. The SM rows corresponding to the center slice at different frequency components and receive channels (Ch) are shown (k denotes the respective SM row index). **(a)** Recovered HR SMs and **(b)** respective error maps are displayed along with the input LR SM, input strided LR SM and reference HR SM. 50
- 3.10 Image reconstruction in Open MPI for $4\times$ SR. Phantom images reconstructed with recovered HR SMs are displayed along with the benchmark reconstruction using the reference HR SM (top row). Respective error maps are also displayed (bottom row). 51
- 3.11 HR SM recovery in the in-house MPI dataset for $4\times$ SR. The SM rows corresponding to different harmonics at 81° are shown. **(a)** Recovered HR SMs and **(b)** respective error maps for each competing method are displayed along with the input LR SM, input strided LR SM and reference HR SM. 51
- 3.12 Phantom images in the in-house dataset reconstructed with recovered HR SMs at $4\times$ SR and respective error maps are displayed, along with the benchmark reconstruction using the reference HR SM. The length, inner radius and outer radius of the tubes were 20 mm, 2 mm and 4 mm, respectively. 51

- 3.13 Prospective versus retrospective HR SM recovery in the in-house dataset for $2\times$ SR. Real and imaginary parts of the SM rows corresponding to different harmonics ((**a**) – (**c**)) at 0° are shown. 52
- 3.14 Phantom images for the in-house dataset based on prospectively versus retrospectively recovered HR SMs are shown along with the benchmark reconstruction using the reference HR SM. The upper tube is 23-mm long, the lower tube is 21-mm long with a 5-mm-long central stenosis region of 1-mm inner radius (1 pixel : 2 mm). 52
- 4.1 (**a**) Conventional unrolled methods versus the deep equilibrium model in DEQ-MPI. Unrolled methods express reconstruction as repeated projections through an explicit operator, $\mathbf{x}_{k+1}=h_\theta(\mathbf{x}_k;\mathbf{y},\mathbf{A})$ where $\mathbf{x}_k$ is the image at iteration k , $\mathbf{y}$ are acquired data, $\mathbf{A}$ is the system matrix, and network parameters θ are shared across iterations. h_θ is trained to maximize performance after N_{it} iterations where N_{it} is fixed and small to limit computational burden, resulting in suboptimal performance. DEQ-MPI instead leverages an implicit mapping $\mathbf{x}_*=h_\theta(\mathbf{x}_*;\mathbf{y},\mathbf{A})$ to compute a convergent output based on repeated injection of acquired data. In this case, h_θ is trained to maximize image quality upon convergence as opposed to an adhoc N_{it} . (**b**) Proposed DEQ-MPI implementation. DEQ-MPI integrates the implicit mapping into an ADMM algorithm with fixed-point iterations expressed in Eq. (4.14) for the image $\mathbf{x}$ and Lagrange multipliers $\mathbf{d}^{(0)}, \mathbf{d}^{(1)}$. Each iteration involves projection through a learned regularization block (Ψ_{RDN}), projection through a learned consistency block (Ψ_{LC}), and reconciliation in a least-squares step to compute the output $\mathbf{x}_{k+1}$ 58
- 4.2 (a) Open-sided ASELSAN FFL scanner. (b) A cylindrical phantom with two parallel tubes was filled with Perimag MNPs at a dilution of 1:20. (c) A Y-shape phantom was filled with Perimag at a dilution of 1:100 and contained a central air bubble. 67

4.3	Comparison of DEQ-MPI against an end-to-end model that used an RDN block and omitted DC, an unrolled model with $N_{it}=5$ iterations, and an LC-ablated variant based on an unlearned DC block. Separate models were trained at measurement SNRs of 5-45 dB, testing was performed under 35 dB SNR. Average pSNR across the test set is shown for each model.	70
4.4	Convergence behaviors of DEQ-MPI, unrolled model, and LC-ablated variant. Models were trained and tested at 35 dB SNR. Average pSNR across the test set is shown for each model. The zoom-in window highlights performance during initial iterations.	71
4.5	Representative reconstructions of three simulated phantoms and respective error maps (see colorbar) are shown for competing methods, along with the reference images. A measurement SNR of 35 dB was used. The SM at 0.5 T/m SF gradient was used for training, and the SM at 0.6 T/m was used for testing.	73
4.6	Representative reconstructions of a simulated phantom at SNR=15-35 dB and respective error maps are shown for competing methods, along with the reference image. The SM at 0.5 T/m was used for training, and the SM at 0.6 T/m was used for testing.	74
4.7	Reconstructions of experimental cylindrical phantom with DEQ-MPI trained at SNR levels in 5-40 dB. (a) Approximate geometry of the phantom. (b) Reconstructions at (top-row) 0.5 T/m and (bottom-row) 0.6 T/m SF gradients. The SM at 0.5 T/m was used for training DEQ-MPI. Testing utilized the SM corresponding to each case. Red box denotes the case where the training and testing SNRs match.	74
4.8	Reconstructions of experimental cylindrical phantom with competing methods. (a) Approximate geometry of the phantom. (b) Reconstructed images at (top-row) 0.5 T/m and (bottom-row) 0.6 T/m SF gradients. The SM at 0.5 T/m was used for training DEQ-MPI. Testing utilized the SM corresponding to each case.	75

4.9	Reconstructions of experimental Y-shape phantom with competing methods. (a) Approximate geometry of the phantom. (b) Reconstructed images based on (top-row) the original 13×26 SM and (bottom-row) the upsampled 26×52 SM. The SM at 0.5 T/m from the second session was used for training DEQ-MPI. The SM at 0.5 T/m from the first session was used for testing.	75
4.10	Performance of competing methods as a function of inference time. ART and DIP that have markedly prolonged run times are omitted. Average pSNR across the test set is shown for each method. . . .	77

List of Tables

3.1	Imaging protocols for HR SM in analyzed datasets.	27
3.2	Average nRMSE (%) in HR SM recovery for ablation studies performed under noisy setting at $2\times-8\times$ SR.	29
3.3	Average nRMSE (%) in HR SM recovery for the simulated dataset under (noiseless/noisy) settings at $2\times-8\times$ SR.	33
3.4	Average pSNR (dB) in image reconstruction for the simulated dataset under noisy setting at $2\times-8\times$ SR.	34
3.5	Average nRMSE (%) in HR SM recovery for Open MPI at $2\times-8\times$ SR. Calibration scans were performed at different times and with unaligned FOVs for SMs in training and test sets.	37
3.6	Average nRMSE (%) in HR SM recovery / pSNR (dB) in image reconstruction for Open MPI at $2\times-8\times$ SR. SF gradient strength and MNP tracer differed for SMs in training and test sets.	37
3.7	Average nRMSE (%) in HR SM recovery / pSNR (dB) in image reconstruction for the in-house dataset at $2\times-8\times$ SR.	39
4.1	Simulated phantoms were reconstructed at varying measurement SNRs. An SM at 0.5 T/m SF gradient was used in training and test sets. pSNR (dB) / SSIM (%) are reported as mean $\pm$ std. across the test set. Boldface marks the top-performing method.	72
4.2	Simulated phantoms were reconstructed at varying measurement SNRs. The SM at 0.5 T/m was used in training, and the SM at 0.6 T/M was used in testing. pSNR (dB) / SSIM (%) are reported as mean $\pm$ std. across the test set.	72

4.3 Inference iterations (number of iterations) and reconstruction times (milliseconds) for a single 13×26 MPI image.	77
--	----



Chapter 1

Introduction

Magnetic particle imaging (MPI) is an imaging modality that allows non-invasive imaging of magnetic nanoparticles (MNP) with the potential for high-resolution & frame rate [1,2]. It has various potential applications such as angiography, cancer imaging, neuroimaging, stem cell tracking, and localized hyperthermia [3–11]. MPI uses a selection field to generate a field free region to encode the position of responsive MNPs. Then, a drive field is applied to excite the MNPs to obtain the magnetization response from the particles inside the field free region. The field free region is then moved to scan a field of view of interest for imaging.

Widely adopted image reconstruction techniques for MPI mainly involve either using a model-based x-space method [12,13] or using a system matrix (SM) based approach [1]. X-space reconstruction derives a response function based on an analytical signal model regarding MNPs and magnetic fields [13,14]. While they enable relatively fast reconstruction, the underlying signal model can neglect system imperfections such as non-ideal particle response. In contrast, SM methods explicitly measure the system function using a calibration scan, where a small MNP sample spanning an imaging voxel is traversed on a spatial grid. Because the system function is directly measured on the imaging system, SM methods offer improved immunity against system imperfections [15,16]. However, the scan is time-consuming and requires over 12 hours for an SM of $32 \times 32 \times 32$ [17].

Furthermore, it needs to be repeated for changes in either system configuration or used MNPs. Hence, accelerated SM calibration is of critical importance for a clinical MPI scanner.

There are several previous techniques for accelerated calibration. First, compressed-sensing-based approaches have been proposed by exploiting the sparsity of SMs in some transformation domains [18, 19]. While these approaches have been shown to accelerate calibration, they use the same MNP sample with a matching voxel size for recovery, which limits their SNR efficiency significantly. Another approach is to simply skip pixels uniformly and recover the overall SM through the super-resolution of SMs [17, 20]. However, these approaches also propose using small MNP samples, which also limits their SNR efficiency.

Next, image reconstruction using the SM-based techniques set up an inverse problem that seeks to reconstruct an image given the measured data and the calibrated SM. Conventional techniques usually adopt an optimization-based approach by enforcing measurement fidelity while reconstructing an image that follows MPI image statistics using some prior information [21]. Conventional techniques use hand-crafted prior information, which does not fully capture the complex nature of MPI images [15, 22–26]. Recent learning-based techniques propose learning prior information by adopting an end-to-end approach that maps data directly to an image [27–29], which is both harder to train for large SMs with large FOVs, and heavily relies on the system parameters. Moreover, they do not generalize well to other SMs. Finally, resultant images are heavily affected by the selected parameters, which may be re-optimized for each experimental data for improved performance [26, 30].

There are several approaches in computational imaging literature that have been shown to outperform previous techniques [31–37]. First, common learned approaches replace an off-the-shelf deep denoiser network as a “plug-and-play” denoiser with an associated proximal mapping of classical optimization-based approaches [38–41]. The resulting algorithm has convergence guarantees provided that the denoiser satisfies mild conditions [42]. However, these approaches usually have suboptimal performance since the denoiser is trained without being

integrated into the downstream task of image reconstruction [43, 44]. One way to integrate the denoiser into the downstream task is to first unroll the optimization algorithm for a fixed number of iterations such as 5 or 10, and train the overall network in an end-to-end fashion [44–46]. Although they outperform previous techniques in the trained task, they do not have the same convergence guarantees as their plug-and-play counterparts, such that testing with a different number of iterations compared to training results in decreased performance [47]. Furthermore, these techniques usually rely on conventional data consistency that does not capture the correlation within the data. Using learned consistency is shown to improve performance over previous approaches [48].

1.1 Contributions

In this thesis, we address the discussed issues in MPI systems. First, we address the time-consuming calibration scan. To accelerate the calibration, we propose measuring a low-resolution SM with a much fewer number of grid points compared to desired high-resolution SM. The proposed technique relies on transformers for SM super-resolution (TranSMS). For improving SNR efficiency, TranSMS uses larger MNP samples spanning an LR voxel. For improving estimation accuracy, TranSMS leverages a combination of both a vision transformer module and a convolutional module-based deep neural network. While the vision transformer module captures long-range dependencies of sinusoidal grating patterns in each SM row, the convolutional module captures the local patterns. Finally, to improve generalizability, TranSMS includes a novel data consistency module which projects the combined output of previous modules to enforce LR consistency. Demonstrations show that while both convolutional and vision transformer modules improve the estimation accuracy, the data consistency module especially improves generalizability to out-of-distribution experiments. TranSMS is demonstrated on both simulated and experimental datasets involving both field-free-point and field-free-line scanners. The experiments suggest an acceleration up to a factor of 64-fold is possible for 2D imaging. Further details and discussion is given in Chapter 3.

Next, we address image reconstruction in MPI. Since MPI is an imaging modality with real-time application potential, we propose a lightweight technique. The technique is inspired by the recent advances in physics-driven machine learning-based computational imaging. Recent approaches interchange the regularization step with a deep denoiser to better account for the distribution of image statistics while keeping the interleaved data-consistency steps intact, which in turn improves reconstruction performance. In this thesis, we propose a deep equilibrium-based image reconstruction technique for MPI (DEQ-MPI). DEQ-MPI leverages implicit layers [49, 50], which train a given network to maximize performance at fixed-point convergence. This process can be viewed as equivalent to using an unlimited number of iterations in unrolled architectures [51, 52]. The key difference is that during training, while unrolled architectures compute the gradient of each iteration and require a large amount of memory that grows with the number of iterations, DEQ-MPI uses implicit differentiation which only requires the convergent gradient to be computed. Furthermore, DEQ-MPI introduces the first learned consistency block used jointly with deep equilibrium techniques. This block exploits the MPI data distribution to better enforce data consistency under low-SNR scenarios. Finally, we propose novel initialization strategies for both regularization and learned consistency blocks for improved model training. Demonstrations on both simulated and experimental phantoms show the superior performance of DEQ-MPI compared to conventional approaches.

1.2 Outline

The rest of the thesis is organized as follows. Chapter 2 describes MPI background and physics. Chapters 3 and 4 discuss the described contributions. Specifically, Chapter 3 introduces a novel accelerated MPI SM calibration scan through deep learning enabled super-resolution. Chapter 4 describes a novel deep equilibrium-based image reconstruction technique that provides high quality, improved reliability, and faster inference. Finally, Chapter 5 includes concluding remarks, and a discussion on potential future directions.

Chapter 2

Magnetic Particle Imaging Basics

2.1 Signal Equation

In MPI, a static magnetic field, i.e. selection field, is used to generate a field free region (FFR). Then, a drive field is applied to excite the MNPs to obtain the magnetization response from the particles inside the FFR. The particles outside the FFR are saturated and do not respond to the dynamic field [1]. MNP response is then usually picked up inductively by receive coils.

The FFR is then moved inside a field of view for received signal collection. Assuming a time-varying (t) and position ($\mathbf{r}$) dependent magnetic field of $H(\mathbf{r}, t)$ and ignoring the relaxation effects, the MNP magnetization response to the magnetic field is described by the relation:

$$M(H(\mathbf{r}, t)) = mc\mathcal{L}(kH(\mathbf{r}, t)), \quad (2.1)$$

where $k = \mu_0 m / (k_B T)$ and c is the particle concentration. Here, m , k_B and T represent the MNP's magnetic moment modulus that depends on the MNP properties, Boltzmann's constant, and the temperature, respectively. μ_0 is the free space magnetic permeability. In Eq. (2.1), $\mathcal{L}(\cdot)$ is the Langevin function

defined as:

$$\mathcal{L}(kH(\mathbf{r}, t)) = \coth(kH(\mathbf{r}, t)) - \frac{1}{kH(\mathbf{r}, t)}. \quad (2.2)$$

The induced voltage on the receive coil depends on the magnetization of the MNPs $\mathbf{M}(\mathbf{r}, t)$ and the receive coil sensitivity $\rho_r(\mathbf{r})$ which only depends on the specific receive coil design:

$$s(t) = -\frac{d}{dt} \int_V \rho_r(\mathbf{r}) \cdot \mathbf{M}(\mathbf{r}, t) d\mathbf{r}, \quad (2.3)$$

where V , $\mathbf{r}$, t , and $s(t)$ denote the volume with MNPs, the position, time, and the received time domain signal, respectively. Combining Eq. (2.1) and Eq. (2.3) yields:

$$s(t) = -\frac{d}{dt} \int_V mc\rho_r(\mathbf{r}) \cdot \mathcal{L}(kH(\mathbf{r}, t)) d\mathbf{r}. \quad (2.4)$$

Finally, the received signal includes an additive background term due to interference. In this thesis, we subtract the background using multiple measurements without the presence of any MNPs within the scanner. We also use these multiple measurements for estimation of standard deviation of the noise present within measurements.

Figure 2.1 depicts the signal acquisition in MPI. Figure 2.1a shows the received signal for an MNP placed within the FFR. Since the MNP sits in the linear part of the magnetization curve, the oscillating drive field induces a time-dependent magnetization $H(\mathbf{r}, t)$. The received signal, which corresponds to the time derivative of the magnetization, has high intensities at the harmonics of a center frequency of the oscillating drive field. In contrast, Fig. 2.1b shows the received signal for an MNP placed outside the FFR. In this case, since the MNP has saturated magnetization, it does not induce a time-varying magnetization. Hence, no MPI signal is collected. Scanning the field of view (FOV) with an FFR provides spatial encoding which can then be used to reconstruct the MNP distribution.

The generated FFR can either be a field free point (FFP) or a field free line (FFL). Since the total volume of an FFP is smaller than an FFL's, FFL-based

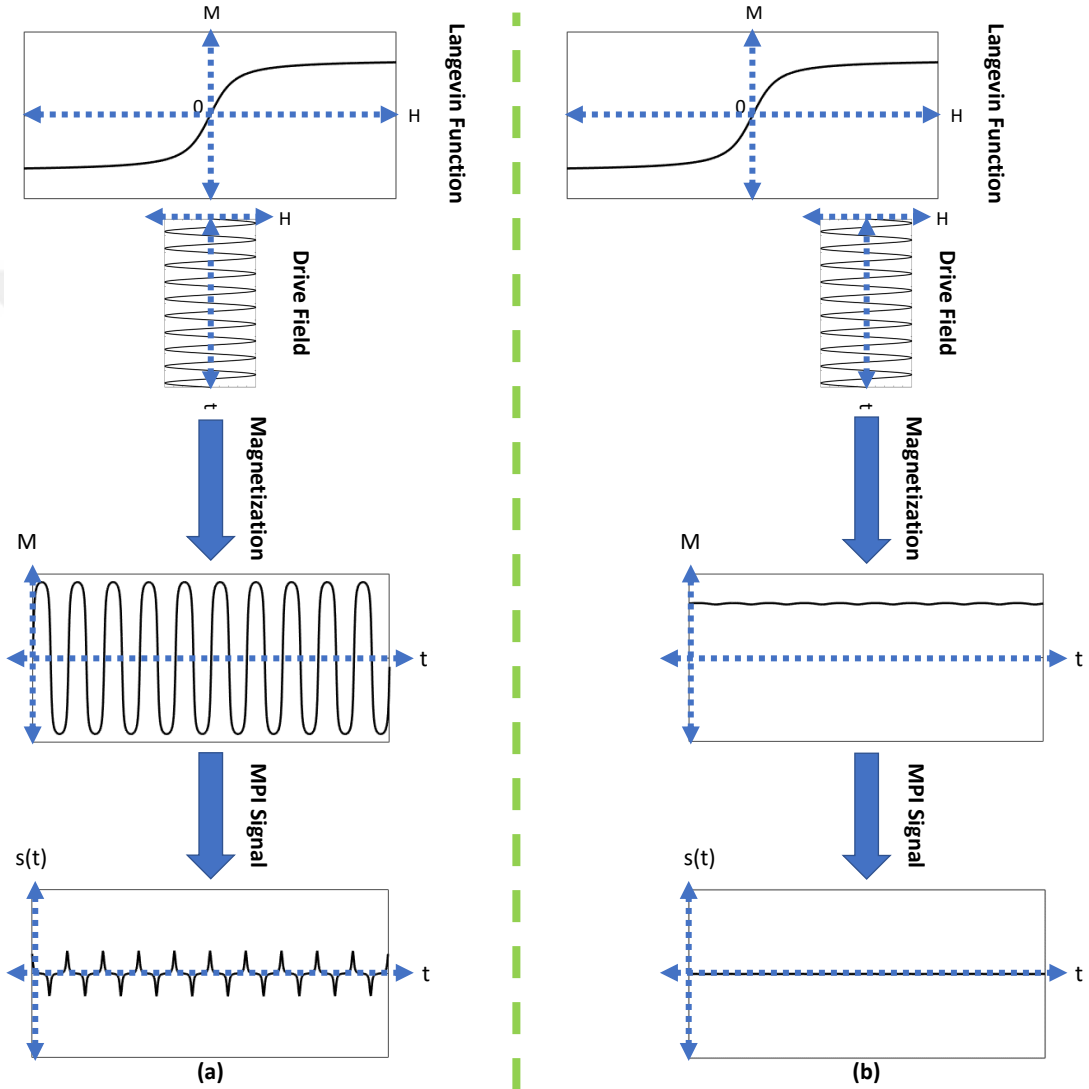


Figure 2.1: MPI encoding signal receiving process. **(a)** A depiction of the received signal for an MNP inside the FFR. The top plot depicts the magnetization of MNPs for a generated H-field. A sinusoidal drive field is applied to an MNP without a field offset. The corresponding magnetization response and the MPI signal, which is proportional to the time derivative of the magnetization, are depicted below. As shown, a high intensity $s(t)$ is received. **(b)** A depiction of the received signal for an MNP outside the FFR. A sinusoidal drive field is applied to an MNP with a field offset. The corresponding magnetization response and the MPI signal are depicted below. As shown, a negligible $s(t)$ is received.

systems provide a higher signal-to-noise ratio (SNR) [53]. They also possess a higher potential for human-sized clinical applications [30]. Within the context of this thesis, to show the wide-range applicability of the proposed techniques, we show results on both FFP and FFL-type systems [30, 54–57].

2.2 Image Reconstruction Approaches for MPI

There are two main approaches for image reconstruction in MPI [12, 58, 59]: signal model-based x-space method, or system matrix (SM) based methods. X-space-based techniques adopt a model-based approach by deriving a linear shift-invariant response function based on the analytical signal model. Although these techniques enable fast reconstruction, the analytical signal model can neglect system imperfections, which degrade overall image quality [60]. The other approach is the SM-based approach, which poses an inverse problem based on the linear forward model for image reconstruction [15]. Within the context of this thesis, we adopt SM-based approaches. The overall system can be modeled linearly as:

$$\mathbf{Ax} + \mathbf{n} = \mathbf{y}, \quad (2.5)$$

where $\mathbf{A} \in \mathbb{R}^{M \times N}$ is the system matrix, $\mathbf{x} \in \mathbb{R}^N$ is the image vector, $\mathbf{y} \in \mathbb{R}^M$ is the data vector and $\mathbf{n} \in \mathbb{R}^M$ is the noise vector. Here, M is the number of data points in frequency domain and N is the number of total grid points. Since a simulated SM cannot reflect the imperfections present within the system, a calibration scan is performed using a voxel-sized MNP phantom to construct the SM. Then, the forward model in Eq. (2.5) is used to reconstruct images. The following sections describe the common approaches and hurdles in calibration and reconstruction.

2.3 Calibration

Due to various imperfections such as inter-batch variability of MNPs, system drift or mismatch between generated and desired magnetic fields, SM is measured through a calibration scan [60,61]. The conventional technique for calibration involves filling a small MNP sample of the desired resolution and measuring its response at each position in the image grid within the FOV. This noisy measurement process can be described by:

$$\mathbf{B} = \mathbf{A} + \tilde{\mathbf{N}}, \quad (2.6)$$

where $\tilde{\mathbf{N}} \in \mathbb{R}^{M \times N}$ and $\mathbf{B} \in \mathbb{R}^{M \times N}$ represent the noise and measured SM, respectively. Since MNPs are only excited at the harmonic frequencies, it is common practice to use the Fourier transform of the gathered time-domain signal, and filter a small bandwidth around harmonic frequencies for computational efficiency [30]. Furthermore, since the SM calibration includes additive noise, frequency components having an SNR lower than some predetermined threshold are discarded to improve reconstruction quality and computational efficiency [62].

However, the conventional approach for calibration is impractically slow for clinical use [30,63]. Furthermore, the calibration needs to be repeated for changes in system parameters, trajectory, or MNP samples. Hence, there is a dire need for accelerating calibration. Moreover, a second issue arises from the noisy calibration measurements. MPI signal is linearly proportional to the total number of MNPs present inside the FOV, which is proportional to MNP concentration and the sample volume for the calibration scan. Hence, using a small MNP sample for a high-resolution (HR) calibration results in relatively low SNR. These issues have been previously addressed by many studies in the literature [18,19,63–65].

2.3.1 Compressed Sensing-based Acceleration for Calibration

First, compressed sensing (CS) based approaches that exploit the sparsity of SM rows in some transformation domains have been proposed [18, 19, 64]. CS is a signal processing technique that allows the recovery of sparse signals from their undersampled measurements. Each SM row in MPI corresponds to a complex-valued spatial sensitivity map. These rows are shown to be compressible in Fourier, cosine or Chebyshev polynomial domains [18]. Hence, in the CS-based approach, some grid points are simply omitted. The omitted points are then recovered for each row separately through the solution of an optimization problem that promotes sparsity. The measurement model is described by:

$$\mathbf{B}_{CS} = \mathbf{A}\mathbf{C} + \tilde{\mathbf{N}} \quad (2.7)$$

where $\mathbf{C}$ is the binary masking operator describing the sampled columns of SM. i^{th} row of the system expressed in Eq. (2.7) is then:

$$\mathbf{b}_{CS,i} = \mathbf{C}^T \mathbf{a}_i + \mathbf{n}_i \quad (2.8)$$

where $\mathbf{a}_i$, $\mathbf{b}_{CS,i}$, $\mathbf{n}_i$ are the reference SM row, measured SM row and noise vectors corresponding to the i^{th} row of $\mathbf{A}$, $\mathbf{B}$ and $\tilde{\mathbf{N}}$, respectively. The sampling procedure in Eq. (2.8) corresponds to the well-known image inpainting problem [34]. The SM can be recovered via the solution of the following optimization problem:

$$\mathbf{a}_i = \arg \min_a \|\mathbf{W}\mathbf{a}\|_1 \text{ s.t. } \|\mathbf{C}^T \mathbf{a} - \mathbf{b}_{CS,i}\|_2 \leq \epsilon_i \quad (2.9)$$

where $\mathbf{W}$ is the sparsifying transformation such as discrete Fourier transform or discrete cosine transform [18, 19]. ϵ_i in the constraint sets the upper bound on the noise level. While CS provides acceleration up to $27\times$ for 3D calibration [19], since the MNP sample size matches the desired SM resolution, CS-based methods inherently suffer from limited SNR efficiency. While using coded calibration scenes for improving SNR efficiency is proposed in a previous technique, it requires special manufacturing of these complex scenes [63].

2.3.2 Super Resolution-based Acceleration for Calibration

A second approach involves super-resolution-based SM calibration. One technique used a hybrid model in which a simulated HR SM and a low-resolution experimentally sampled SM is used jointly for image reconstruction [65]. The technique jointly used both SMs during image reconstruction formulated as a weighted optimization problem. The accuracy of this technique relies heavily on the fidelity of the simulated model. Furthermore, this technique does not capture the imperfections of the SM, since it uses simulated SM for high-resolution image generation. Finally, the tackled optimization problem is not only a highly non-convex problem, but also requires selection of multiple weight parameters, which will have a dramatic effect on the reconstructed image.

Baltruschat et al. proposed another technique using a similar approach to CS-based calibration [20]. However, instead of randomly undersampling the SM, they proposed uniformly subsampling the SM. Then, the SM is recovered by a deep learning-based super-resolution method. The method is trained by first gathering an HR SM, and omitting some pixels of the SM. The sampling equation is similar to Eq. (2.8):

$$\mathbf{b}_{SR,i} = \mathbf{C}_S^T \mathbf{a}_i + \mathbf{n}_i, \quad (2.10)$$

where $\mathbf{b}_{SR,i}$ is equal to uniformly subsampled elements of $\mathbf{a}_i$ with $\mathbf{C}_S$ representing the uniform subsampling matrix. The SM is estimated using a function $f_\theta(\mathbf{b}_{SR,i})$ by first training it on a set:

$$\arg \min_{\theta} \|f_\theta(\mathbf{b}_{SR,i}) - \mathbf{a}_i\|_1, \quad (2.11)$$

then test its performance on a differing SM. However, this method still suffers from the same SNR-efficiency problem that exists in CS-based acceleration.

2.4 Image Reconstruction

Image reconstruction in MPI is posed as an inverse problem, in which an image estimate is obtained through the solution of an optimization problem. However, since the SM is not full-rank, the optimization problem is ill-posed. Hence, prior information has a great effect on the image estimate. Furthermore, the measurement noise in Eq. (2.5) is colored [66]. Hence, a pre-whitening step by estimating the standard deviation in each row is usually recommended [23, 30, 55]. After whitening, previous approaches solve an optimization problem of the form:

$$\arg \min_{\mathbf{x}} R(\mathbf{x}) + \lambda \|\mathbf{Ax} - \mathbf{y}\|_2^2, \quad (2.12)$$

where $R(\cdot)$ is the prior information on $\mathbf{x}$ and λ is the regularization between the prior information and the data consistency term [15, 39, 59, 67, 68]. Moreover, previous techniques usually rely on re-optimization of hyper-parameters [26, 30]. Therefore, a robust, stable and fast reconstruction algorithm for MPI is of critical importance. State of the art techniques for MPI image reconstruction are summarized below.

2.4.1 Algebraic Reconstruction Technique (ART)

Algebraic reconstruction technique (ART) iteratively solves the regularized least squares problem by setting $R(\mathbf{x}) = \|\mathbf{x}\|_2^2$ in Eq. (2.12). Here, λ is a regularization constant that needs to be chosen based on the average image energy and the SNR of the measurements. ART reconstructs images through the iterations:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{a}_i^T \frac{\mathbf{y}[i] - \mathbf{a}_i \mathbf{x}_k}{\|\mathbf{a}_i\|_2^2}, \text{ for } 1 \leq i \leq M \quad (2.13)$$

where k is the iteration index, $\mathbf{y}[i]$ is the i -th element of the measurement vector and $\mathbf{a}_i$ is the i -th row of the SM. However, ℓ_2 -norm-based reconstruction usually tends to over-smooth the reconstructed images. Moreover, due to its sequential nature, ART requires high computational time for reconstruction.

2.4.2 Alternating Direction Method of Multipliers (ADMM)

Alternating direction method of multipliers (ADMM) is an iterative technique that solves an optimization problem by splitting it into multiple easier subproblems [34, 39]. The technique is shown to converge under mild conditions. Within the context of MPI, ADMM is employed both to improve image quality and to accelerate reconstruction [15]. While there are many adaptations in a generic context, here we only describe a specific implementation that has been used in MPI [15, 30].

Ilbey et al. proposed a previous approach by setting $R(\mathbf{x}) = \alpha_1 \|\mathbf{x}\|_1 + \alpha_{TV} TV(\mathbf{x})$ in constrained formulation, where $TV(\cdot)$ is the total variation, and $\alpha_{1,TV}$ are the regularization weights [15]:

$$\arg \min_{\mathbf{x}} \alpha_1 \|\mathbf{x}\|_1 + \alpha_{TV} TV(\mathbf{x}) \text{ subject to } \|\mathbf{Ax} - \mathbf{y}\|_2 \leq \epsilon, \quad (2.14)$$

where ϵ can be chosen according to the bound on the expectation of ℓ_2 -norm of the noise vector, $E[\|\mathbf{n}\|_2]$. This approach has been proved to be useful for many inverse problems involving other modalities [34, 69–71]. ADMM solves problems of type:

$$\arg \min_{\mathbf{x}, \mathbf{z}} f(\mathbf{x}) + g(\mathbf{z}), \text{ subject to } \mathbf{Gx} + \mathbf{Qz} = \mathbf{c}, \quad (2.15)$$

through the iterations:

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{Gx} + \mathbf{Qz}_k - \mathbf{c} - \mathbf{d}_k\|_2^2, \quad (2.16)$$

$$\mathbf{z}_{k+1} = \arg \min_{\mathbf{z}} g(\mathbf{z}) + \frac{\mu}{2} \|\mathbf{Gx}_{k+1} + \mathbf{Qz} - \mathbf{c} - \mathbf{d}_k\|_2^2, \quad (2.17)$$

$$\mathbf{d}_{k+1} = \mathbf{d}_k - \mathbf{Gx}_{k+1} - \mathbf{Qz}_{k+1} + \mathbf{c}, \quad (2.18)$$

where μ is proportional to the inverse of the step size, and k is the iteration index. To solve the given problem in Eq. (2.14), the following equalities are defined:

$$\mathbf{G} = \begin{bmatrix} \mathbf{A} \\ \mathbf{I} \\ \mathbf{I} \end{bmatrix}, \mathbf{z} = \begin{bmatrix} \mathbf{z}^{(0)} \\ \mathbf{z}^{(1)} \\ \mathbf{z}^{(2)} \end{bmatrix} \mathbf{Q} = -\mathbf{I}, \mathbf{c} = 0, \quad (2.19)$$

$$f(\mathbf{x}) = 0, g(\mathbf{z}) = \alpha_1 \|\mathbf{z}^{(1)}\|_1 + \alpha_{TV} TV(\mathbf{z}^{(2)}) + I(\mathbf{z}^{(0)} \|\mathbf{z}^{(0)} - \mathbf{y}\|_2 \leq \epsilon), \quad (2.20)$$

where $I(\mathbf{z}^{(0)}|\|\mathbf{z}^{(0)} - \mathbf{y}\|_2 \leq \epsilon)$ is the indicator function equal to 0 if the condition is satisfied, and ∞ otherwise. ADMM then solves:

$$\begin{aligned} & \arg \min_{\mathbf{x}, \mathbf{z}} \alpha_1 \|\mathbf{z}^{(1)}\|_1 + \alpha_{TV} TV(\mathbf{z}^{(2)}), \\ & \text{subject to } \mathbf{x} = \mathbf{z}^{(1)} = \mathbf{z}^{(2)}, \mathbf{A}\mathbf{x} = \mathbf{z}^{(0)}, \|\mathbf{z}^{(0)} - \mathbf{y}\|_2 \leq \epsilon, \end{aligned} \quad (2.21)$$

which is equivalent to Eq. (2.14). The resulting iterations are then:

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} (2\mathbf{I} + \mathbf{A}^T \mathbf{A})^{-1} \left(\left(\sum_{i=1}^2 \mathbf{d}_k^{(i)} + \mathbf{z}_k^{(i)} \right) + \mathbf{A}^T (\mathbf{d}_k^{(0)} + \mathbf{z}_k^{(0)}) \right), \quad (2.22)$$

$$\mathbf{z}_{k+1}^{(0)} = \arg \min_{\mathbf{z}^{(0)}} \|\mathbf{z}^{(0)} - (\mathbf{A}\mathbf{x}_{k+1} - \mathbf{d}_k^{(0)})\|_2^2 \text{ subject to } \|\mathbf{z}^{(0)} - \mathbf{y}\|_2 \leq \epsilon, \quad (2.23)$$

$$\mathbf{z}_{k+1}^{(1)} = \arg \min_{\mathbf{z}^{(1)}} \alpha_1 \|\mathbf{z}^{(1)}\|_1 + \frac{\mu}{2} \|\mathbf{z}^{(1)} - (\mathbf{x}_{k+1} - \mathbf{d}_k^{(1)})\|_2^2, \quad (2.24)$$

$$\mathbf{z}_{k+1}^{(2)} = \arg \min_{\mathbf{z}^{(2)}} \alpha_{TV} TV(\mathbf{z}^{(2)}) + \frac{\mu}{2} \|\mathbf{z}^{(2)} - (\mathbf{x}_{k+1} - \mathbf{d}_k^{(2)})\|_2^2, \quad (2.25)$$

$$\mathbf{d}_{k+1}^{(0)} = \mathbf{d}_k^{(0)} + \mathbf{z}_{k+1}^{(0)} - \mathbf{A}\mathbf{x}, \quad (2.26)$$

$$\mathbf{d}_{k+1}^{(1)} = \mathbf{d}_k^{(1)} + \mathbf{z}_{k+1}^{(1)} - \mathbf{x}, \quad (2.27)$$

$$\mathbf{d}_{k+1}^{(2)} = \mathbf{d}_k^{(2)} + \mathbf{z}_{k+1}^{(2)} - \mathbf{x}. \quad (2.28)$$

While this approach yields superior performance compared to ART, the parameters are still required to be re-optimized for each experimental phantom [30].

2.4.3 Deep Image Prior (DIP)

The final technique discussed in this section uses learning based regularization. Deep image prior uses regularization by architecture [72]. This approach removes the explicit regularization function $R(\cdot)$, and constrains the reconstructed image to the images that can be generated by deep neural networks. Hence, the untrained network architecture serves as an implicit regularizer for reconstruction. This approach does not require a training dataset, and is shown to perform well in absence of training data. Within the context of MPI, assuming a forward mapping of $f_\theta(\mathbf{g})$ where $\mathbf{g}$ can be generated from i.i.d. Gaussian distribution or simply selected as least squares solution, a neural network is trained for each

phantom reconstruction with the loss function [73, 74]:

$$\arg \min_{\theta} \|\mathbf{A}f_{\theta}(\mathbf{g}) - \mathbf{y}\|_1. \quad (2.29)$$

Here, setting $f_{\theta}(\mathbf{g}) = \mathbf{x}$ provides regularization [72]. However, image reconstruction requires thousands of iterations for training the network. Furthermore, selection of specific model architecture is challenging [75].



Chapter 3

Super-Resolution based Accelerated Calibration for Magnetic Particle Imaging

The contents of this chapter reflect the work reported in the following publications:

- A. Güngör, B. Askin, D. A. Soydan, E. U. Saritas, C. B. Top and T. Çukur, “TranSMS: Transformers for Super-Resolution Calibration in Magnetic Particle Imaging,” in arXiv preprint, 2022.
- A. Güngör, B. Askin, D. A. Soydan, C. Barış Top and T. Cukur, “Deep Learned Super Resolution of System Matrices for Magnetic Particle Imaging,” 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 3749-3752, doi: 10.1109/EMBC46164.2021.9630601.
- A. Güngör, B. Askin, D. A. Soydan, E. U. Saritas, C. B. Top and T. Çukur, “TranSMS: Transformers for Super-Resolution Calibration in Magnetic Particle Imaging,” in IEEE Transactions on Medical Imaging, vol. 41, no. 12, pp. 3562-3574, Dec. 2022, doi: 10.1109/TMI.2022.3189693.

3.1 Introduction

Two fundamental approaches for MPI reconstruction differ in their treatment of the system function: X-space and system matrix (SM) methods. In X-space methods, a point spread function is used to characterize the MPI signal, typically based on an analytical model or alternatively based on measurements [13]. In contrast, while analytical approaches have been proposed [65], SM methods commonly measure the system function using a calibration scan where a small magnetic nanoparticle (MNP) sample spanning an imaging voxel is traversed on a spatial grid. When the system function is directly measured on the imaging system, SM methods offer improved immunity against system imperfections [15, 16]. Yet, a time-consuming calibration (e.g., ~ 12 hours for a typical $32 \times 32 \times 32$ grid) must be repeated whenever the MNP type or scanning trajectory is changed [17]. Pioneering studies in this domain have proposed accelerating calibration based on compressed sensing (CS) [18, 19, 63, 64, 76] and super-resolution (SR) [20, 65, 77]. However, calibration using a small MNP sample can manifest low signal-to-noise ratio (SNR) efficiency, limiting the quality of recovered SMs at high acceleration rates [63, 78].

Here we propose a novel deep learning (DL) method for accelerated MPI calibration, named Transformers for SM Super-resolution (TranSMS). For improved SNR efficiency, TranSMS performs low-resolution (LR) SM measurements using a larger MNP sample spanning across an LR voxel as opposed to a high-resolution (HR) voxel. HR SM is then recovered by a deep network that super-resolves LR SM. SM rows follow spatial patterns that visually resemble sinusoidal grating functions, windowed by an envelope whose shape is linearly proportional to the receive coil sensitivity. Thus, pixel intensities in each SM intrinsically show long-range correlations. Prior SR methods are predominantly based on convolutional neural networks (CNNs) that improve performance over traditional methods [20, 79]. However, due to their local kernels, CNNs cannot capture the long-range contextual features in SM rows. To address this limitation, we introduce a novel architecture that aggregates the localization power of CNNs with the contextual expressiveness of vision transformers. TranSMS leverages a novel

data-consistency module to ensure fidelity to LR SM measurements based on the signal model between LR and HR SM. Comprehensive demonstrations are performed on simulated and experimental data from field-free point (FFP) and field-free line (FFL) scanners, for up to 64-fold acceleration during two-dimensional (2D) imaging. Our results indicate that TranSMS offers substantial performance improvements in both SM recovery and image reconstruction over the state-of-the-art CS- and SR-based calibration methods. Our main contributions are:

- We introduce a novel deep learning approach to super-resolve MPI SMs for accelerated calibration.
- To our knowledge, we introduce the first vision transformer model for MPI.
- We introduce a novel data-consistency module to incorporate the physical signal model relating HR SM to LR SM measurements.

3.2 Related Work

Several successful approaches have previously been introduced for accelerating MPI calibration. A hybrid approach combines model- and measurement-based SM estimates [65]. An HR SM is analytically simulated while an LR SM is measured. This method was used to achieve 4-fold accelerated 2D calibration by 2-fold increase of voxel dimensions in LR SM. Reconstruction is then formulated as a weighted optimization with terms reflecting the simulated and measured SMs. This model-assisted SR approach derives HR information from an analytical model. Thus, accuracy might be compromised by system imperfections that are not captured by the model.

The CS framework instead exploits sparsity to recover signals from undersampled measurements [18, 19, 64]. Each SM row is a complex-valued spatial sensitivity map with compressible representations in Fourier, cosine or Chebyshev polynomial domains [18]. Thus, calibrations can be accelerated by randomly undersampling the grid locations during HR SM measurements (Fig. 3.1b). The full

SM can then be recovered by simultaneously enforcing consistency to available measurements, transform-domain sparsity, and additional priors regarding the symmetry of sensitivity maps. For 3D FFP imaging, accelerations up to 27-fold were demonstrated under relatively limited noise [19]. Other optimization-based methods have exploited non-equispaced sampling and low-rank approximations through higher-order singular value decomposition [76, 80]. Yet, HR SM measurements using small MNP samples spanning a single imaging voxel limit SNR efficiency. Although calibration with multiple MNP samples per measurement can improve efficiency [63], there are practical challenges in positioning of a large calibration phantom.

Recently, DL techniques have also been considered for MPI calibration [20]. A previous study has proposed using strided LR SM measurements, as shown in Fig. 3.1c, followed by upsampling to recover an HR SM [20]. Upsampling is performed via a CNN trained using paired sets of LR and HR SMs. This SR approach was employed to recover HR SM from 64-fold accelerated 3D calibration scans (analogously 16-fold acceleration for 2D calibration). That said, strided LR SM measurements performed using an MNP sample spanning a single HR voxel can potentially limit SNR efficiency.

Super-resolving HR SM in MPI is closely related to the single-image super-resolution (SISR) problem in computer vision [20, 65, 79]. DL methods have recently become the gold standard in this domain. Successful SISR models have been introduced for natural images based on convolutional architectures, such as Super-Resolution Convolutional Neural Network (SRCNN) [81], Very Deep Super-Resolution (VDSR) [82], and Residual-Dense Network (RDN) [83]. That said, characteristics of SM rows notably differ from natural images. First, SM sizes are relatively compact compared to natural images. Second, SMs contain a single broadly distributed pattern whereas multiple nonoverlapping objects can exist in natural images. Third, SMs are corrupted by higher levels of noise compared to natural images. Due to these differences, deep architectures with high-dimensional feature maps can be suboptimal for SM recovery. Furthermore, CNN

models with compact filters have limited sensitivity to long-range contextual features [84]. This has motivated the adoption of hybrid CNN-transformer models for improved contextual sensitivity in SISR [85, 86]. However, prior models used serially-connected CNNs and vanilla vision transformers [87]. In contrast, TranSMS uses parallel-connected convolutional and transformer modules, and novel transformer blocks with convolutional patch embedding. Few recent studies have introduced efficient transformers for medical imaging tasks [88–91]. Yet, this is the first study to introduce a vision transformer for MPI, and the first study to introduce a model-based transformer for SR.

3.3 Theory

3.3.1 Magnetic Particle Imaging

3.3.1.1 Signal Model

In MPI, excitation is performed via an oscillating drive field (DF) at a fundamental frequency, while spatial encoding is enabled via a static SF that creates a field-free region (FFR) within the field-of-view (FOV). Voltage waveforms captured by receive coils reflect MNP responses within the FFR. SM-based reconstruction is performed by transforming the received time domain signal to frequency domain. The signal model can be expressed as [1]:

$$\mathbf{A}\mathbf{x} + \mathbf{n} = \mathbf{y} \quad (3.1)$$

where $\mathbf{A} \in C^{M \times N}$ is the SM; $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{n} \in C^M$ and $\mathbf{y} \in C^M$ are the image vector, noise vector and frequency domain signal, respectively. Here, M is the number of frequency components and N is the number of voxels in the grid. MPI data contain thermal noise that varies as a function of frequency [66], and the overall signal intensity decreases with increasing harmonic frequency [1]. Thus, the additive noise in Eq. (3.1) can be taken as colored Gaussian [30, 66].

3.3.1.2 SM Calibration

The SM is affected by imperfections such as system drift or inter-batch variability in MNP response [61], and is spatially variant [60]. To account for such non-idealities, SM calibration is performed by traversing a small MNP sample across the imaging grid [17], as shown in Fig. 3.1a:

$$\mathbf{B} = \mathbf{A} + \tilde{\mathbf{N}}. \quad (3.2)$$

Here, $\tilde{\mathbf{N}}$ and $\mathbf{B}$ denote additive noise and the measured SM, respectively. Measurements are typically taken with an MNP sample size that matches the intended imaging voxel size. This limits SNR efficiency in HR SM measurements.

3.3.1.3 Image Reconstruction

Once an SM estimate $\hat{\mathbf{A}}$ is obtained from measurements $\mathbf{B}$, image reconstruction can be cast as an inverse problem to solve Eq. (3.1). Two common optimization approaches are the Kaczmarz [59] and alternating direction method of multipliers (ADMM) methods [15, 63]. Here, we consider the ADMM-based formulation [39, 69]:

$$\arg \min_{\mathbf{x}} \alpha_1 \|\mathbf{x}\|_1 + \alpha_{TV} \text{TV}(\mathbf{x}) \text{ s.t. } \|\hat{\mathbf{A}}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon, \quad (3.3)$$

where $\alpha_{1,TV}$ are regularization weights. Eq. (3.3) promotes sparsity in image and finite-differences (TV) domains, while constraining fidelity of reconstructed image and acquired signals with an ϵ bound based on noise level. To account for varying noise levels across SM rows in Eq. (3.3), repeated background measurements can be used for noise whitening, and filtering out frequency components with low SNR [30, 62, 92, 93]. For whitening, $\hat{\mathbf{A}}$ and $\mathbf{y}$ can be multiplied by a diagonal matrix with elements $\mathbf{E}_{i,i} = 1/\sigma_i$, where σ_i is the standard deviation (std.) of noise for the i^{th} SM row (i.e., std. of noise in i^{th} component in frequency domain) [30, 62]. Here, we pre-whiten all SM rows to unit std.

3.3.2 TranSMS

To improve SNR efficiency, TranSMS performs LR SM measurements using larger MNP samples as opposed to smaller samples used in strided LR SM measurements (Fig. 3.1). To improve HR SM recovery, the SR problem is solved via a novel transformer architecture.

3.3.2.1 Super-Resolution MPI Calibration

Inspired by recent reports [65, 78], here we propose to measure LR SM with MNP samples matching the LR SM voxel size. A large MNP sample spans multiple voxels on the HR SM grid, so the proposed measurements correspond to box-car downsampling of the HR SM with a matrix $\mathbf{D}$:

$$\mathbf{B} = \mathbf{A}\mathbf{D}^T + \tilde{\mathbf{N}} \quad (3.4)$$

For each SM row, the sampling relationship is $\mathbf{b}_i = \mathbf{D}\mathbf{a}_i + \mathbf{n}_i$, so recovery of HR SM row $\mathbf{a}_i = (\mathbf{A}_{HR})_i$ given LR SM row $\mathbf{b}_i$ is an SR task. Since noise does not depend on the MNP sample size, the proposed measurements yield improved SNR efficiency. For instance, measurements with a 16-fold larger MNP sample would result in a 16-fold SNR increase.

3.3.2.2 Network Model

We solve Eq. (3.4) to recover HR SM from LR SM measurements. MPI SMs resemble sinusoidal gratings with strong inter-voxel correlations over broad distances. The MPI signal also shows spatially-variant characteristics due to coil inhomogeneities. Taken together, these factors render CNN-based SR with local, translation-invariant filters suboptimal for SM recovery. Instead, we propose a hybrid architecture to super-resolve HR SM at an SR factor of S , corresponding to S^2 -fold acceleration for 2D calibration, leveraging transformer, convolutional, and data-consistency modules. The transformer module captures contextual image features, the convolutional module captures localized features, whereas the

data-consistency module enforces fidelity to LR SM measurements. The network input is a reformatted row of LR SM $\mathbf{b}_i \in \mathbb{R}^{2 \times W \times H}$, with two channels in the first dimension representing real/imaginary parts, W and H denoting the width and height of the LR grid (i.e., $WH = N/S^2$). An initial convolutional layer $Z_{init}(\cdot)$ expands the channels to produce $\mathbf{U}_{init} = Z_{init}(\mathbf{b}_i) \in \mathbb{R}^{C_1 \times W \times H}$ (with C_1 channels), which is then fed to the transformer and convolutional modules.

Transformer module: The transformer module contains a cascade of n_{TB} submodules that capture contextual features in SM rows (Fig. 3.2). Each submodule has a token-embedding block ($B_{TE,j}(\cdot)$) followed by a transformer block ($B_{T,j}(\cdot)$), where j denotes submodule index. Token-embedding is achieved via a sequence of convolutional, reshaping and layer normalization (LN) layers [94]. Assuming the input is $\mathbf{U}_{T,j-1}$ (where $\mathbf{U}_{T,0} = \mathbf{U}_{init}$), the token-embedded feature map is:

$$\mathbf{u}_{\text{encoded},j} = \text{LN}(\text{Reshape}_{1D}(\text{Conv}_{2D}(\mathbf{U}_{T,j-1}))), \quad (3.5)$$

where Conv_{2D} has 3×3 kernels with stride s_j . Unlike non-overlapping patch embeddings [87], overlapping token embeddings in Eq. (3.5) can improve the capture of local and global context [94]. $\mathbf{u}_{\text{encoded},j} \in \mathbb{R}^{WH/s_j^2 \times C_T}$ (C_T denoting number of channels) is processed via a *convolutional transformer* block comprising convolutional projection, multi-head self-attention (MHSA) and multi-layer perceptron (MLP) sub-blocks. The projection extracts query, key and value for each token via depth-wise separable convolution (DWSC) [95]:

$$\mathbf{Q}_j, \mathbf{K}_j, \mathbf{V}_j = \text{Flatten}(\text{DWSC}(\text{Reshape}_{2D}(\mathbf{u}_{\text{encoded},j}))), \quad (3.6)$$

where the reshape operator formats the dimensions to $\mathbb{R}^{C_T \times W/s_j \times H/s_j}$, $\mathbf{Q}_j, \mathbf{K}_j, \mathbf{V}_j \in \mathbb{R}^{n_A \times WH/s_j^2 \times C_T}$ (with n_A denoting the number of attention heads) are query, key and value, respectively. DWSC is implemented as:

$$\text{DWSC}(\mathbf{v}) = \text{Conv}_1(\text{BN}(\text{Conv}_{2D}(\mathbf{v}))), \quad (3.7)$$

where $\text{Conv}_1(\cdot)$ is 1×1 convolution, BN is batch normalization. The projected maps are then residually fed to MHSA:

$$\mathbf{u}_{\text{MHSA},j} = \text{MHSA}(\mathbf{Q}_j, \mathbf{K}_j, \mathbf{V}_j) + \mathbf{u}_{\text{encoded},j}, \quad (3.8)$$

where $\mathbf{u}_{\text{MHSA},j} \in \mathbb{R}^{WH/s_j^2 \times C_T}$ is the output and MHSA is:

$$\text{MHSA} = \text{Lin} \left(\text{Reshape}_{1\text{D}} \left(\text{SftMx} \left(\mathbf{Q}_j \mathbf{K}_j^T \right) \mathbf{V}_j \right) \right), \quad (3.9)$$

In Eq. (3.9), the linear layer (Lin) compresses the number of channels from $n_A C_T$ to C_T . Finally, the feature maps are projected with residual connection through an MLP:

$$\mathbf{u}_{T,j+1} = \text{Reshape}_{2\text{D}}(\text{MLP}(\text{LN}(\mathbf{u}_{\text{MHSA},j})) + \mathbf{u}_{\text{MHSA},j}), \quad (3.10)$$

where $\mathbf{u}_{T,j+1} \in \mathbb{R}^{C_T \times W/s_j \times H/s_j}$ is the output map, and MLP has a single hidden layer of size $2C_T$. A stride of $s_j > 1$ broadens the convolutional kernels to permit analysis at a larger scale. Here, we used $n_{TB} = 3$ submodules with $s_j = [1, 1, 2]$. Since the last submodule halved the image size, an upsampler block (see Fusion Module for details) was used to restore the original dimensions of $W \times H$.

Convolutional module: The convolutional module is composed of convolutional layers ($Z_{C,0}(\cdot)$ & $Z_{C,1}(\cdot)$) followed by a cascade of n_{RDB} Residual-Dense Blocks (RDBs). Convolution layers project $\mathbf{U}_{init}$ to feature map $\mathbf{u}_0 \in \mathbb{R}^{C_C \times W \times H}$.

$$\mathbf{u}_{-1} = Z_{C,0}(\mathbf{U}_{init}), \quad \mathbf{u}_0 = Z_{C,1}(\mathbf{u}_{-1}) \quad (3.11)$$

Then, the RDBs sequentially process the feature maps as:

$$\mathbf{u}_d = B_d(\mathbf{u}_{d-1}) \quad (3.12)$$

where $B_d(\cdot)$ is the projection through the d^{th} RDB with n_{CL} convolutional layers each performing a projection $Z_{d,k}(\cdot)$. Each layer concatenates all feature maps from previous layers:

$$\mathbf{u}_{d,k} = Z_{d,k}([\mathbf{u}_{d-1}; \mathbf{u}_{d,1}; \mathbf{u}_{d,2}; \cdots ; \mathbf{u}_{d,k-1}]) \quad (3.13)$$

where $\mathbf{u}_{d,k}$ has n_{GR} channels denoting the growth-rate of the block across a layer. The last layer implements $Z_{d,out}(\cdot)$ to reduce the number of channels to C_C via a 1×1 convolution:

$$\mathbf{u}_d = Z_{d,out}([\mathbf{u}_{d-1}; \mathbf{u}_{d,1}; \cdots ; \mathbf{u}_{d,r}]) + \mathbf{u}_{d-1} \quad (3.14)$$

Finally, the output of all RDBs are concatenated through a 1×1 convolution operator $Z_{C,ct}$ as:

$$\mathbf{u}_{ct} = Z_{C,ct}([\mathbf{u}_1; \mathbf{u}_2; \cdots; \mathbf{u}_{n_{RDB}}]) \quad (3.15)$$

The output is then combined with a global residual from the initial layer using a final convolutional layer $Z_{C,last}(\cdot)$ as:

$$\mathbf{U}_C = Z_{C,last}(\mathbf{u}_{ct}) + \mathbf{u}_{-1} \quad (3.16)$$

where $\mathbf{U}_C$ denotes the output of the convolutional module.

Fusion module: The fusion module first aggregates localized and contextual representations, $\mathbf{U}_C$ and $\mathbf{U}_T$, respectively from the convolutional and transformer modules via a convolutional layer ($Z_{cat}(\cdot)$) to generate $\mathbf{U}_{cat} \in \mathbb{R}^{C_{cat} \times H \times W}$:

$$\mathbf{U}_{cat} = Z_{cat}([\mathbf{U}_C; \mathbf{U}_T]) \quad (3.17)$$

The fused map is then processed via a cascade of 2-fold upsampling blocks. Assuming that S is a power of 2, a feature map $\mathbf{U}_{ups} \in \mathbb{R}^{C_{cat} \times SH \times SW}$ is generated:

$$\mathbf{U}_{ups} = \text{Upsampler}_{2 \times}(\cdots(\text{Upsampler}_{2 \times}(\mathbf{U}_{cat}))) \quad (3.18)$$

The upsampling blocks consist of convolutional layers that expand the number of channels 4-fold, and a pixel shuffler to map channels onto higher resolution tensors [96]. A final convolutional layer ($Z_{fin}(\cdot)$) produces $\tilde{\mathbf{a}}_i \in \mathbb{R}^{2 \times SH \times SW}$, the i^{th} row of SM super-resolved by a factor of S :

$$\tilde{\mathbf{a}}_i = Z_{fin}(\mathbf{U}_{ups}) \quad (3.19)$$

Data-Consistency Module: DL methods for SR tasks typically perform a direct mapping between LR and HR images without observing a signal model, posing a risk for inconsistency between measured LR and predicted HR images. Although the training loss can implicitly satisfy consistency on the training set, it offers no guarantees on the test set. To obtain consistent LR-HR SMs, we introduce a novel data-consistency (DC) module in TranSMS to enforce fidelity between recovered and measured data as inspired by recent MRI methods [45, 48].

Based on the signal model in Eq. (3.4), DC can be enforced by comparing box-car downsampled HR SM with measured LR SM. Receiving as input $\tilde{\mathbf{a}}_i$:

$$\hat{\mathbf{a}}_i = \arg \min_{\mathbf{a}} \|\mathbf{a} - \tilde{\mathbf{a}}_i\|_2^2 \text{ s.t. } \|\mathbf{D}\mathbf{a} - \mathbf{b}_i\|_2 \leq \sqrt{m}\sigma_i \quad (3.20)$$

where $\hat{\mathbf{a}}_i$ is the estimated HR SM row, $m = WH$. This projection enforces HR SM to be consistent with measured LR SM within an error bound determined by noise levels. Here, we implement the DC projection efficiently by a closed-form analytical solution (see Appendix):

$$\tilde{\mathbf{b}}_i = \mathbf{b}_i - \mathbf{D}\tilde{\mathbf{a}}_i \quad (3.21)$$

$$\hat{\mathbf{a}}_i = \begin{cases} \tilde{\mathbf{a}}_i, & \text{if } \|\tilde{\mathbf{b}}_i\|_2 \leq \sqrt{m}\sigma_i \\ \tilde{\mathbf{a}}_i + \left(1 - \frac{\sqrt{m}\sigma_i}{\|\tilde{\mathbf{b}}_i\|_2}\right) \mathbf{D}^T \tilde{\mathbf{b}}_i, & \text{otherwise.} \end{cases} \quad (3.22)$$

Network Loss: The overall network can be cast as a feed-forward function $f_\theta(\mathbf{b}_i) = \hat{\mathbf{a}}_i$, where θ denotes the set of learnable parameters. The network is trained using ℓ_1 -norm loss between the recovered and reference HR SM rows:

$$\arg \min_{\theta} \sum_i \|f_\theta(\mathbf{b}_i) - \mathbf{a}_i\|_1 \quad (3.23)$$

3.4 Methods

3.4.1 Datasets

As experimental datasets are inherently noisy and of practically limited size, we first conducted controlled analyses on a simulated dataset. These analyses included ablation studies to demonstrate the components of the network architecture, ablation studies to assess the influence of the size and diversity of the training set, as well as comparisons to competing methods in noisy and noise-free settings. Next, we conducted analyses on two experimental datasets to demonstrate TranSMS on both FFP and FFL scanners. These analyses aimed to examine reliability against variability in selection field (SF) strength, FOV position

Table 3.1: Imaging protocols for HR SM in analyzed datasets.

	Simulated Dataset	Open MPI Dataset (#7) / (#8, #9, #10)	In-house MPI Dataset
FOV Size (mm)	$32 \times 32 \times 1$	$37 \times 37 \times 18.5$ / $66 \times 66 \times 27$	$32 \times 64 \times 2$
Grid Size (px)	$32 \times 32 \times 1$	$37 \times 37 \times 37$ / $33 \times 33 \times 27$	$16 \times 32 \times 1$
Sequence	FFL - 3°	FFP - Lissajous	FFL - 3°
MNP	[14.1-34.47] nm	Synomag / Perimag	Perimag
Sample Sz. (mm)	$1 \times 1 \times 1$	$2 \times 2 \times 1$	$2 \times 2 \times 2$
SF Grad. (T/m)	[0.4-1.03] $\perp$ to FFL	$-1 \times -1 \times 2$ / $-0.5 \times -0.5 \times 1$	[0.3-0.6] $\perp$ to FFL
DF Freq.	23.2 kHz	2.5/102 MHz 2.5/96 MHz 2.5/99 MHz	23.2 kHz
DF Amp. (mT)	$0 \times 0 \times [6.4-16]$	$12 \times 12 \times 12$	$0 \times 0 \times 9$

and MNP tracer type, as well as variability over time. In this section, we describe the details regarding the calibration measurements (Table 3.1). The voxel size of each LR SM is $S \times S$ times larger compared to the respective HR SM.

3.4.1.1 Simulated Dataset

We used an in-house MPI simulator modeling the open-sided FFL scanner topology in [30]. A 32×32 mm² FOV was scanned with a sample of size 1×1 mm² by rotating the FFL in the x-y plane. MNP saturation magnetization and temperature were $0.55 T/\mu_0$ and 300 °K. A sinusoidal DF with varying amplitude (depending on the SF gradient and FOV size) was applied in the z-direction. The received signal was sampled at 5 MS/s for a receive coil with homogeneous sensitivity in the z-direction. Following Fourier transformation of time-domain signals, the components corresponding to harmonics 2-to-9 were captured. Finally, imaging data were generated for a numerical phantom.

Simulated SMs were split to use 66 for training, 30 for validation and 34 for

testing. The validation set contained SF gradients $[0.70:0.265:1.03]$ T/m, and MNP diameters $[15.17:2.145:34.47]$ nm. For the training and test sets, 100 SMs were generated with SF gradients of $[0.40:0.075:1.00]$ T/m, MNP diameters of $[14.10:2.145:33.40]$ nm. For the test set, we exclusively reserved 19 of these SMs with SF gradient 1 T/m or MNP diameter 33.40 nm, and 15 additional randomly selected SMs were added for improving diversity.

3.4.1.2 Open MPI Dataset

Open MPI is a public dataset with calibration and imaging measurements on a preclinical FFP scanner (Bruker, Ettlingen) [54]. We analyzed calibration data collected at $37 \times 37 \times 18.5$ mm³ and $66 \times 66 \times 27$ mm³ FOVs, a sample of size $2 \times 2 \times 1$ mm³ and a 3D Lissajous sequence. Signals from 3 coils were sampled at 2.5 MS/s. SM rows with high-SNR responses were selected (SNR>5) [62, 65, 73], and whitened based on the noise power computed using the provided SNR levels. Noise was estimated to induce an intrinsic 3.38% normalized root-mean-squared-error (nRMSE), serving as a performance bound for SM recovery. Imaging data collected from a Perimag-filled resolution phantom for $[-0.5, -0.5, 1]$ T/m gradients were used. To assess generalizability across time, longitudinally measured SMs with identical imaging protocols albeit spatially-misaligned FOVs were used to construct the training (#8, #9) and test (#10) sets. To assess generalizability across MNP type and SF gradients, an SM sampled using Synomag-D (Micromod GmbH, Germany) and $[-1, -1, 2]$ T/m SF gradients was reserved for training (#7), and another SM using Perimag (Micromod GmbH, Germany) and $[-0.5, -0.5, 1]$ T/m SF gradients was reserved for testing (#10). SM volumes were split into separate slices.

3.4.1.3 In-house MPI Dataset

Data were collected on the open-sided ASELSAN FFL scanner (Fig. 3.3a) [30]. A 32×64 mm² FOV was scanned with an undiluted Perimag sample, with a sample of size 2×2 mm² by rotating the FFL in the x-y plane. SF gradients

Table 3.2: Average nRMSE (%) in HR SM recovery for ablation studies performed under noisy setting at $2\times$ - $8\times$ SR.

	$2\times$	$4\times$	$8\times$
RDSR	0.84%	1.18%	4.32%
CTSR	0.45%	0.70%	4.06%
CNN-ViT	0.67%	0.81%	3.91%
RD-ViT	0.57%	0.76%	4.14%
CNN-CT	0.51%	0.68%	4.31%
TranSMS	0.41%	0.58%	3.71%

of $[0.3, 0.4, 0.5, 0.6]$ T/m were applied, with 9 mT DF amplitude for 10 ms per angle. A 10% duty cycle was used to prevent drive coil overheating. The received signal was amplified (gain = 5) and filtered (cutoffs: 10-300 kHz) on an SR-560 pre-amplifier (SRS, MA, USA), and sampled at 5 MS/s. Frequency components around harmonics 2-to-11 were selected with a 500 Hz bandwidth. High-SNR rows were selected ($\text{SNR} > 5$) and whitened. Noise was estimated to induce an intrinsic 4.36% nRMSE. For retrospective analyses that assessed generalizability across SF gradient strength, the training set contained 3 SMs at $[0.3, 0.4, 0.6]$ T/m, the test set contained 1 SM at 0.5 T/m. To improve reconstruction quality, SMs were cropped to the sensitive region of the receive coil (13×23 grid size). For a prospective analysis to demonstrate the full SNR benefits of TranSMS, we separately measured an LR SM using a larger 4×4 mm² MNP sample, and super-resolved HR SM using the retrospectively trained model. Imaging data were collected by applying a 0.5 T/m SF gradient to cylindrical tube (inner radius: 2 mm, outer radius: 4 mm) phantoms filled with diluted Perimag (Fig. 3.3b). In retrospective experiments, a phantom with two 20-mm-long tubes with 8-12 mm inter-tube spacing was used. In prospective experiments, a 23-mm-long tube, and a 21-mm-long tube with a 5-mm-long central stenosis region of 1-mm inner radius with 6-7 mm inter-tube spacing was used.

3.4.2 TranSMS Implementation

Hyperparameters including number of network layers, number of channels, and learning rate were selected via one-fold cross-validation on the simulated dataset.

Non-overlapping training, validation and test sets were constructed. Parameters were optimized separately for each SR factor ($2\times$, $4\times$, $8\times$) to maximize validation performance in terms of nRMSE. The common set of parameters among SR factors included $n_{RDB} = 4$, $C_C = 24$, $n_{GR} = 6$, $C_T = 64$, $n_A = 4$, $C_{cat} = 48$. Meanwhile, $C_1 = [24, 24, 64]$ and $n_{CL} = [5, 8, 9]$ were selected for $2\times$, $4\times$, and $8\times$ SR, respectively. The transformer module contained $n_{TB} = 3$ stages, where the first two used convolutions of stride 1, and the last stage used stride 2. MLP blocks used Gaussian error linear unit (GELU) activation functions, RDBs used rectified linear unit (ReLU) activation functions, and upsampler blocks used leaky ReLU.

3.4.3 Competing Methods

TranSMS was comparatively demonstrated against several state-of-the-art techniques for SM calibration including traditional, CS-based and SR-based methods.

3.4.3.1 Bicubic & Strided Bicubic

As a traditional method, bicubic interpolation was used to upsample LR SM [78]. Two variants were implemented based on LR SM measurements with large MNP sample size versus strided subsampling of HR SM measurements with small MNP sample size. Real and imaginary parts of SM were treated separately.

3.4.3.2 Compressed Sensing

HR SM measurements were randomly undersampled in 2D, and data were assumed to be sparse in the Fourier domain [19]. An ADMM-based algorithm was used to solve the CS optimization problem [63]. To prevent edge artifacts, the FOV was extended by 4 pixels in all directions, the original FOV was retained in recovered SM.

3.4.3.3 SRCNN

For SR-based recovery, LR SM rows were upsampled with the popular SRCNN model [81, 97]. SRCNN performs bicubic upsampling followed by projection through 3 cascaded convolutional layers. Real and imaginary parts of SM were provided as separate channels at input and output.

3.4.3.4 VDSR

As another SR method, LR SM rows were upsampled with VDSR [82]. VDSR performs bicubic upsampling followed by 20 cascaded convolution layers to better preserve high-frequency information. Real and imaginary parts of SM were provided as separate channels at input and output.

3.4.3.5 2d-SMRnet

Finally, the 3d-SMRnet model in [20] was adapted for 2D SR-based calibration by converting convolution kernels from 3D to 2D. During calibration, 2d-SMRnet performs strided LR SM measurements with a small MNP sample. Real and imaginary parts of SM were provided as separate channels at the model input and output.

3.4.4 Analysis Procedures

3.4.4.1 SM Recovery

DL models were implemented using the PyTorch framework, and trained via the Adam optimizer. For TranSMS, the learning rate was halved at every $N_{epoch}/5$ epochs, where N_{epoch} is number of epochs. For 2d-SMRnet, SRCNN and VDSR, learning rate schedules proposed in the respective manuscripts were adopted [20,

81, 82]. The weight decay parameter was set to 10^{-8} for simulated, and to 0 for experimental datasets as in [20]. The learning rate was selected for each dataset and SR factor via cross-validated search across the range $[10^{-8}, 1]$. Bicubic and CS methods were implemented in MATLAB (MathWorks, Natick USA). Bicubic interpolation assumed default settings in the *imresize* function. In CS, intensity of each whitened SM row was normalized separately for numerical stability, and restored to its original scale following recovery. Step size $\mu = 1$ and # of iterations = 1000 were selected via cross-validation. The bound on the noise level for the constrained problem was estimated from background measurements.

Methods were compared in super-resolving LR SM. Background measurements were used to estimate σ_i in Eq. (3.20), and to perform background subtraction. In the simulated dataset, LR SM was simulated for a large MNP sample spanning an LR voxel. Gaussian noise was added to maintain 30 dB SNR [30]. Networks were trained to estimate noise-free HR SM given noisy LR SM. We assessed performance at $2\times-8\times$ SR, reducing calibration time by 4–64-fold, respectively.

In experimental datasets, independent LR and HR SM measurements can carry inconsistencies due to nuisance variability in the system (e.g., sample misalignment, temperature drift), biasing trained models and degrading recovery performance. To alleviate this issue, LR SMs were generated retrospectively using the linearity of the MPI signal, through box-car downsampling of HR SMs to permit unbiased reporting of performance metrics. Training sets were expanded via data augmentation by horizontal and vertical flipping of SM rows. Networks were trained to estimate measured HR SM given LR SM, considering $2\times-8\times$ SR. To prospectively demonstrate TransSMS, an LR SM measured on the in-house scanner using a large MNP sample spanning an LR voxel was super-resolved using the retrospectively trained model.

Table 3.3: Average nRMSE (%) in HR SM recovery for the simulated dataset under (noiseless/noisy) settings at $2\times$ - $8\times$ SR.

	$2\times$	$4\times$	$8\times$
Bicubic	3.53% / 4.02%	14.82% / 14.95%	46.88% / 46.92%
Bicub. (str.)	14.9% / 16.6%	42.61% / 50.20%	80.20% / 93.34%
CS	4.15% / 9.81%	23.58% / 37.94%	80.60% / 101.39%
SRCNN	33.5% / 38.8%	36.48% / 35.84%	53.93% / 50.00%
VDSR	2.66% / 6.51%	5.49% / 13.36%	20.02% / 48.16%
2d-SMRnet	0.12% / 1.34%	0.36% / 5.30%	1.11% / 64.02%
TranSMS	0.09% / 0.41%	0.13% / 0.58%	0.40% / 3.71%

3.4.4.2 Image Reconstruction

Recovered HR SMs were used to perform image reconstruction. Among advanced methods [62, 73, 98], we opted for an ADMM-based algorithm with fast convergence implemented in PyTorch [15, 39]. Data were whitened using background measurements prior to reconstruction, so ϵ was set to $\sqrt{M}$ where M is the number of HR SM rows. Reconstruction using the reference HR SM was taken as a benchmark for performance evaluations. Hyperparameters were selected via cross-validated search to optimize quality of the reference image in the simulated dataset. The selected parameters were $\alpha_1 = 0.95$, $\alpha_{TV} = 0.05$, step size of $\mu = 10$, and 5000 iterations.

3.4.5 Quantitative Assessments

Performances in HR SM recovery and image reconstruction were assessed via quantitative metrics. Assuming $\hat{\mathbf{A}}$ is the recovered HR SM and $\mathbf{A}$ is the reference HR SM, the data-consistency constraint in Eq. (3.3) can be expressed as:

$$\|\hat{\mathbf{A}}\mathbf{x} - \mathbf{y}\|_2^2 = \|\mathbf{R}\mathbf{x}\|_2^2 + \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + 2\mathbf{x}^T \mathbf{R}^T (\mathbf{A}\mathbf{x} - \mathbf{y}) \quad (3.24)$$

where $\mathbf{R} = \hat{\mathbf{A}} - \mathbf{A}$ denotes the residual error in HR SM recovery. Given zero-mean measurement noise, the expected value of $\mathbf{A}\mathbf{x} - \mathbf{y}$ is 0. Data-fidelity error is then proportional to the Frobenius norm of the residual error, $\|\mathbf{R}\|_F^2$. Thus,

Table 3.4: Average pSNR (dB) in image reconstruction for the simulated dataset under noisy setting at $2\times$ - $8\times$ SR.

	$2\times$	$4\times$	$8\times$
Bicubic	24.85	14.05	9.70
Bicubic (str.)	15.00	9.86	8.36
CS	19.92	-4.15	-14.20
SRCNN	16.56	12.39	-8.25
VDSR	23.07	17.32	-9.62
2d-SMRnet	24.77	20.38	-11.60
TranSMS	26.35	25.72	20.89

we measured the quality of HR SM recovery via nRMSE:

$$nRMSE(\hat{\mathbf{A}}) = \|\hat{\mathbf{A}} - \mathbf{A}\|_F / \|\mathbf{A}\|_F \quad (3.25)$$

For image reconstruction, we adopted the popular peak Signal-to-Noise-Ratio (pSNR). Assuming $\mathbf{x}_{ref}$ is the reference image:

$$pSNR(\mathbf{x}) = 20 \log_{10} \left(\sqrt{N} \|\mathbf{x}_{ref}\|_{\infty} / \|\mathbf{x} - \mathbf{x}_{ref}\|_2 \right) \quad (3.26)$$

3.5 Results

3.5.1 Ablation Studies

3.5.1.1 Network Ablation

We first demonstrated the role of each network component in TranSMS. Analyses were performed on the simulated dataset where a noise-free ground truth is available for absolute performance evaluation. First, we compared TranSMS against variants where the convolutional transformer (CT) module was ablated (residual-dense super-resolution; RDSR) or the residual-dense (RD) convolutional module was ablated (convolutional transformer super-resolution; CTSR). Second, we evaluated the specific implementations of transformer and convolutional modules. To evaluate the benefit of CT blocks, we implemented a variant containing

vanilla vision transformer (ViT) modules (residual-dense vision transformer; RD-ViT). To evaluate the benefit of RD blocks, we implemented a variant with vanilla CNN modules (CNN with convolutional transformer, CNN-CT). The combination of vanilla CNN and vanilla transformer was also considered (CNN with vision transformer, CNN-ViT). Table 3.2 lists performance for TranSMS and its variants. Overall, TranSMS yields higher performance than both RDSR and CTSR, indicating the importance of its transformer and convolutional modules. TranSMS outperforms RD-ViT indicating the benefit of CT over ViT, and it also outperforms CNN-CT indicating the benefit of RD over CNN. Finally, we evaluated the benefit of the DC module by comparing TranSMS variants with and without DC projection. Fig. 3.4 plots nRMSE for $4\times$ SR, showing that the DC module enables faster and more stable convergence.

3.5.1.2 Training Set Ablation

Next, we assessed the influence of the number and diversity of the SMs in the training set on model performance. These analyses were conducted on the simulated dataset. First, we varied the number of SMs used for training in the range [6 66] via random selection, while the original test set was maintained. Figure 3.5 shows nRMSE in SM recovery as a function of the number of training SMs. Near-optimal results are achieved around 15 SMs with diminishing returns for larger training sets. Second, we compared two training sets with 6 SMs each, where one was restricted to a narrow range of SF and MNP parameters (gradient strength in T/m, diameter in nm): (0.3, 18.39), (0.3, 20.53), (0.3, 24.82), (0.7, 16.24), (0.7, 22.67), (0.7, 24.82). The other training set was randomly selected. We find that the restricted and randomized training sets yield moderate performance differences, with an average nRMSE of 2.54% for randomized and 3.42% for restricted sets.

3.5.2 Analyses on Simulated Dataset

3.5.2.1 SF & MNP Tracer Variability

We first demonstrated TranSMS on the simulated dataset for HR SM recovery at $2\times$ - $8\times$ SR. We assessed reliability against variability in SF gradients and MNP diameter between training and test sets, under both noise-free and noisy settings. Table 3.3 lists nRMSE of recovered HR SMs. TranSMS yields superior performance against competing methods in all cases, with 21.96% lower nRMSE over the second-best method on average in noisy settings. Representative SM rows are displayed in Fig. 3.6 for $4\times$ SR. Bicubic interpolation and VDSR have visible blurring, CS shows prominent artifacts and noise, and SRCNN shows edge artifacts and blurring. While 2d-SMRnet is often the top contender, it has a high error rate at $8\times$ in the noisy setting. This could be attributed to the relatively high complexity of 2d-SMRnet rendering it more susceptible to overfitting with noisy training data.

We then investigated the effects of recovered HR SMs on image reconstruction performance. Table 3.4 lists pSNR for data simulated at 30 dB SNR. TranSMS outperforms all competing methods with 6.04 dB higher pSNR over the second-best method on average. Representative phantom images at $4\times$ SR are shown in Fig. 3.7. DL methods outperform CS in terms of residual artifacts, and bicubic interpolation in terms of visual acuity. Among DL models, TranSMS offers the reconstruction with the highest similarity to the reference image, whereas 2d-SMRnet is the second best with occasional intensity errors. Methods successful in SM recovery generally yield higher downstream reconstruction performance.

3.5.2.2 Super-Resolving Higher Grid-Size

To assess the performance of TranSMS in recovering a higher grid size than the reference HR SM it has been trained to recover, we trained TranSMS to estimate

Table 3.5: Average nRMSE (%) in HR SM recovery for Open MPI at $2\times$ - $8\times$ SR. Calibration scans were performed at different times and with unaligned FOVs for SMs in training and test sets.

	$2\times$	$4\times$	$8\times$
Bicubic	4.55%	18.13%	52.02%
Bicubic (str.)	16.86%	47.41%	92.08%
CS	8.81%	51.48%	101.31%
SRCNN	53.65%	53.16%	79.29%
VDSR	4.55%	12.62%	39.19%
2d-SMRnet	19.69%	48.94%	75.37%
TranSMS	3.15%	6.19%	20.58%

Table 3.6: Average nRMSE (%) in HR SM recovery / pSNR (dB) in image reconstruction for Open MPI at $2\times$ - $8\times$ SR. SF gradient strength and MNP tracer differed for SMs in training and test sets.

	$2\times$	$4\times$	$8\times$
Bicubic	4.55% / 34.77	18.13% / 17.15	52.02% / 11.34
Bicubic (str.)	16.86% / 33.30	47.41% / 18.62	92.08% / 13.88
CS	8.81% / 33.89	51.48% / 20.47	101.31% / 16.12
SRCNN	50.88% / 32.21	62.81% / 19.39	106.76% / –
VDSR	3.34% / 38.23	11.83% / 30.59	113.81% / 10.01
2d-SMRnet	6.85% / 33.79	17.22% / 24.81	78.88% / 12.60
TranSMS	3.32% / 38.54	10.66% / 31.96	114.45% / 13.38

32×32 HR SM given 16×16 LR SM. We then used the trained model to super-resolve a 64×64 HR SM by processing overlapping 16×16 patches in a given 32×32 SM. Representative results for recovered SMs at 30 dB SNR are shown in Fig. 3.8. TranSMS yields highly similar SM estimates to the reference HR SM although it has not been trained using data at the target SM resolution. Note that prominent block artifacts are evident when the DC module is ablated, suggesting the importance of explicit DC projections for out-of-domain generalization.

3.5.3 Analyses on Open MPI Dataset

3.5.3.1 Temporal & FOV Variability

Next, we demonstrated TranSMS for HR SM recovery in Open MPI using two separate analyses with different training sets albeit a common test set. First, we assessed reliability against variability across time and FOV positioning between the training and test SMs. Table 3.5 lists nRMSE when the training and test sets contained SMs measured with the same protocol albeit at different time points and with unaligned FOVs. TranSMS yields superior performance against competing methods, with 8.81% lower nRMSE on average over the closest competitor.

3.5.3.2 SF & MNP Tracer Variability

We also assessed reliability against variability in SF gradients and MNP tracer. Table 3.6 lists respective nRMSE in HR SM recovery when the training and test SMs differed in their SF and MNP parameters. TranSMS performs better than competing methods, except for $8\times$ SR where all methods perform poorly and bicubic interpolation yields the lowest nRMSE. TranSMS lowers nRMSE by 0.60% over the second-best method across $2\times$ - $4\times$ SR.

Recovered HR SMs at $4\times$ SR are displayed in Fig. 3.9. TranSMS generates a more similar estimate to the reference HR SM than competing methods. Bicubic interpolation suffers from visible blurring, and CS shows artifacts. SRCNN performs poorly with visible boundary artifacts. While 2d-SMRnet achieves high spatial acuity, it shows some geometric warping. VDSR yields a geometrically accurate estimate with some intensity errors. In contrast, TranSMS yields an HR SM of high spatial acuity and consistency to the reference HR SM. Table 3.6 lists pSNR for subsequent image reconstructions except for SRCNN that did not converge and returned an empty image at $8\times$ SR. On average, TranSMS outperforms the closest competitor by 1.67 dB pSNR. A representative cross-section from the reconstructed phantom at $4\times$ SR is shown in Fig. 3.10, where methods

with successful SM recovery demonstrate improved reconstruction results.

3.5.4 Analyses on In-house MPI Dataset

3.5.4.1 SF Variability

Next, we demonstrated TranSMS on the in-house MPI dataset acquired on an FFL scanner. We first assessed reliability against variability in SF gradients between SMs in the training versus test sets. Table 3.7 lists nRMSE in SM recovery. TranSMS outperforms competing methods, with elevated benefits at higher SR factors. On average, it yields 3.99% lower nRMSE than the second-best method. Recovered SMs are displayed in Fig. 3.11 at $4\times$ SR. TranSMS yields the closest estimates to the reference HR SM, while 2d-SMRnet offers the second-best performance. Table 3.7 lists pSNR for image reconstruction using recovered HR SMs. TranSMS yields the highest pSNR, outperforming the closest competitor by 1.59 dB pSNR on average. Reconstructions at $4\times$ SR are shown in Fig. 3.12 for the phantom in Fig. 3.3b. TranSMS produces a reconstruction with a high degree of similarity to the reference phantom image.

Table 3.7: Average nRMSE (%) in HR SM recovery / pSNR (dB) in image reconstruction for the in-house dataset at $2\times$ - $8\times$ SR.

	$2\times$	$4\times$	$8\times$
Bicubic	10.97% / 16.25	36.26% / 13.75	69.76% / 4.25
Bicubic (str.)	25.99% / 15.84	68.40% / 12.88	101.52% / 5.63
CS	15.31% / 17.30	62.54% / 7.68	104.00% / 2.59
SRCNN	43.79% / 13.78	75.40% / 6.46	437.66% / 6.08
VDSR	7.46% / 20.24	14.78% / 18.39	30.46% / 10.21
2d-SMRnet	6.65% / 22.71	11.27% / 19.38	56.06% / 6.22
TranSMS	6.03% / 24.89	8.92% / 19.68	21.47% / 12.51

3.5.4.2 Prospective Experiment

Finally, we demonstrated TranSMS on an LR SM prospectively measured using a large MNP sample spanning an LR voxel. The model trained on retrospectively downsampled LR SMs was used for inference. HR SMs recovered from retrospectively versus prospectively sampled LR SMs are displayed in Fig. 3.13 at $2\times$ SR. As expected, HR SM based on prospectively sampling shows lower noise levels. Reconstructions using recovered HR SMs are shown in Fig. 3.14. Reconstructions based on prospective sampling better preserve the stenosis in the phantom compared to retrospective sampling. A quantitative metric cannot be reported for prospective sampling due to absence of ground truth. The theoretical improvement in measurement SNR with prospective sampling is 6dB at $2\times$ SR. Based on simulation analyses, a crude expectation is that this improvement will translate onto 1-2dB difference in reconstruction performance (results not shown).

3.6 Discussion

To accelerate MPI calibration, here we proposed to recover HR SM via a vision transformer model given LR SM measured with large MNP samples. Clear benefits were observed over strided or randomly undersampled measurements with small MNP samples [65, 77, 78]. In the simulated dataset, LR SMs were obtained by prospective downsampling of HR SMs by a factor of S in each dimension. Signal power increases S^4 -fold while the noise power remains constant, improving SNR by $10\log_{10}(S^4)$, e.g., ~ 24 dB at $4\times$ SR. In experimental datasets, we trained TranSMS on LR SMs obtained by retrospective downsampling of HR SMs to mitigate potential biases due to nuisance variability between SM measurements. In that case, noise power increases S^2 -fold, limiting the SNR improvement to $10\log_{10}(S^2)$, e.g., ~ 12 dB at $4\times$ SR. In prospective demonstrations with an LR SM measured using a large MNP sample, further benefits were observed due to improved SNR. An imaging system with robust mechanical positioning might further permit prospectively sampled LR SMs to be used for model training.

Alternatively, a hybrid SM calibration can be performed to measure the MNP response via a magnetic particle spectrometer (MPS) where sample position is emulated using an offset field [17], or physical movement of the sample can be avoided via homogeneous focus fields to move the FFR [99]. Such SM measurements might also be super-resolved with TranSMS for further increasing the achievable resolution in HR SM.

Here, we observe that HR SM recovery performance is higher in simulated versus experimental data, likely due to several factors. First, retrospective sampling of LR SM limits SNR benefits on experimental data. While non-stationary system parameters between separate measurements including spatial misalignment render quantitative evaluations of prospective sampling challenging, our preliminary results suggest further improvements in this case. Second, the training sets are relatively compact for experimental datasets due to practical limitations regarding SM measurement, whereas the simulated dataset compiles a diverse collection of SMs that helps boost performance of learning-based models. The ablation study on training set size suggests diminishing returns beyond a certain number of training samples, yet future work is warranted to find the precise number of SMs needed for near-optimal performance in experimental datasets. Third, our simulations do not consider nonidealities such as relaxation that can lower task difficulty. Other influential factors include SF gradient strength that controls the spatial rate of change in SM, and MNP diameter that alters sensitivity to magnetic field and broadening of the response from FFR. In this work, we find elevated SM recovery errors at relatively high frequency components that show high spatial frequency information. This observation implies that factors that increase spatial resolution would also elicit increased task difficulty. Thus, in theory, task difficulty should elevate with higher SF gradients and larger MNP diameters.

All methods naturally show performance degradation towards higher SR factors as the spatial information in input measurements is constrained. For each method, performance becomes notably poor (e.g., $>50\%$ nRMSE, $<10\text{dB}$ PSNR) beyond an SR factor, which inevitably depends on the characteristics of SM measurements and varies across datasets. Yet, TranSMS generally shows higher

resilience than competing methods against high SR factors. CS theory guarantees signal recovery for a near-isometric sampling matrix, and the recovery error depends critically on the number of measurements and SNR level. Here, we observe that CS performs competitively at $2\times$ SR (i.e., 4-fold acceleration), albeit poorly at $4\times$ - $8\times$ SR (i.e., 16-fold to 64-fold acceleration). Our results are aligned with a recent study that assumes similar measurement SNRs for 2D imaging [63]. While performant recovery using CS has been reported for 3D imaging under higher SNRs [19, 80], different imaging scenarios are considered here. In Open MPI, we observed that 2d-SMRnet shows relatively higher performance differences between the two analyses that examine SF and MNP variability versus temporal variability compared to other methods. This could reflect an interaction between the sampling scheme and the training set. 2d-SMRnet uses strided LR SMs, and other methods use box-car downsampled LR SMs. In Open MPI, the training and test sets that we used for temporal variability analysis were collected under matching system parameters. In contrast, the training set that we used for SF and MNP variability analysis were collected with half-pixel shifts in the sampling grid, differing from the data that we used for the test set. The half-pixel shift during SM acquisition generates spatially overlapping measurements in the latter case, which does not strictly obey the downsampling model that TranSMS assumes. In contrast, these overlaps are not a concern for the strided measurements used by 2d-SMRnet. These differences likely contribute to the relative performance boost for 2d-SMRnet.

Prior CS methods for accelerated MPI calibration perform randomly under-sampled HR SM measurements with a small MNP sample, and recover SM by enforcing sparsity in a transform domain [18, 19]. In contrast, TranSMS performs LR SM measurements with a large MNP sample to improve measurement SNR, and a data-driven recovery is performed between LR and HR SM instead of relying on sparsity assumptions. An alternative for accelerated calibration is based on SR. In [20], HR SM measurements are uniformly subsampled in a strided fashion, and HR SM is then recovered via a CNN. In [65], an LR SM is supplemented via an analytically-simulated HR SM, and a nonlinear optimization problem is solved to directly reconstruct an MPI image. Instead, we perform LR SM measurements

with large MNP samples to improve SNR efficiency, we rely solely on measured SMs for reliability against system imperfections, and we leverage a novel vision transformer for SM recovery.

It is possible to adopt the sampling strategy of TranSMS in previous CS- and SR-based methods. Note, however, that CS relies on measurements with incoherent interference in the sparsifying-transform domain. Large MNP samples would result in more structured spatial sampling, elevating coherence and degrading recovery. On the other hand, learning-based SR methods such as 2d-SMRnet can benefit from improved SNR efficiency due to large MNP samples similar to TranSMS.

Voxel-sized undiluted MNP samples help improve SNR during calibration. Instead, clinically admissible intravenous administration of MNPs result in notably lower concentration per voxel during imaging. Particle interactions may alter MPI signal characteristics for highly concentrated samples, eliciting a discrepancy between calibration and imaging experiments [100]. Because TranSMS enables SNR benefits due to larger MNP sample spanning an LR voxel, it might mitigate this issue by allowing calibration with diluted MNP samples.

Several developments can improve performance of TranSMS. TranSMS can be trained to perform SR on multiplexed LR SM measurements with multiple MNP samples to further calibration efficiency [63]. Instead of using a uniformly filled LR voxel, optimizing the spatial distribution within an LR voxel could also increase the level of spatial information captured during calibration, and thereby improve SM recovery. Furthermore, measuring HR SM with larger MNP samples can increase SNR of reference images used to train TranSMS. However, the measured HR SM would be blurred due to box-car downsampling, so the DC module would have to be modified and deconvolution would be required to resolve HR SM at the intended resolution.

3.7 Conclusion

In this work, we introduced a novel MPI calibration approach that captures an LR SM with up to 64-fold larger MNP sample that spans across an LR voxel for improving measurement SNR. A novel transformer model is then leveraged to sensitively recover HR SM. TranSMS achieves lower error in SM recovery and higher quality in image reconstruction compared to state-of-the-art CS and SR methods. Therefore, it is a promising candidate to improve practicality of MPI.

3.8 Appendix

Proof: Let us first define $\epsilon_i = \sqrt{m}\sigma_i$, $\mathbf{c}_i = \mathbf{a} - \tilde{\mathbf{a}}_i$ and $\tilde{\mathbf{b}}_i = \mathbf{b}_i - \mathbf{D}\tilde{\mathbf{a}}_i$. The associated Lagrangian is,

$$L(\mathbf{c}_i, \lambda) = \|\mathbf{c}_i\|_2^2 - \lambda \left(\epsilon_i^2 - \|\mathbf{D}\mathbf{c}_i - \tilde{\mathbf{b}}_i\|_2^2 \right). \quad (3.27)$$

Considering the Lagrangian function, we get two conditions. In the first case, if $\|\tilde{\mathbf{b}}_i\|_2 \leq \epsilon_i$, then $\mathbf{c}_i = 0$ and $\hat{\mathbf{a}}_i = \tilde{\mathbf{a}}_i$. In the second case, we get the optimality conditions as:

$$0 = \mathbf{c}_i + \lambda \mathbf{D}^T (\mathbf{D}\mathbf{c}_i - \tilde{\mathbf{b}}_i), \quad \epsilon_i^2 = \|\mathbf{D}\mathbf{c}_i - \tilde{\mathbf{b}}_i\|_2^2 \quad (3.28)$$

Now using Eq. (3.28),

$$\lambda \mathbf{D}^T \tilde{\mathbf{b}}_i = (\lambda \mathbf{D}^T \mathbf{D} + \mathbf{I}) \mathbf{c}_i, \quad (3.29)$$

$$\mathbf{c}_i = \lambda (\lambda \mathbf{D}^T \mathbf{D} + \mathbf{I})^{-1} \mathbf{D}^T \tilde{\mathbf{b}}_i. \quad (3.30)$$

Using the Woodbury matrix identity and $\mathbf{D}\mathbf{D}^T = \mathbf{I}$:

$$\mathbf{c}_i = \lambda/(\lambda + 1) \mathbf{D}^T \tilde{\mathbf{b}}_i. \quad (3.31)$$

Now, let us put $\mathbf{c}_i$ back into Eq. (3.28):

$$\epsilon_i^2 = \|\lambda/(\lambda + 1) \mathbf{D}\mathbf{D}^T \tilde{\mathbf{b}}_i - \tilde{\mathbf{b}}_i\|_2^2, \quad (3.32)$$

$$\lambda/(\lambda + 1) = 1 - \epsilon_i/\|\tilde{\mathbf{b}}_i\|_2 \quad (3.33)$$

Finally, the result becomes:

$$\mathbf{c}_i = \left(1 - \epsilon_i / \|\tilde{\mathbf{b}}_i\|_2\right) \mathbf{D}^T \tilde{\mathbf{b}}_i, \quad (3.34)$$

$$\hat{\mathbf{a}}_i = \tilde{\mathbf{a}}_i + \left(1 - \sqrt{m} \sigma_i / \|\tilde{\mathbf{b}}_i\|_2\right) \mathbf{D}^T \tilde{\mathbf{b}}_i. \quad (3.35)$$



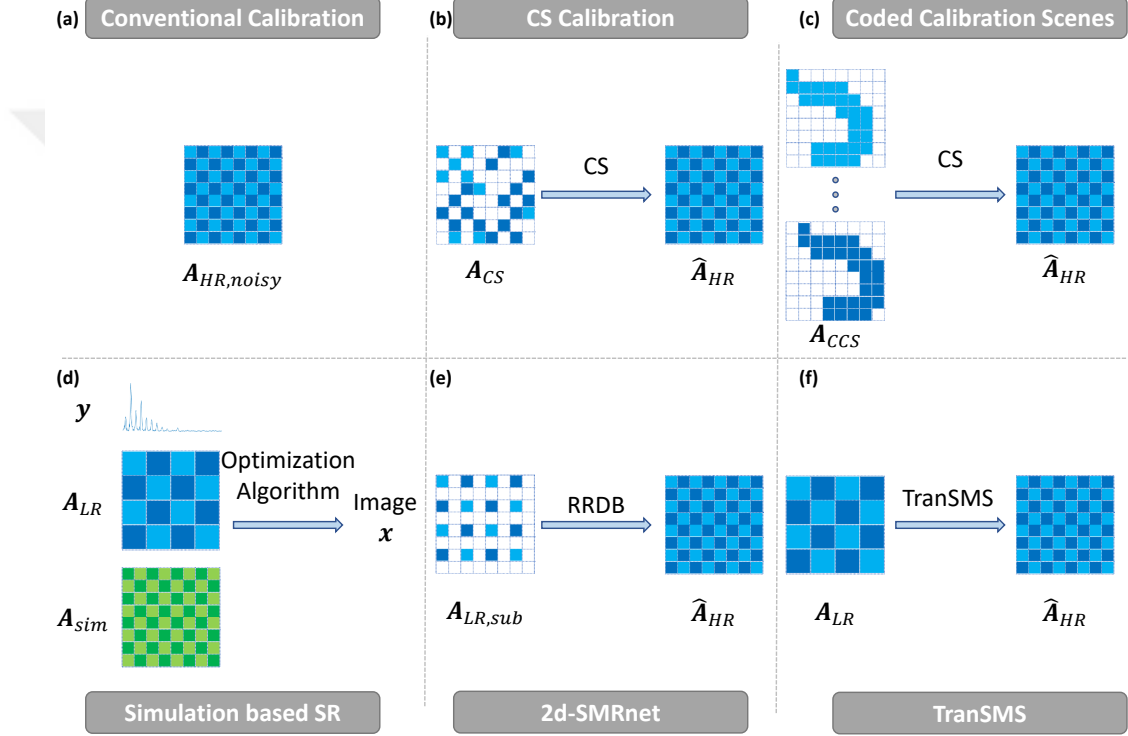


Figure 3.1: Overview of prior and proposed techniques for SM calibration. **(a)** Conventional calibration samples all of the grid points in HR SM. **(b)** CS calibration randomly subsamples HR SM and uses CS optimization for recovery. **(c)** Coded calibration scenes samples HR SM using a specially tailored calibration scene and uses CS optimization for recovery. **(d)** Simulation based SR jointly solves an optimization problem involving both a simulated SM and a sampled LR SM as well as the phantom to be imaged. **(e)** 2d-SMRnet uniformly subsamples HR SM in a strided fashion and recovers HR SM via a CNN. **(f)** TranSMS performs LR SM measurements with a larger MNP sample that fills an LR grid, and recovers HR SM via a transformer.

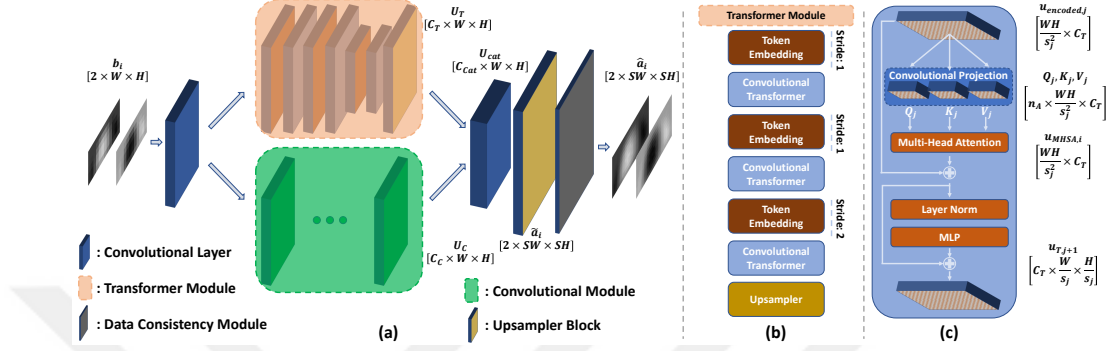


Figure 3.2: (a) Schematic of the network architecture in TransSMS. After an initial convolutional layer, encoded feature maps are fed to parallel transformer and convolutional modules. The transformer module extracts contextual representations over longer distances, while the convolutional module captures localized representations. Contextual and localized feature maps are fused and upsampled to the resolution of the HR SM. Finally, a data-consistency module enforces fidelity of recovered HR SM to input LR SM. (b) The transformer module contains a cascade of token-embedding and convolutional transformer sub-blocks, and a final upsampler to restore dimensionality. (c) The transformer performs convolutional token projection, followed by MHA and MLP blocks to extract contextual features.

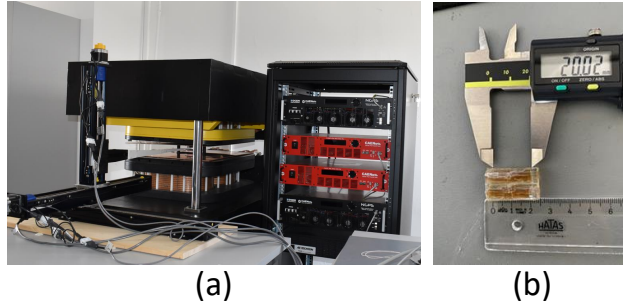


Figure 3.3: (a) In-house open-sided FFL scanner. (b) The imaging phantom used to acquire in-house MPI data. Cylindrical tubes were prepared with Perimag (Micromod GmbH, Germany) MNPs at a dilution ratio of 1:20 (405 $\mu\text{g}/\text{ml}$ Fe). The length, inner radius and outer radius of the tubes were 20 mm, 2 mm and 4 mm, respectively.

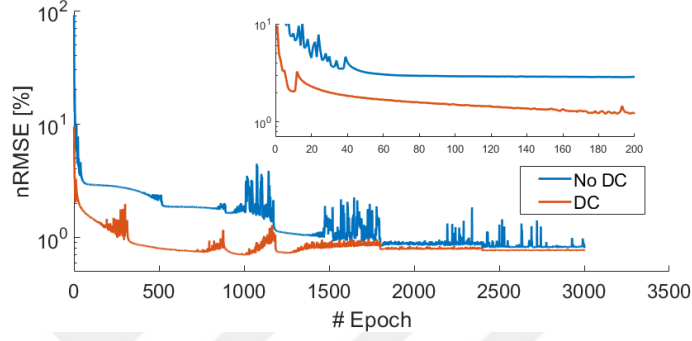


Figure 3.4: Validation performance of TranSMS with (red curve) and without (blue curve) the data-consistency (DC) module. The HR SM recovery performance is reported as nRMSE across training epochs. The upper-right panel displays a zoomed-in portion of nRMSE curves during the initial stages of training.

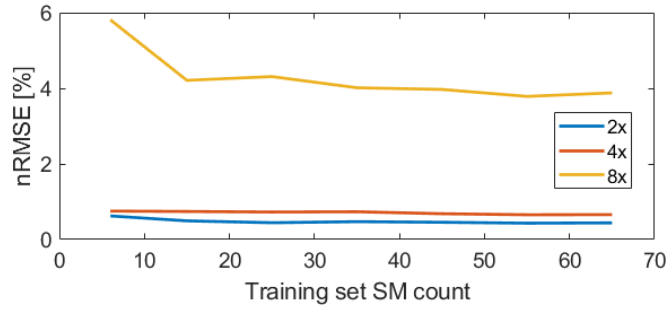


Figure 3.5: SM recovery performance of TranSMS as a function of the number of SMs included in the training set for SR factors of $2\times$ - $8\times$.

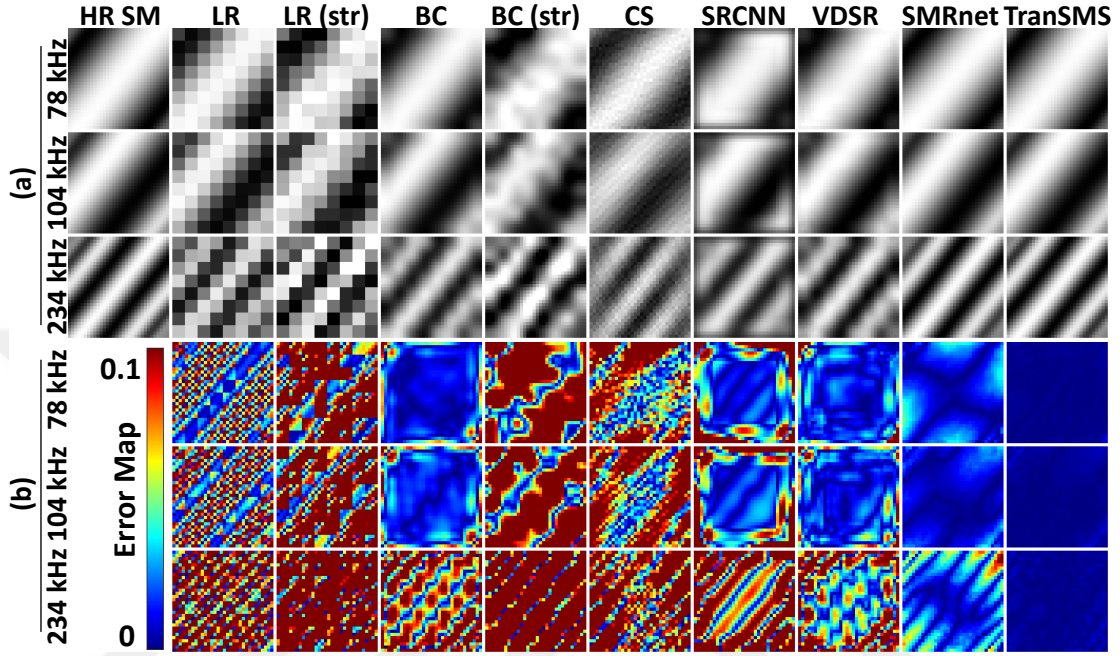


Figure 3.6: HR SM recovery in the simulated dataset for $4\times$ SR using noisy measurements. SF gradient strength and MNP size differed between the training and test sets. SM rows corresponding to different harmonics at 45° are shown. (a) Recovered HR SMs are displayed along with the input LR SM, and input strided (str) LR SM. Strided bicubic (BC) interpolation and 2d-SMRnet receive as input the strided LR SM. (b) Respective error maps (see colorbar).

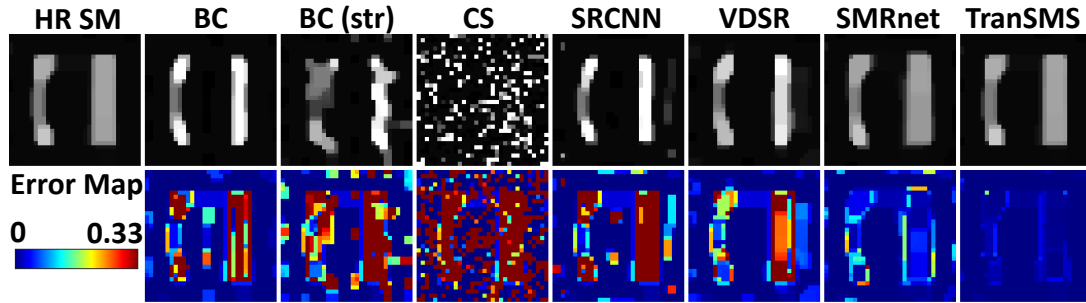


Figure 3.7: Phantom images from the simulated dataset at 1 T/m SF gradient strength reconstructed with recovered HR SMs at $4\times$ SR are displayed along with the reference reconstruction using the reference HR SM (top row). Error maps for each method are also displayed (bottom row). The phantom comprised two 22-mm long tubes, the right tube's width fixed at 6 mm, and the left tube's width ranging in 2-6 mm. Inter-tube spacing varied in 8-12 mm.

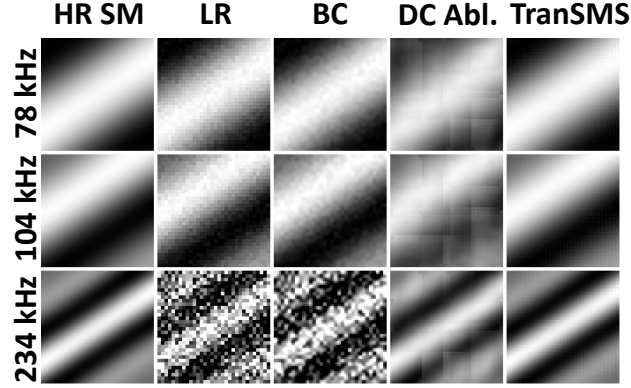


Figure 3.8: A pre-trained TranSMS model was used to super-resolve a 64×64 HR SM by processing overlapping 16×16 patches in the input 32×32 LR SM. SM rows for different harmonics at 45° are shown. Recovered HR SMs are shown under 30 dB measurement SNR, along with the input LR SM and the reference HR SM.

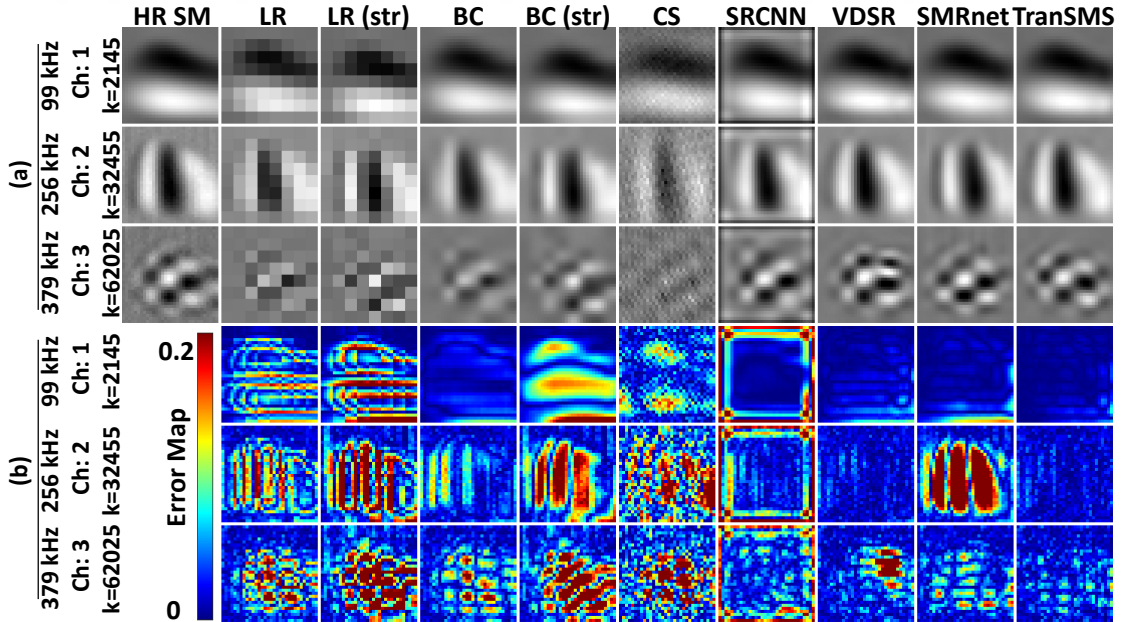


Figure 3.9: HR SM recovery in Open MPI for $4 \times$ SR. Both the SF gradient strength and MNP tracer differed for SMs in training and test sets. The SM rows corresponding to the center slice at different frequency components and receive channels (Ch) are shown (k denotes the respective SM row index). (a) Recovered HR SMs and (b) respective error maps are displayed along with the input LR SM, input strided LR SM and reference HR SM.

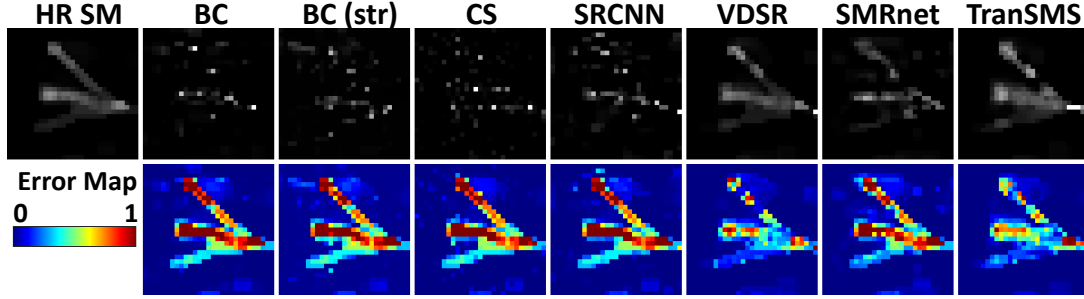


Figure 3.10: Image reconstruction in Open MPI for $4\times$ SR. Phantom images reconstructed with recovered HR SMs are displayed along with the benchmark reconstruction using the reference HR SM (top row). Respective error maps are also displayed (bottom row).

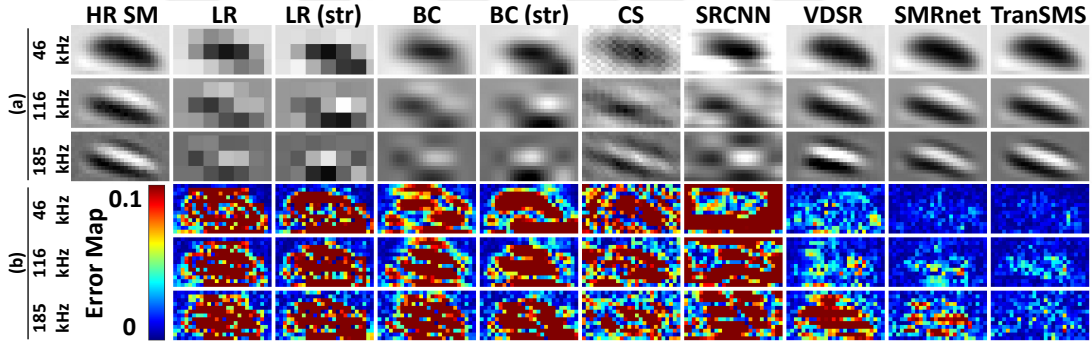


Figure 3.11: HR SM recovery in the in-house MPI dataset for $4\times$ SR. The SM rows corresponding to different harmonics at 81° are shown. (a) Recovered HR SMs and (b) respective error maps for each competing method are displayed along with the input LR SM, input strided LR SM and reference HR SM.

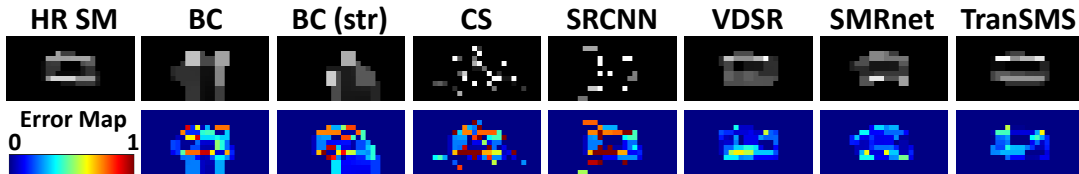


Figure 3.12: Phantom images in the in-house dataset reconstructed with recovered HR SMs at $4\times$ SR and respective error maps are displayed, along with the benchmark reconstruction using the reference HR SM. The length, inner radius and outer radius of the tubes were 20 mm, 2 mm and 4 mm, respectively.

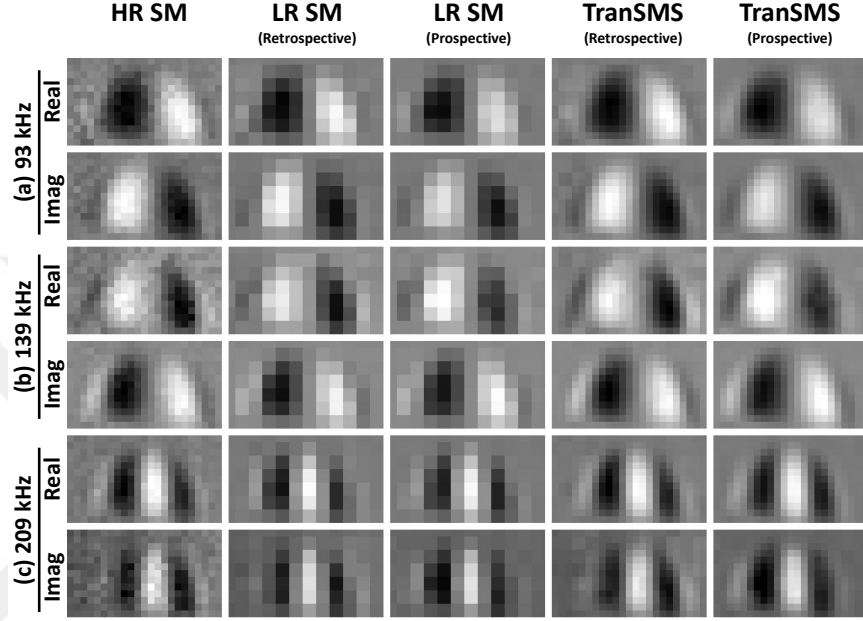


Figure 3.13: Prospective versus retrospective HR SM recovery in the in-house dataset for $2\times$ SR. Real and imaginary parts of the SM rows corresponding to different harmonics ((a) – (c)) at 0° are shown.

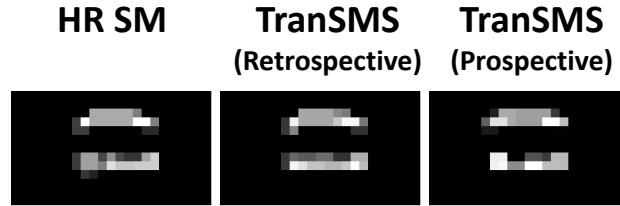


Figure 3.14: Phantom images for the in-house dataset based on prospectively versus retrospectively recovered HR SMs are shown along with the benchmark reconstruction using the reference HR SM. The upper tube is 23-mm long, the lower tube is 21-mm long with a 5-mm-long central stenosis region of 1-mm inner radius (1 pixel : 2 mm).

Chapter 4

Deep Equilibrium based Image Reconstruction for Magnetic Particle Imaging

The contents of this chapter reflect the work reported in the following publications:

- Baris Askin, Alper Güngör, Damla Alptekin Soydan, Emine Ulku Saritas, Can Barış Top, Tolga Cukur, “PP-MPI: A Deep Plug-and-Play Prior for Magnetic Particle Imaging Reconstruction,” In: Machine Learning for Medical Image Reconstruction. MLMIR 2022. Lecture Notes in Computer Science, vol 13587. Springer, Cham. https://doi.org/10.1007/978-3-031-17247-2_11.
- Alper Güngör, Baris Askin, Damla Alptekin Soydan, Can Barış Top, Emine Ulku Saritas, Tolga Cukur, “DEQ-MPI: A Deep Equilibrium Reconstruction with Learned Consistency for Magnetic Particle Imaging,” arXiv preprint, 2022.
- Alper Güngör, Baris Askin, Damla Alptekin Soydan, Can Barış Top, Emine

Ulku Saritas, Tolga Cukur, “DEQ-MPI: A Deep Equilibrium Reconstruction with Learned Consistency for Magnetic Particle Imaging,” *submitted to Transactions on Medical Imaging*, 2022.

4.1 Introduction

In magnetic particle imaging (MPI), magnetic nanoparticle (MNP) type and scanner hardware influence the spatially-varying response across the field-of-view (FOV). To account for these variations, a system matrix (or analogously a point spread function) is typically utilized [13, 22, 58]. The system matrix (SM) can be either analytically estimated [65] or experimentally measured for reliability against non-idealities [15, 16]. The SM is then used to setup an ill-posed problem to reconstruct images from acquired data [22].

Traditional reconstructions seek a solution that embodies both the physical constraints related to the SM and acquired data, and the attributes of high-quality MPI images. While non-iterative solvers exist [101, 102], optimization algorithms are prominent that iteratively enforce data consistency (DC) based on the SM and regularize the image [15, 62]. For DC, hand-crafted measures based on energy or intensity differences between reconstructed and acquired data samples are common [22, 23]. For regularization, hand-crafted priors are used to promote desired attributes (e.g., smoothness, sparsity) via ℓ_2 [24, 25], ℓ_1 [98], TV [15] losses or their combinations [26]. While they have been pervasive in MPI reconstruction, hand-crafted priors cannot fully capture the image distribution and their performance depends on careful tuning of regularization weights that can vary substantially across scans [26, 30, 103].

In recent years, learning-based priors have received interest in MPI reconstruction as a powerful alternative. Purely data-driven methods train neural networks to directly recover images from frequency- or time-domain data [27–29, 104]. Although they enable efficient inference, neglecting the physical constraints embodied in the SM can limit generalizability. To improve generalization, deep

image prior (DIP) methods instead use untrained networks whose parameters are learned at test time by minimizing a DC loss [73, 103]. Yet, extensive inference times and difficulty in identifying appropriate architectures per dataset can limit utility [74]. A recent plug-and-play method (PP-MPI) pre-trains an image prior for denoising and later combines it with the SM for reconstruction [56]. While it offers a compromise between efficiency and generalization, transferring a prior from the denoising to the reconstruction task can potentially limit performance.

Here, we introduce a novel deep equilibrium model, DEQ-MPI, for improved performance and efficiency in MPI reconstruction. Inspired by physics-driven unrolling methods [43], DEQ-MPI augments neural networks into an iterative optimization to rapidly alternate between regularization and DC projections. Conventional unrolled methods produce non-convergent solutions following a small number of iterations due to computational and memory constraints [49], and they use hand-crafted DC measures that can elicit suboptimal performance [48]. DEQ-MPI instead trains an iterative architecture to maximize image quality at convergence for improved performance, and it introduces a novel learned consistency block to better conform to the MPI data distribution. Initialization strategies are also proposed for both regularization and learned consistency blocks to improve model training. Demonstrations show that DEQ-MPI achieves superior performance to state-of-the-art methods for MPI reconstruction, while also maintaining superior or on par efficiency.

Contributions:

- We introduce the first physics-driven deep iterative architecture for performant and efficient MPI reconstruction.
- To our knowledge, DEQ-MPI is the first deep equilibrium model with learned consistency in the literature.
- We propose initialization strategies for regularization and learned consistency blocks in DEQ-MPI to improve model training.

4.2 Related Work

Learned image priors have recently been adopted as a promising approach in MPI tasks such as SM or image super-resolution [20, 55, 105], view imputation in projection imaging [106], and image reconstruction [27–29, 56, 73, 74, 104]. For image reconstruction, purely data-driven methods provide fast inference by directly mapping acquired data onto images without explicitly considering the SM [27–29, 104]. As these methods do not explicitly integrate physical constraints, reliability against system variability can be limited. Moreover, previous data-driven methods include dense layers whose complexity grows substantially with data dimensions. DIP methods instead use untrained networks with convolution filters serving as native regularizers, and learn network parameters to optimize DC on individual test scans [73, 74]. Although DIP methods promise enhanced generalization by incorporating the SM, they require thousands of inference iterations and face challenges in network selection as ideal architectures are often image specific [75]. PP-MPI pre-trains a convolutional network for image denoising, and combines it with the SM during an inference optimization [56]. While PP-MPI offers improved efficiency compared to DIP, transferring a model from the denoising to the reconstruction task can limit performance [107]. Thus, learning-based methods with improved efficiency and generalization are needed in MPI reconstruction.

A powerful framework for learning-based reconstruction employs physics-driven unrolled methods that perform a fixed number of iterated projections through a convolutional network block to regularize the image and a DC block to enforce the system’s physical constraints [43]. While no previous study has considered unrolled methods for MPI, state-of-the-art results have been reported with them in other modalities [43, 45]. That said, as computational complexity grows rapidly when more blocks are cascaded, unrolled methods are typically trained to optimize image quality after a small number of iterations. This limitation results in suboptimal performance, and image quality degrades significantly when inference is sought at a different number of iterations than that prescribed for training [51, 108]. Moreover, DC in MPI and other modalities is commonly

performed by projecting reconstructed data onto the ℓ_2 -ball of acquired data to alleviate bias due to noise [22, 23, 109]. This procedure ignores the underlying data distribution as it does not consider correlations among acquired data samples that can help lower such biases more effectively. In turn, the use of suboptimal DC measures can elicit performance losses during image reconstruction [48].

Our proposed DEQ-MPI model leverages two technical novelties to address the limitations of conventional unrolled methods in the context of MPI reconstruction. First, DEQ-MPI is not trained to optimize performance within a fixed number of iterations, but rather upon convergence as inspired by recent deep equilibrium models in machine learning [49]. Second, DEQ-MPI introduces a novel learned consistency block based on a convolutional module as opposed to hand-crafted measures. In the MPI domain, DEQ-MPI employs the first physics-driven iterative architecture, and the first learned consistency measure for image reconstruction. In the broader machine learning domain, DEQ-MPI is the first method that integrates a learned consistency measure into a deep equilibrium model.

4.3 Theory

4.3.1 MPI Reconstruction

Receive coils in MPI measure time-domain voltage waveforms that reflect the magnetization responses of MNPs. The acquired data can be transformed to frequency domain to define a linear system of equations [1]:

$$\mathbf{A}\mathbf{x} + \mathbf{n} = \mathbf{y} \quad (4.1)$$

where $\mathbf{A} \in \mathbb{C}^{M \times N}$ is the SM, $\mathbf{x} \in \mathbb{R}^N$ is the image vector, $\mathbf{n} \in \mathbb{C}^M$ is the noise vector, $\mathbf{y} \in \mathbb{C}^M$ are frequency-domain data, M is the number of frequency components, and N is the number of voxels in the imaging grid. Since noise in MPI is frequency dependent, $\mathbf{n}$ follows a multi-variate distribution with potential correlations [30, 66]. A common approach to solve Eq. (4.1) uses iterative

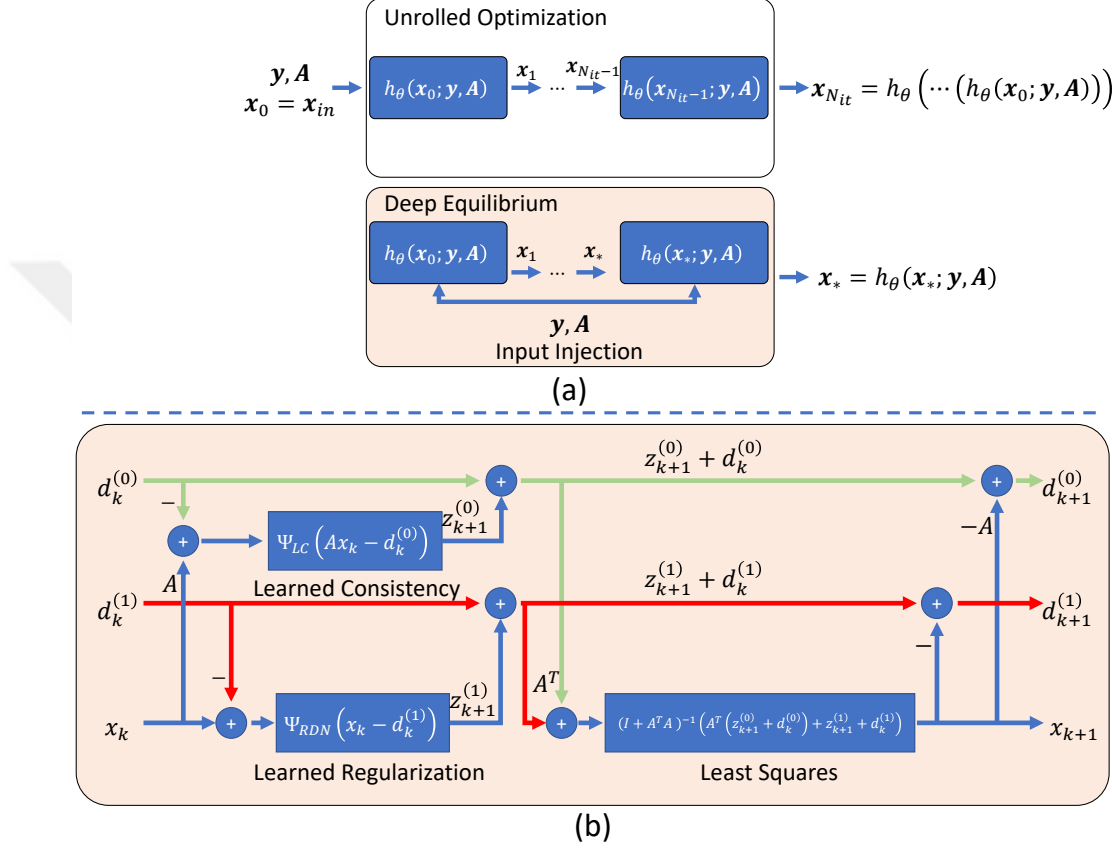


Figure 4.1: **(a)** Conventional unrolled methods versus the deep equilibrium model in DEQ-MPI. Unrolled methods express reconstruction as repeated projections through an explicit operator, $\mathbf{x}_{k+1} = h_\theta(\mathbf{x}_k; \mathbf{y}, \mathbf{A})$ where $\mathbf{x}_k$ is the image at iteration k , $\mathbf{y}$ are acquired data, $\mathbf{A}$ is the system matrix, and network parameters θ are shared across iterations. h_θ is trained to maximize performance after N_{it} iterations where N_{it} is fixed and small to limit computational burden, resulting in suboptimal performance. DEQ-MPI instead leverages an implicit mapping $\mathbf{x}_* = h_\theta(\mathbf{x}_*; \mathbf{y}, \mathbf{A})$ to compute a convergent output based on repeated injection of acquired data. In this case, h_θ is trained to maximize image quality upon convergence as opposed to an adhoc N_{it} . **(b)** Proposed DEQ-MPI implementation. DEQ-MPI integrates the implicit mapping into an ADMM algorithm with fixed-point iterations expressed in Eq. (4.14) for the image $\mathbf{x}$ and Lagrange multipliers $\mathbf{d}^{(0)}, \mathbf{d}^{(1)}$. Each iteration involves projection through a learned regularization block (Ψ_{RDN}), projection through a learned consistency block (Ψ_{LC}), and reconciliation in a least-squares step to compute the output $\mathbf{x}_{k+1}$.

optimization [15, 62]:

$$\arg \min_{\mathbf{x} \geq 0} R(\mathbf{x}) \text{ s.t. } \|\mathbf{Ax} - \mathbf{y}\|_2 < \epsilon, \quad (4.2)$$

where $R(\cdot)$ is the regularization operator, ϵ is the error bound for the ℓ_2 -based DC measure that can be selected based on the estimated signal-to-noise ratio (SNR) [15, 30]. Conventional methods adopt $R(\mathbf{x}) = \sum_i \alpha_i r_i(\mathbf{x})$, where $r_i(\mathbf{x})$ is a hand-crafted function such as $\|\mathbf{x}\|_2^2$, $\|\mathbf{x}\|_1$, or $TV(\mathbf{x})$ [24–26, 98].

An efficient algorithm is alternating direction method of multipliers (ADMM) that solves problems of type [15]:

$$\arg \min_{\mathbf{x}, \mathbf{z}} g(\mathbf{x}) + f(\mathbf{z}) \text{ s.t. } \mathbf{Hx} + \mathbf{Gz} = \mathbf{c}, \quad (4.3)$$

by splitting them into simpler sub-problems [69, 71]. To arrive at an ADMM formulation equivalent to Eq. (4.2), $\mathbf{H} = [\mathbf{A}^T, \mathbf{I}]^T \in \mathbb{R}^{(M+N) \times N}$, $\mathbf{G} = -\mathbf{I} \in \mathbb{R}^{(M+N) \times (M+N)}$, $\mathbf{c} = 0$ with $\mathbf{z} = [(\mathbf{z}^{(0)})^T, (\mathbf{z}^{(1)})^T]^T$ and $g(\mathbf{x}) = 0$ can be selected:

$$\arg \min_{\mathbf{x}, \mathbf{z}} f(\mathbf{z}) \text{ s.t. } \mathbf{x} = \mathbf{z}^{(1)}, \text{ and } \mathbf{Ax} = \mathbf{z}^{(0)}, \quad (4.4)$$

where $\mathbf{z} \in \mathbb{R}^{M+N}$ is the auxiliary variable vector used for splitting, and $f(\mathbf{z}) = \chi(\mathbf{z}^{(0)}) + R(\mathbf{z}^{(1)})$ where $\chi(\mathbf{t})$ is the indicator function of the set $\{\mathbf{t} \mid \|\mathbf{t} - \mathbf{y}\|_2 \leq \epsilon\}$ for the DC constraint. The following iterations are used to solve Eq. 4.4:

$$\mathbf{z}_{k+1}^{(0)} = \Psi_{\chi}(\mathbf{Ax}_k - \mathbf{d}_k^{(0)}, \mathbf{y}), \quad (4.5)$$

$$\mathbf{z}_{k+1}^{(1)} = \Psi_R(\mathbf{x}_k - \mathbf{d}_k^{(1)}), \quad (4.6)$$

$$\mathbf{x}_{k+1} = \mathbf{M}(\mathbf{A}^T(\mathbf{z}_{k+1}^{(0)} + \mathbf{d}_k^{(0)}) + \mathbf{z}_{k+1}^{(1)} + \mathbf{d}_k^{(1)}), \quad (4.7)$$

$$\mathbf{d}_{k+1}^{(0)} = \mathbf{d}_k^{(0)} + \mathbf{z}_{k+1}^{(0)} - \mathbf{Ax}_{k+1}, \quad (4.8)$$

$$\mathbf{d}_{k+1}^{(1)} = \mathbf{d}_k^{(1)} + \mathbf{z}_{k+1}^{(1)} - \mathbf{x}_{k+1}, \quad (4.9)$$

where k is the iteration index, $\mathbf{M} = (\mathbf{I} + \mathbf{A}^T \mathbf{A})^{-1}$ can be precomputed, $\mathbf{d} = [(\mathbf{d}^{(0)})^T, (\mathbf{d}^{(1)})^T]^T \in \mathbb{R}^{M+N}$ contains Lagrange multiplier terms for the constraints in Eq. 4.4. To incorporate the constraints flexibly, $\mathbf{d}^{(0)}$ captures data residuals due to deviation of $\mathbf{Ax}$ from $\mathbf{y}$ following DC, and $\mathbf{d}^{(1)}$ captures image residuals due to regularization of $\mathbf{x}$. The proximal mappings for DC and

regularization are given as:

$$\Psi_\chi(\mathbf{v}, \mathbf{y}) = \mathbf{y} + \begin{cases} \mathbf{v} - \mathbf{y} & \text{if } \|\mathbf{v} - \mathbf{y}\|_2 \leq \epsilon \\ \epsilon \frac{\mathbf{v} - \mathbf{y}}{\|\mathbf{v} - \mathbf{y}\|} & \text{o.w.} \end{cases}, \quad (4.10)$$

$$\Psi_R(\mathbf{v}) = \arg \min_{\mathbf{x}} R(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{x} - \mathbf{v}\|_2^2. \quad (4.11)$$

with μ scaled inversely with step size. Performance is limited by the capacity of Ψ_χ to describe MPI data distribution and the capacity of Ψ_R to describe MPI image features.

4.3.2 DEQ-MPI

Unrolled methods use iterated projections through a network operator, $\mathbf{x}_{k+1} = h_\theta(\mathbf{x}_k; \mathbf{y}, \mathbf{A})$, with parameters θ commonly shared across iterations [43]. After a fixed number of iterations N_{it} , h_θ is trained to optimize the quality of $\mathbf{x}_{N_{it}}$:

$$\arg \min_{\theta} \|h_\theta(\dots h_\theta(h_\theta(\mathbf{x}_0; \cdot); \cdot); \cdot) - \hat{\mathbf{x}}_r\|_1, \quad (4.12)$$

where $\hat{\mathbf{x}}_r$ is the ground truth image, and $\mathbf{x}_0$ is an initial reconstruction estimate provided to the network at $k=1$ (Fig. 4.1a). Using large N_{it} improves performance by yielding solutions closer to the convergence point. Yet, while forward passes can be computed efficiently, backpropagation requires computation and storage of model gradients across all iterations, rendering large N_{it} prohibitive [43]. Thus, a small N_{it} is typically used that yields non-convergent solutions of limited quality.

Unlike unrolled methods, DEQ-MPI leverages an implicit mapping $\mathbf{x}_* = h_\theta(\mathbf{x}_*; \mathbf{y}, \mathbf{A})$ based on a convergent solution $\mathbf{x}_*$, as inspired by recent deep equilibrium models in machine learning [49]. In theory, an infinite number of iterations through h_θ might be required to obtain $\mathbf{x}_*$. For efficiency, here we adopt an empirical convergence criterion to stop iterations when the relative change in $\mathbf{x}$ between consecutive iterations falls below a small non-zero threshold [49, 90]. Training is then performed to maximize the quality of $\mathbf{x}_*$ (Fig. 4.1a):

$$\arg \min_{\theta} \|h_\theta(\mathbf{x}_*; \mathbf{y}, \mathbf{A}) - \hat{\mathbf{x}}_r\|_1, \quad (4.13)$$

Because a convergent solution is attained, DEQ-MPI can perform efficient back-propagation via implicit differentiation, where gradients have to be computed only at the convergent iteration for $\mathbf{x}_*$ [50]. Since gradient terms for other iterations are not required, DEQ-MPI can improve performance without the computational overhead of unrolled methods.

Here we integrate the implicit mapping in DEQ-MPI into an ADMM algorithm with fixed-point iterations given as:

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{d}^{(0)} \\ \mathbf{d}^{(1)} \end{bmatrix}_{k+1} = h_\theta \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{d}^{(0)} \\ \mathbf{d}^{(1)} \end{bmatrix}_k ; \mathbf{y}, \mathbf{A} \right). \quad (4.14)$$

Mapping through $h_\theta(\cdot)$ is then operationalized as (Fig. 4.1b):

$$\mathbf{z}_{k+1}^{(0)} = \Psi_{LC}(\mathbf{A}\mathbf{x}_k - \mathbf{d}_k^{(0)}, \mathbf{y}), \quad (4.15)$$

$$\mathbf{z}_{k+1}^{(1)} = \Psi_{RDN}(\mathbf{x}_k - \mathbf{d}_k^{(1)}), \quad (4.16)$$

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{d}^{(0)} \\ \mathbf{d}^{(1)} \end{bmatrix}_{k+1} = \begin{bmatrix} \mathbf{M}(\mathbf{A}^T(\mathbf{z}_{k+1}^{(0)} + \mathbf{d}_k^{(0)}) + \mathbf{z}_{k+1}^{(1)} + \mathbf{d}_k^{(1)}) \\ \mathbf{d}_k^{(0)} + \mathbf{z}_{k+1}^{(0)} - \mathbf{A}\mathbf{x}_{k+1} \\ \mathbf{d}_k^{(1)} + \mathbf{z}_{k+1}^{(1)} - \mathbf{x}_{k+1} \end{bmatrix}. \quad (4.17)$$

A solution for convergent $[\mathbf{x}^T, \mathbf{d}^{(0),T}, \mathbf{d}^{(1),T}]_*^T$ is computed via fixed-point iterations accelerated with Anderson’s method for efficiency [49, 108, 110]. During these iterations, the proximal mapping $\Psi_{RDN}(\cdot)$ is implemented as projection through a residual dense network (RDN) block, where $\mathbf{d}^{(1)}$ captures image residuals after regularization. The proximal mapping $\Psi_{LC}(\cdot)$ is implemented as projection through a novel learned consistency (LC) block, where $\mathbf{d}^{(0)}$ captures data residuals after enforcement of consistency. As deep equilibrium methods can be sensitive to model initialization, we also introduce initialization strategies for both blocks. Details of model architecture and training procedures are discussed below.

4.3.2.1 Model Architecture

RDN block: $\Psi_{RDN}(\cdot)$ projects its input through a cascade of residual dense modules [83]. The input in 2D form $\mathbf{v} = \mathbf{x}_k - \mathbf{d}_k^{(1)} \in \mathbb{R}^{H \times W}$, where $\mathbf{d}_k^{(1)}$ captures

image residuals, and H , W are image height and width, passes through two convolutional layers $Z_0(\cdot)$:

$$\mathbf{u}_0 = Z_0(\mathbf{v}). \quad (4.18)$$

The feature map $\mathbf{u}_0 \in \mathbb{R}^{F_R \times H \times W}$, where F_R is the number of channels, is then processed with n_{res} residual modules:

$$\mathbf{u}_m = Z_m(\mathbf{u}_{m-1}), \quad (4.19)$$

where $Z_m(\cdot)$ is the m^{th} module with n_{conv} convolutional layers that receive concatenated outputs from previous layers:

$$\mathbf{u}_{m,l} = Z_{m,l}([\mathbf{u}_{m-1}; \mathbf{u}_{m,1}; \mathbf{u}_{m,2}; \cdots; \mathbf{u}_{m,l-1}]), \quad (4.20)$$

where $\mathbf{u}_{m,l} \in \mathbb{R}^{F_S \times H \times W}$ is the output of l^{th} convolutional layer, $Z_{m,l}$ with $1 \leq l \leq n_{conv}$. The output of m^{th} residual module $\mathbf{u}_m$ is then computed by adding the module input to the output of a final convolutional layer, $Z_{m,out}$:

$$\mathbf{u}_m = Z_{m,out}([\mathbf{u}_{m-1}; \mathbf{u}_{m,1}; \cdots; \mathbf{u}_{m,n_{conv}}]) + \mathbf{u}_{m-1}. \quad (4.21)$$

The outputs of all residual modules are fused via a 1×1 convolutional layer, Z_{fuse} :

$$\mathbf{u}_{fuse} = Z_{fuse}([\mathbf{u}_1; \mathbf{u}_2; \cdots; \mathbf{u}_{n_{res}}]), \quad (4.22)$$

where $\mathbf{u}_{fuse} \in \mathbb{R}^{F_R \times H \times W}$. The output image $\mathbf{z}_{k+1}^{(1)} \in \mathbb{R}^{H \times W}$ is computed by a convolutional layer, Z_{out} , with ReLU activation to integrate a non-negativity constraint for MPI:

$$\mathbf{z}_{k+1}^{(1)} = \text{ReLU}(Z_{out}(\mathbf{u}_{fuse}) + \mathbf{v}) \quad (4.23)$$

LC block: A common approach to implement $\Psi_\chi(\cdot)$ in Eq. (4.5) is to project onto the ℓ_2 -ball of acquired data $\mathbf{y}$. In contrast, DEQ-MPI leverages the LC block based on a convolutional module to better account for the MPI data distribution. LC receives a frequency-domain input $\mathbf{v} = \mathbf{A}\mathbf{x}_k - \mathbf{d}_k^{(0)} \in \mathbb{C}^M$, where $\mathbf{d}_k^{(0)}$ captures data residuals, along with $\mathbf{y}$:

$$\mathbf{z}_{k+1}^{(0)} = \Psi_{LC}(\mathbf{v}, \mathbf{y}), \quad (4.24)$$

where $\mathbf{z}_{k+1}^{(0)}$ are output data, and LC is implemented as:

$$\Psi_{LC}(\mathbf{v}, \mathbf{y}) = \mathbf{y} + \begin{cases} Z(\mathbf{v}, \mathbf{y}) - \mathbf{y} & \text{if } \|Z(\mathbf{v}, \mathbf{y}) - \mathbf{y}\|_2 \leq \epsilon \\ \epsilon \frac{Z(\mathbf{v}, \mathbf{y}) - \mathbf{y}}{\|Z(\mathbf{v}, \mathbf{y}) - \mathbf{y}\|} & \text{o.w.} \end{cases} \quad (4.25)$$

In Eq. (4.25), $Z(\cdot)$ is a convolutional module with n_{LC} hidden layers and F_{LC} hidden units per layer. The ϵ -bounded constraint prevents the output from diverging away from acquired data undesirably. Assuming field-free-line (FFL) scans with a single receive channel, data can be ordered in two dimensions as $\mathbf{y}(f, \phi)$, where f is the frequency component and ϕ is the FFL angle, and processed with 1D convolutional kernels across the frequency dimension. For field-free-point (FFP) scans with multiple receive channels, 2D kernels may instead be used over frequency and receive channel dimensions.

4.3.2.2 Training Procedures

Model initialization: Multiple convergent outputs $\mathbf{x}_*$ can exist for the implicit mapping in DEQ-MPI, and the quality of a particular solution depends on the initialization of model parameters. Here, we propose to initialize the RDN block based on a plug-and-play approach as inspired by [108]. To do this, independent identically distributed (IID) Gaussian noise $\mathbf{n}_1 \in \mathbb{R}^N$ with standard deviation σ_1 is added onto a training set of MPI images $\hat{\mathbf{x}}_r$, to generate images $\mathbf{x}_n = \hat{\mathbf{x}}_r + \mathbf{n}_1$. RDN is pre-trained to suppress the additive noise in $\mathbf{x}_n$:

$$\arg \min_{\theta_{RDN}} \|\Psi_{RDN}(\mathbf{x}_n) - \hat{\mathbf{x}}_r\|_1. \quad (4.26)$$

For the LC block, we propose a novel initialization procedure based on noise-added MPI data. First, noise-free data are generated using the SM and training MPI images, $\hat{\mathbf{y}}_r = \mathbf{A}\hat{\mathbf{x}}_r$. IID Gaussian noise is added at σ_2 and σ_3 to generate $\mathbf{y}_n = \hat{\mathbf{y}}_r + \mathbf{n}_2$ and $\mathbf{v}_n = \hat{\mathbf{y}}_r + \mathbf{n}_3$ with $\mathbf{n}_{2,3} \in \mathbb{C}^M$, respectively. LC is pre-trained to mimic a canonical unlearned DC block:

$$\arg \min_{\theta_{LC}} \|\Psi_{LC}(\mathbf{v}_n, \mathbf{y}_n) - \Psi_{\chi}(\mathbf{v}_n, \mathbf{y}_n)\|_1, \quad (4.27)$$

where Ψ_{χ} is implemented as in Eq. (4.10) based on projections onto the ℓ_2 -ball. We observed that pre-training to align the outputs of Ψ_{LC} and Ψ_{χ} improves

performance over pre-training to strictly align the output of Ψ_{LC} with $\hat{\mathbf{y}}_r$. While RDN and LC are initialized with the pre-trained weights for $\Psi_{RDN}(\cdot)$ and $\Psi_{LC}(\cdot)$, $\mathbf{x}$ is initialized with the least-squares solution $\mathbf{x}_{LS} = \mathbf{A}^\dagger \mathbf{y}$ based on the pseudo-inverse of the SM, and $\mathbf{d}^{(0)}, \mathbf{d}^{(1)}$ are initialized as zero vectors.

Implicit differentiation: In a forward pass, a convergent solution of $\mathbf{x}_* = h_\theta(\mathbf{x}_*; \mathbf{y}, \mathbf{A})$ is computed via fixed-point iterations in Eq. (4.14) accelerated using Anderson’s method [110]. An empirical convergence criterion is set as the ℓ_2 -norm difference between consecutive iterations falling below 10^{-4} . Here, a maximum of 25 iterations were observed to be sufficient for reaching convergence. In a backward pass based on Eq. (4.13), the Jacobian of the convergent solution $\partial \mathbf{x}_* / \partial \theta$ is computed by differentiating the implicit mapping:

$$\frac{\partial \mathbf{x}_*}{\partial \theta} = \frac{\partial h_\theta(\mathbf{x}_*)}{\partial \mathbf{x}_*} \frac{\partial \mathbf{x}_*}{\partial \theta} + \frac{\partial h_\theta(\mathbf{x}_*)}{\partial \theta}, \quad (4.28)$$

where the arguments $\mathbf{y}, \mathbf{A}$ are omitted for brevity. The following solution for the Jacobian $\partial \mathbf{x}_* / \partial \theta$ is then obtained:

$$\frac{\partial \mathbf{x}_*}{\partial \theta} = \left(\mathbf{I} - \frac{\partial h_\theta(\mathbf{x}_*)}{\partial \mathbf{x}_*} \right)^{-1} \frac{\partial h_\theta(\mathbf{x}_*)}{\partial \theta}. \quad (4.29)$$

Automatic differentiation tools for backpropagation require multiplication of the Jacobian with an arbitrary vector $\mathbf{b}$ [50]:

$$\left(\frac{\partial \mathbf{x}_*}{\partial \theta} \right)^T \mathbf{b} = \left(\frac{\partial h_\theta(\mathbf{x}_*)}{\partial \theta} \right)^T \left(\mathbf{I} - \frac{\partial h_\theta(\mathbf{x}_*)}{\partial \mathbf{x}_*} \right)^{-T} \mathbf{b}. \quad (4.30)$$

To solve Eq. (4.30), an intermediate vector can be defined as:

$$\mathbf{s}_* = (\mathbf{I} - \partial h_\theta(\mathbf{x}_*) / \partial \mathbf{x}_*)^{-T} \mathbf{b}, \quad (4.31)$$

where $\partial h_\theta(\mathbf{x}_*) / \partial \mathbf{x}_*$ can be computed trivially. Eq. (4.31) can be rearranged to compute $\mathbf{s}$ via fixed-point iterations [50]:

$$\mathbf{s}_{i+1} = (\partial h_\theta(\mathbf{x}_*) / \partial \mathbf{x}_*)^T \mathbf{s}_i + \mathbf{b}. \quad (4.32)$$

The expression in Eq. (4.30) can then be evaluated based on $\mathbf{s}_*$:

$$(\partial \mathbf{x}_* / \partial \theta)^T \mathbf{b} = (\partial h_\theta(\mathbf{x}_*) / \partial \theta)^T \mathbf{s}_*. \quad (4.33)$$

As such, the implicit mapping enables calculation of the Jacobian $\partial \mathbf{x}_* / \partial \theta$ in terms of $\partial h_\theta(\mathbf{x}_*) / \partial \mathbf{x}_*$ and $\partial h_\theta(\mathbf{x}_*) / \partial \theta$.

4.4 Methods

4.4.1 Competing Methods

DEQ-MPI was demonstrated against state-of-the-art methods based on hand-crafted and learned priors. Model hyperparameters were selected based on peak SNR (pSNR) performance on a validation set. For each method, the number of inference iterations was selected according to the L-curve criterion to achieve a favorable trade-off between performance and computation time [90]. Methods were implemented in PyTorch on a Tesla V100 GPU. Except DIP, learning-based models were trained for 200 epochs via the ADAM optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$). Code will be available at <https://github.com/icon-lab/DEQ-MPI> upon publication.

DEQ-MPI: Architectural parameters were taken as $n_{res} = 4$, $F_R = 12$, $n_{conv} = 12$ for the RDN block, $n_{LC} = 1$, $F_{LC} = 8$ for the LC block. Cross-validated parameters included a learning rate of 10^{-3} , 25 inference iterations. $\epsilon = \sqrt{M}$ for data consistency, and $\sigma_1 = 0.1$, $\sigma_2 = 0.05$, $\sigma_3 = 0.02$ for model initialization were used.

ADMM: Three ADMM-based methods were implemented with ℓ_1 , TV, and a hybrid ℓ_1 +TV regularizer in the image domain, respectively [15]. For each variant and SNR, μ was selected to ensure convergence. $\mu = 250$ for the ℓ_1 , $\mu = 50$ for the TV, and $\mu = 10$ for the hybrid variant were used. For the hybrid variant, $\alpha_{TV} = 1 - \alpha_1$ was prescribed, and cross-validated values were $\alpha_1 = 0.1$ for $SNR < 20$, $\alpha_1 = 0.8$ for $20 \leq SNR < 30$, $\alpha_1 = 0.9$ for $30 \leq SNR$. Cross-validated number of iterations were 200 for the ℓ_1 , and 100 for the TV and hybrid variants. $\epsilon = \sqrt{M}$ was used.

ART: Algebraic reconstruction technique (ART), i.e. the Kaczmarz method, with a Tikhonov regularizer was implemented [24]. Cross-validated parameters included 10 inference iterations, and a regularization weight of $\lambda = 10$ for $SNR < 15$, $\lambda = 1$ for $15 \leq SNR < 35$, $\lambda = 0.1$ for $35 \leq SNR$.

DIP: The DIP method based on an untrained network was implemented as described in [73], albeit a mean-square error loss was adopted as it was observed to yield higher performance. Inference was performed with the ADAM optimizer. Cross-validated parameters were 10^{-3} learning rate and 20000 inference iterations.

PP-MPI: The PP-MPI method based on a denoising prior was implemented [56]. The network architecture and loss function were adopted from [56]. Cross-validated parameters were 10^{-3} learning rate, additive noise with a standard deviation of 0.1, and 150 inference iterations.

4.4.2 MPI Phantoms

Vessel phantoms are commonly used in demonstrating reconstruction methods in MPI. Here, we generated simulated vessel phantoms based on time-of-flight magnetic resonance angiograms (MRA), following [56]. MRA images from 95 healthy subjects in the ITKTubeTK dataset were used [111]. Data were split into non-overlapping training, validation, test sets of 77, 9, 9 subjects, respectively. $10 \times 26 \times 52$ volumetric patches were randomly cropped, followed by a maximum-intensity projection along the first dimension, and two-fold downsampling in other dimensions to obtain 13×26 images. The maximum pixel intensity in each image was randomly scaled to a number between 0.5 and 1.5. A total of 33692 training, 3377 validation and 3730 test images were obtained.

For the experiments, two different phantoms were used (Fig. 4.2b,c). The first phantom included two parallel cylindrical tubes filled with Perimag (Micromod GmbH, Germany) MNPs at a dilution ratio of 1:20. Each tube had 20-mm length, 2-mm inner radius, 4-mm outer radius, and the tubes were separated by 4 mm. The second phantom included a 3D-printed ‘Y-shape’ filled with Perimag MNPs at a dilution ratio of 1:100, and contained a central air bubble. Two arms of 10.7-mm length with 3.5-mm spacing at one end, and one arm of 8.8-mm length formed the Y-shape. All arms were 3.5-mm wide.

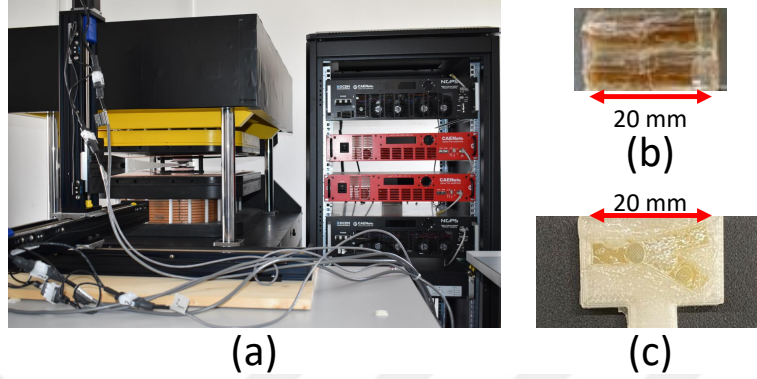


Figure 4.2: (a) Open-sided ASELSAN FFL scanner. (b) A cylindrical phantom with two parallel tubes was filled with Perimag MNPs at a dilution of 1:20. (c) A Y-shape phantom was filled with Perimag at a dilution of 1:100 and contained a central air bubble.

4.4.3 Experimental Procedures

Experimental SM and phantom measurements were performed on the open-sided ASELSAN FFL scanner (Fig. 4.2a) [30]. For the SM acquisition, an undiluted Perimag sample of size $2 \times 2 \times 2 \text{ mm}^3$ was scanned with 2 mm steps while the FFL was rotated in the transverse plane over a $26 \times 52 \text{ mm}^2$ FOV. A drive field of 9 mT amplitude and 10 ms duration per angle was applied at a 10% duty cycle. An SR-560 pre-amplifier (SRS, MA, USA) amplified the signal at a gain of 5, filtered it at a frequency cut-off of 10-300 kHz, and the signal was then sampled at 5 MS/s. Frequency components around the 2nd-to-11th harmonics were selected over 500 Hz bandwidths. Whitening and background subtraction were performed based on background measurements. High-SNR rows were selected ($\text{SNR} > 5$). Two separate experimental sessions were conducted. In the first session, SM and Y-shape phantom measurements were performed at a selection field (SF) gradient strength of 0.5 T/m. In the second session that was held two months later, SM and cylindrical phantom measurements were performed separately at SF gradient strengths of 0.5 and 0.6 T/m.

4.4.4 Simulation and Analysis Procedures

For experimental phantoms, no ground-truth images exist to quantify reconstruction performance. Thus, to perform quantitative assessments in a setup that respects system non-idealities, we emulated MPI data by coupling simulated phantoms with experimental SMs. MPI measurements reflect a superposition integral between the continuously-varying system function and MNP distribution. Assuming that $\mathbf{A}_c$ and $\mathbf{x}_c$ denote finely discretized SM and MNP distribution that closely approximate the underlying continuous variables:

$$\mathbf{y} = \mathbf{A}_c \mathbf{x}_c + \mathbf{n}. \quad (4.34)$$

We do not have access to $\mathbf{A}_c$, but instead calibration scans capture SM on a discretized grid for a relatively large MNP sample size, $\mathbf{A}_{meas} = \mathbf{A}_c \mathbf{D}$ where $\mathbf{D} \in \mathbb{R}^{s^2 N \times N}$ is the box-downsampling matrix by a factor of s in two-dimensions [55]. Given $\mathbf{A}_{meas}$, one can approximate the system function by bicubic up-sampling, $\tilde{\mathbf{A}}_c = \mathbf{A}_{meas} \mathbf{U}$ where $\mathbf{U} \in \mathbb{R}^{N \times s^2 N}$. However, during reconstruction $\mathbf{x}_{rec} = \mathbf{m}(\mathbf{y}, \tilde{\mathbf{A}}_c)$, this elicits a discrepancy between the underlying SM that gives rise to the MPI data versus the estimated SM input to the reconstruction:

$$\mathbf{x}_{rec} = \mathbf{m}(\mathbf{A}_c \mathbf{x}_c + \mathbf{n}, \mathbf{A}_c \mathbf{D} \mathbf{U}). \quad (4.35)$$

Ignoring this discrepancy can lead to an inverse crime for simulation studies involving image reconstruction [112]. To avoid this problem, we generated MPI data by multiplying the simulated phantom with the SM measured during a calibration scan, i.e., $\hat{\mathbf{y}}_r = \mathbf{A}_{meas} \hat{\mathbf{x}}_r$. We then reconstructed the noise-added data $\mathbf{y} = \hat{\mathbf{y}}_r + \mathbf{n}$ assuming a modified SM, $\mathbf{A}_{meas} \mathbf{U} \mathbf{D}$, to mimic the discrepancy highlighted in Eq. (4.35).

For training DEQ-MPI, simulated phantoms from the training and validation sets were coupled with the SM measured at 0.5 T/m SF gradient from the second experimental session. Simulated phantoms from the test set were coupled with the measured SMs from both experimental sessions to quantify performance. The DEQ-MPI model trained on simulated phantoms was also tested on experimental phantoms.

To describe the noise level in the generated MPI data, the measurement SNR was computed as:

$$SNR(\mathbf{y}) = 20 \log_{10} (\|\hat{\mathbf{y}}_r\|_2 / \|\mathbf{y} - \hat{\mathbf{y}}_r\|_2), \quad (4.36)$$

where $\hat{\mathbf{y}}_r$ and $\mathbf{y}$ are noise-free and noisy data. To assess reconstruction performance for simulated phantoms, pSNR and structural similarity (SSIM) were computed:

$$pSNR(\mathbf{x}) = 20 \log_{10} \left(\sqrt{N} \|\hat{\mathbf{x}}_r\|_{\infty} / \|\mathbf{x} - \hat{\mathbf{x}}_r\|_2 \right), \quad (4.37)$$

$$SSIM(\mathbf{x}) = \frac{(2\mu_x\mu_{\hat{x}_r} + c_1)(2\sigma_{x\hat{x}_r} + c_2)}{(\mu_x^2 + \mu_{\hat{x}_r}^2 + c_1)(\sigma_x^2 + \sigma_{\hat{x}_r}^2 + c_2)}, \quad (4.38)$$

where $\mathbf{x}$ is the reconstructed image, $\hat{\mathbf{x}}_r$ is the ground truth, μ , σ^2 denote image mean and variance, $\sigma_{x\hat{x}_r}$ is the covariance of $\mathbf{x}$ and $\hat{\mathbf{x}}_r$, $c_{1,2}$ are scalars that prevent division by zero [113].

4.5 Results

4.5.1 Ablation Studies

We conducted a set of ablation studies to assess the value of the individual design elements in DEQ-MPI. To assess the value of physics-driven learning, an end-to-end model was built where an RDN block without DC was trained to directly map the least-squares solution onto ground-truth images. To assess the value of deep equilibrium modeling, an unrolled model with conventional unlearned DC block was built with $N_{it} = 5$ iterations (selected via cross validation). To assess the value of learned consistency, an LC-ablated variant of DEQ-MPI was built with a conventional unlearned DC block. Using simulated phantoms, training was performed at measurement SNRs of 5-45 dB, and testing was performed under 35 dB SNR. For brevity, assessments based on pSNR are reported, although the same conclusions are valid based on SSIM. DEQ-MPI generally outperforms all ablated models across training SNRs (Fig. 4.3). The only exception is at

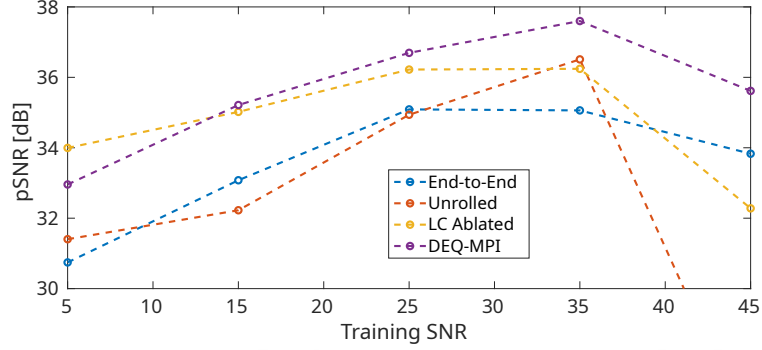


Figure 4.3: Comparison of DEQ-MPI against an end-to-end model that used an RDN block and omitted DC, an unrolled model with $N_{it}=5$ iterations, and an LC-ablated variant based on an unlearned DC block. Separate models were trained at measurement SNRs of 5-45 dB, testing was performed under 35 dB SNR. Average pSNR across the test set is shown for each model.

SNR=5 dB where the LC-ablated variant yields a moderately higher pSNR, best attributed to the relatively low test SNR levels that mismatch the training SNR limiting the performance of the LC block.

We also examined the convergence behaviors of the iterative models, when the training and test SNRs were both 35 dB (Fig. 4.4). The non-iterative end-to-end model was not considered. The unrolled model begins to suffer dramatically when the number of iterations exceeds N_{it} assumed during training, and the LC-ablated variant has relatively slow convergence to a suboptimal performance level. In contrast, DEQ-MPI shows fast convergence where it exceeds the performance of the unrolled model beyond $N_{it} = 5$.

Next, we assessed the importance of the proposed initializations for the RDN and LC blocks in DEQ-MPI. Accordingly, a variant based on a randomly initialized RDN, a variant based on a randomly initialized LC, and a variant with randomly initialized RDN and LC were built. To demonstrate the proposed LC initialization, an additional variant with LC pre-trained to estimate noise-free data was also built. The training and test SNRs matched (35 dB). We find that the average pSNR is 37.6 dB for DEQ-MPI, 29.9 dB when RDN is randomly initialized, 20.2 dB when LC is randomly initialized, 20.1 dB when both RDN and LC are randomly initialized, and 29.9 dB with LC pre-trained to estimate

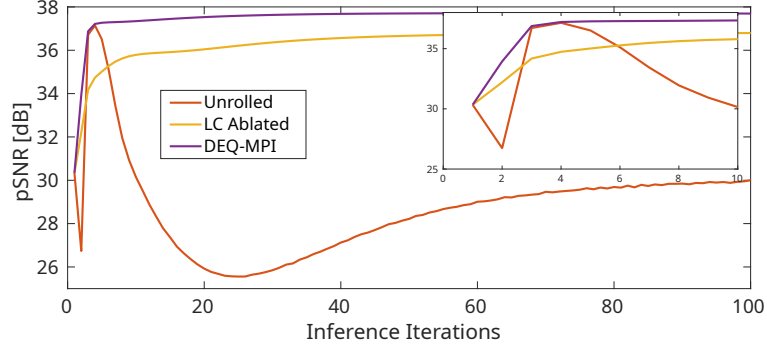


Figure 4.4: Convergence behaviors of DEQ-MPI, unrolled model, and LC-ablated variant. Models were trained and tested at 35 dB SNR. Average pSNR across the test set is shown for each model. The zoom-in window highlights performance during initial iterations.

noise-free data. These results indicate that the proposed model initializations contribute substantially to reconstruction performance.

4.5.2 Simulated Phantoms

DEQ-MPI was demonstrated against traditional (ADMM_{ℓ_1} , ADMM_{TV} , ADMM_{hyb} , ART) and learning-based methods (DIP, PP-MPI) via quantitative assessments on simulated phantoms (Sec. 4.4.4). pSNR and SSIM were computed across the test set for variable measurement SNRs, with training and test sets having matching SNR for each case for DEQ-MPI. When the same SM was used for both training and testing (the SM at 0.5 T/m SF gradient from the second session), DEQ-MPI outperforms the top-contending traditional method by 2.8 dB pSNR / 6.8% SSIM, and the top-contending learning-based method by 1.7 dB pSNR / 4.5% SSIM (Table 4.1). When the SM differed across training-test sets (the SMs at 0.5 versus 0.6 T/m SF gradient), DEQ-MPI again outperforms the top-contending traditional method by 2.5 dB pSNR / 6.2% SSIM, and the top-contending learning-based method by 1.3 dB pSNR / 3.7% SSIM (Table 4.2). For each competing method, performance levels are comparable across Tables 4.1-4.2 because both cases utilized the same underlying phantoms in the test set and the same SNR levels. Note, however, that this does not imply that the SMs

Table 4.1: Simulated phantoms were reconstructed at varying measurement SNRs. An SM at 0.5 T/m SF gradient was used in training and test sets. pSNR (dB) / SSIM (%) are reported as mean $\pm$ std. across the test set. Boldface marks the top-performing method.

	SNR=15 dB	SNR=25 dB	SNR=35 dB
ADMM $_{\ell_1}$	18.9 $\pm$ 3.5 / 16.4 $\pm$ 9.5	23.2 $\pm$ 2.2 / 32.6 $\pm$ 14.0	28.9 $\pm$ 1.3 / 58.6 $\pm$ 13.8
ADMM $_{TV}$	27.3 $\pm$ 6.1 / 56.2 $\pm$ 20.0	31.2 $\pm$ 4.4 / 71.9 $\pm$ 11.3	35.5 $\pm$ 3.0 / 84.5 $\pm$ 6.5
ADMM $_{hyb}$	28.7 $\pm$ 5.1 / 63.9 $\pm$ 9.5	31.7 $\pm$ 3.6 / 74.8 $\pm$ 7.5	35.9 $\pm$ 2.7 / 85.6 $\pm$ 5.6
ART	26.7 $\pm$ 4.3 / 49.6 $\pm$ 7.8	28.5 $\pm$ 5.7 / 64.3 $\pm$ 8.8	31.9 $\pm$ 4.3 / 76.1 $\pm$ 4.4
DIP	16.7 $\pm$ 2.6 / 13.5 $\pm$ 9.1	22.6 $\pm$ 1.7 / 30.9 $\pm$ 16.6	28.2 $\pm$ 2.9 / 54.6 $\pm$ 21.0
PP-MPI	30.6 $\pm$ 5.1 / 70.8 $\pm$ 11.5	33.4 $\pm$ 4.1 / 77.8 $\pm$ 9.6	35.7 $\pm$ 3.1 / 82.8 $\pm$ 8.2
DEQ-MPI	32.1$\pm$4.5 / 75.2$\pm$9.9	34.8$\pm$3.5 / 81.5$\pm$8.3	37.7$\pm$2.6 / 88.1$\pm$5.7

Table 4.2: Simulated phantoms were reconstructed at varying measurement SNRs. The SM at 0.5 T/m was used in training, and the SM at 0.6 T/M was used in testing. pSNR (dB) / SSIM (%) are reported as mean $\pm$ std. across the test set.

	SNR=15 dB	SNR=25 dB	SNR=35 dB
ADMM $_{\ell_1}$	19.1 $\pm$ 3.5 / 17.1 $\pm$ 9.5	23.3 $\pm$ 2.2 / 33.4 $\pm$ 14.1	29.1 $\pm$ 1.3 / 59.3 $\pm$ 13.6
ADMM $_{TV}$	27.5 $\pm$ 6.1 / 57.4 $\pm$ 19.5	31.3 $\pm$ 4.4 / 72.8 $\pm$ 10.8	35.7 $\pm$ 3.0 / 84.9 $\pm$ 6.3
ADMM $_{hyb}$	28.7 $\pm$ 5.0 / 64.4 $\pm$ 9.4	31.8 $\pm$ 3.6 / 75.3 $\pm$ 7.3	36.0 $\pm$ 2.7 / 86.0 $\pm$ 5.4
ART	26.7 $\pm$ 4.3 / 49.8 $\pm$ 8.0	28.5 $\pm$ 5.8 / 64.6 $\pm$ 8.8	32.0 $\pm$ 4.3 / 76.5 $\pm$ 4.4
DIP	16.9 $\pm$ 2.6 / 13.9 $\pm$ 9.0	22.5 $\pm$ 1.8 / 31.0 $\pm$ 16.7	28.0 $\pm$ 3.1 / 54.1 $\pm$ 21.8
PP-MPI	30.8 $\pm$ 5.0 / 71.3 $\pm$ 11.4	33.5 $\pm$ 4.1 / 78.4 $\pm$ 9.2	35.8 $\pm$ 3.1 / 83.4 $\pm$ 7.8
DEQ-MPI	31.9$\pm$4.5 / 75.0$\pm$9.8	34.5$\pm$3.4 / 81.3$\pm$8.1	37.6$\pm$2.6 / 88.0$\pm$5.7

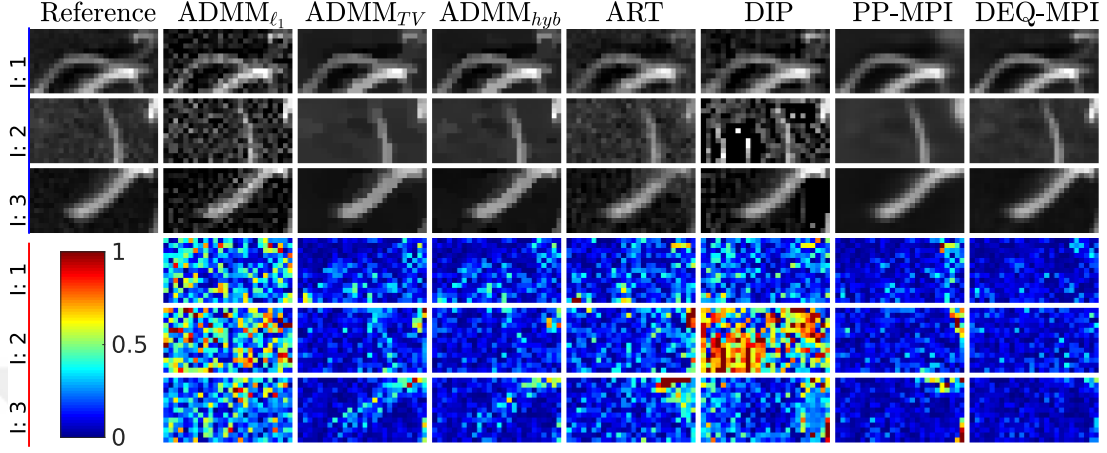


Figure 4.5: Representative reconstructions of three simulated phantoms and respective error maps (see colorbar) are shown for competing methods, along with the reference images. A measurement SNR of 35 dB was used. The SM at 0.5 T/m SF gradient was used for training, and the SM at 0.6 T/m was used for testing.

at 0.5 and 0.6 T/m are interchangeable, since reconstructing the data measured with the SM at 0.5 T/m using the SM at 0.6 T/m would result in substantial performance loss (e.g. 15.5 dB pSNR / 7.3% SSIM loss for ADMM_{hyb}).

Representative reconstructions and the respective error maps from competing methods under 35 dB SNR are displayed in Fig. 4.5 for the case with different SMs across the training-test sets. ADMM_{ℓ_1} yields a grainy image with residual noise and background signal; ADMM_{TV} , ADMM_{hyb} , and ART suffer from spatial blurring; and DIP can suffer from noise amplification. While PP-MPI yields relatively higher performance, it shows elevated errors near the upper and lower right corners where the experimental SM has limited sensitivity. In contrast, DEQ-MPI yields superior performance with lower reconstruction errors than competing methods. Reconstructions for varying measurement SNRs for the case with different SMs across the training-test sets are shown in Fig. 4.6. As expected, performance improves for all methods as measurement SNR increases. Among competing methods, ADMM_{ℓ_1} and particularly DIP that are relatively amenable to noise amplification show poor performance towards lower SNR levels. Overall, DEQ-MPI produces high image quality with lower artifacts and noise than competing methods.

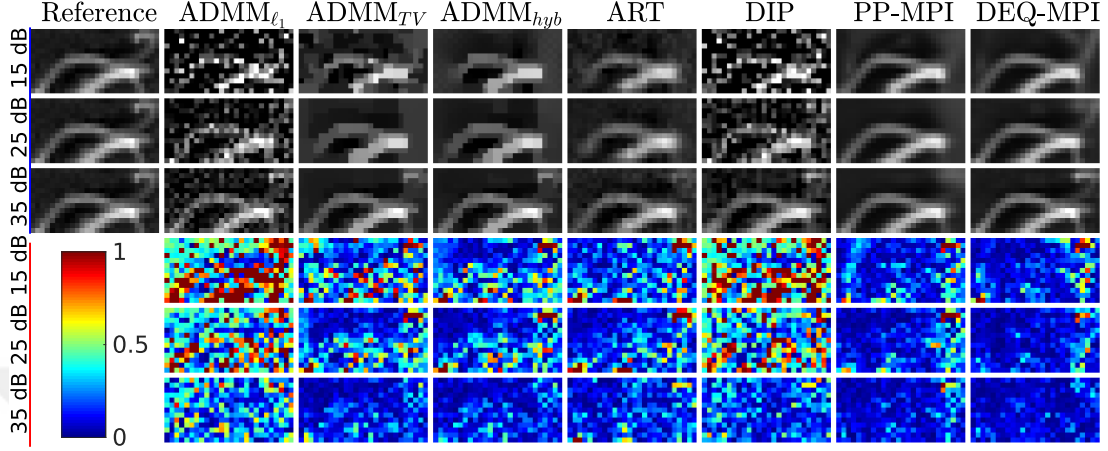


Figure 4.6: Representative reconstructions of a simulated phantom at SNR=15-35 dB and respective error maps are shown for competing methods, along with the reference image. The SM at 0.5 T/m was used for training, and the SM at 0.6 T/m was used for testing.

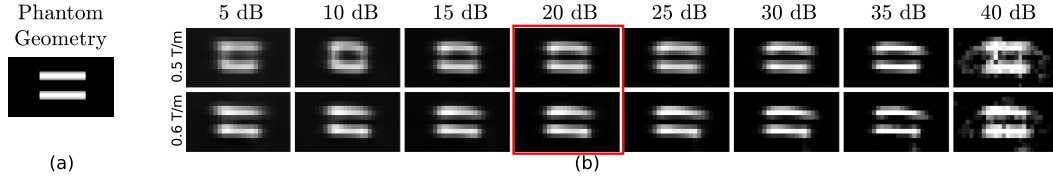


Figure 4.7: Reconstructions of experimental cylindrical phantom with DEQ-MPI trained at SNR levels in 5-40 dB. (a) Approximate geometry of the phantom. (b) Reconstructions at (top-row) 0.5 T/m and (bottom-row) 0.6 T/m SF gradients. The SM at 0.5 T/m was used for training DEQ-MPI. Testing utilized the SM corresponding to each case. Red box denotes the case where the training and testing SNRs match.

4.5.3 Experimental Phantoms

Next, DEQ-MPI was demonstrated on two experimental phantoms. Due to lack of ground-truth images in experimental settings, assessments were performed qualitatively via visual inspection [26, 30]. DEQ-MPI was trained using emulated MPI data from simulated phantoms at an assumed SNR level, whereas testing was performed on experimental MPI data at $\text{SNR} \approx 20$ dB (estimated based on multiple signal and background measurements). To assess reliability against SNR and SM mismatches between the training-test sets, reconstructions of measurements at

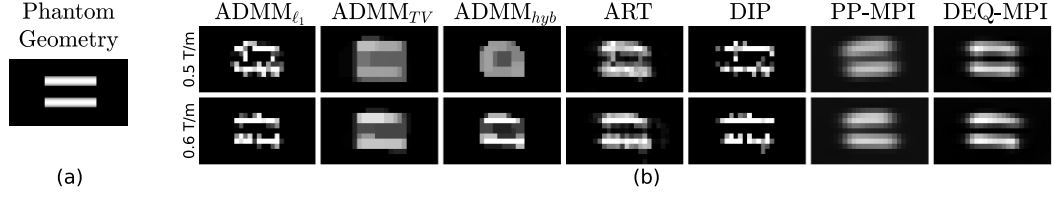


Figure 4.8: Reconstructions of experimental cylindrical phantom with competing methods. **(a)** Approximate geometry of the phantom. **(b)** Reconstructed images at **(top-row)** 0.5 T/m and **(bottom-row)** 0.6 T/m SF gradients. The SM at 0.5 T/m was used for training DEQ-MPI. Testing utilized the SM corresponding to each case.

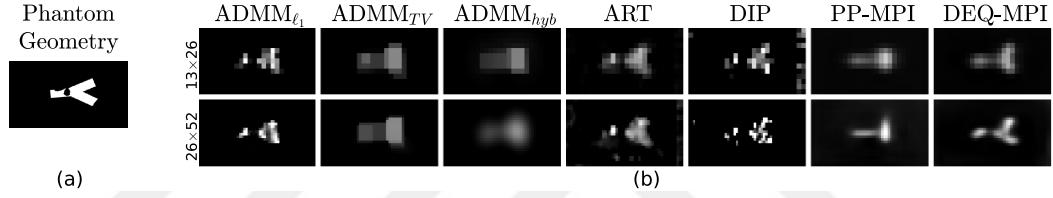


Figure 4.9: Reconstructions of experimental Y-shape phantom with competing methods. **(a)** Approximate geometry of the phantom. **(b)** Reconstructed images based on **(top-row)** the original 13×26 SM and **(bottom-row)** the upsampled 26×52 SM. The SM at 0.5 T/m from the second session was used for training DEQ-MPI. The SM at 0.5 T/m from the first session was used for testing.

0.5 and 0.6 T/m SF gradients were obtained separately using models trained for SNRs in 5-40 dB (Fig. 4.7). Training was performed using the SM at 0.5 T/m, whereas testing was performed using the SM corresponding to each SF gradient. In general, DEQ-MPI shows reliability against moderate deviations between the training and test SNRs, albeit residual reconstruction errors occur when the difference between the two SNRs reaches 20 dB. DEQ-MPI also demonstrates robustness against deviations between the training and test SMs.

Reconstructions of the cylindrical phantom were then compared for the competing methods at 0.5 and 0.6 T/m (Fig. 4.8). Again, training for DEQ-MPI was performed using the SM at 0.5 T/m, whereas testing was performed using the SM corresponding to each case. The other methods utilized the SM corresponding to each case, as well. $ADMM_{\ell_1}$ and DIP yield over-sparsified images with artefactual bright/dark pixels due to amplified noise; $ADMM_{TV}$ and $ADMM_{hyb}$ yield

over-smoothed images with block artifacts; and ART shows blurring and residual noise. While PP-MPI mostly avoids these issues, it reconstructs cylindrical tubes at an incorrect geometric orientation compared to remaining methods. In contrast, DEQ-MPI yields lower artifacts/noise and higher resemblance to the designed phantom than competing methods.

Images of the Y-shape phantom were also reconstructed (Fig. 4.9). The SM at 0.5 T/m from the second experimental session was used for training DEQ-MPI, whereas testing was performed using the SM at 0.5 T/m from the first session. A recalibration was performed on the FFL system between the two sessions that were 2 months apart, so the resultant SMs differed. The measured SM (sampled at 2-mm/pixel) was used to reconstruct images at the original 13×26 resolution, and bicubic interpolated version of the SM (upsampled to 1-mm/pixel) was used to reconstruct images at 26×52 resolution [30, 78]. ADMM_{ℓ_1} and DIP suffer from artefactual pixels due to noise amplification; ADMM_{TV} , ADMM_{hyb} , and ART suffer from spatial blurring; and PP-MPI does not faithfully capture the geometry of the phantom including the central air bubble. In comparison, DEQ-MPI offers high quality reconstructions in both the original and upsampled resolutions.

4.5.4 Reconstruction Time

The number of inference iterations and reconstruction times for all competing methods are listed in Table 4.3. Reconstruction performance as a function of run time is plotted in Fig. 4.10. Among competing methods, ART and particularly DIP require prolonged inference, and ADMM variants with TV regularization (ADMM_{TV} , ADMM_{hyb}) and PP-MPI have moderate run times. In comparison, DEQ-MPI yields efficient reconstructions with relatively fast convergence and run times competitive with ADMM_{ℓ_1} .

Table 4.3: Inference iterations (number of iterations) and reconstruction times (milliseconds) for a single 13×26 MPI image.

	ADMM_{ℓ_1}	ADMM_{TV}	ADMM_{hyb}	ART	DIP	PP-MPI	DEQ-MPI
Iters.	200	100	100	10	20000	150	25
Time	55	221	236	6966	235162	379	64

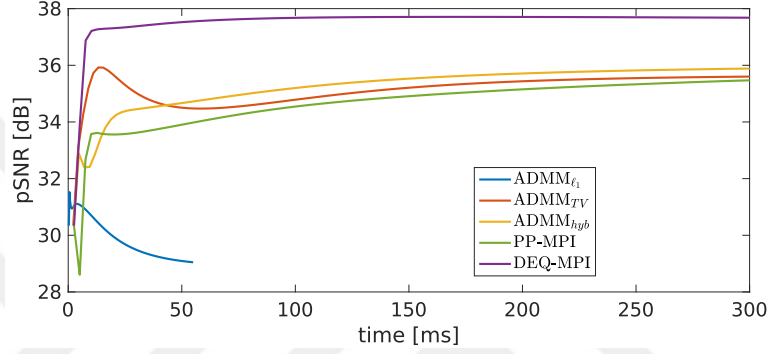


Figure 4.10: Performance of competing methods as a function of inference time. ART and DIP that have markedly prolonged run times are omitted. Average pSNR across the test set is shown for each method.

4.6 Discussion

DEQ-MPI integrates an implicit mapping into an optimization algorithm for performance and efficiency in MPI reconstruction. The implicit mapping is based on learned regularization and DC blocks to better conform to the data distribution, and accelerated fixed-point iterations are used to rapidly compute a convergent solution. Demonstrations on simulated and experimental phantoms indicate that DEQ-MPI trained using a single acquired SM outperforms previous traditional and learning-based methods. As a physics-driven method, DEQ-MPI shows reliability against deviations in the SM and in SNR levels between the training and test sets. While reconstruction errors occur when the training SNR is dramatically higher than the test SNR, this scenario can be avoided by performing a rough SNR estimation on given data.

Traditional methods can show high sensitivity to weights for hand-crafted regularizers [45]. MPI studies have reported that ideal weights can vary drastically

across scans, suggesting that parameter tuning on each test image might be useful [26, 30]. Such optimization is infeasible in pre-clinical or clinical scenarios as no a priori knowledge would be available on the MNP distribution. To address this challenge, here we optimized model hyperparameters on a validation set, and the selected values were used thereafter in the test set. Learning-based methods were observed to be more forgiving against suboptimal parameters (results not shown), so they might alleviate the need for exhaustive parameter tuning.

While performant reconstructions have been reported based on untrained networks in MPI literature [73, 74], here we observed relatively limited performance with DIP. Note that DIP directly minimizes a conventional DC loss between recovered and acquired test data. This loss function intrinsically assumes that data contain negligible noise compared to the signal. Because this assumption is violated for moderate to low SNR levels as considered in the current study, DIP can perform suboptimally in relatively limited SNR regimes.

Few recent studies have considered deep equilibrium models for undersampled MRI reconstruction [51, 108], with similar aims to the current study in terms of improving performance over unrolled methods. Yet, our proposed approach is unique in the following aspects: (1) Instead of integrating an implicit mapping into a projection-onto-convex-sets algorithm as in [51], DEQ-MPI leverages an ADMM algorithm that can offer improved reliability for non-convex or non-smooth problems. (2) Both [51] and [108] employ a conventional unlearned DC block. In contrast, DEQ-MPI leverages a learned consistency (LC) block to better conform to the data distribution. (3) While [51] does not report non-standard initialization and [108] only considers initialization for the regularization block, DEQ-MPI employs dedicated initialization methods for its regularization and LC blocks that improve model performance. (4) Lastly, [51] and [108] both use cascaded convolutional architectures with residual learning at the output layer. Instead, DEQ-MPI adopts multiple residual connections densely distributed across convolutional layers that have been reported to offer performance benefits [83]. It remains an important future work to comparatively evaluate the efficacy of various network architectures in MPI reconstruction.

Several developments can be considered to improve DEQ-MPI. First, we generated training data using MRA images under the assumption that they have similar features to MPI images. When the imaged anatomy is non-vascular, this approach might yield suboptimal performance. While public datasets of MPI images are rare, DEQ-MPI can in principle be enhanced by training the model on large amounts of experimental data to better capture application-specific image features. Second, MPI data include a non-stationary background that was separately measured and subtracted from acquired data prior to reconstruction. The need for background measurements can be avoided by extending DEQ-MPI to separately reconstruct the foreground and background signals. To do this, a dictionary-based approach can be adopted to estimate the background signal from acquired data [93]. Third, DEQ-MPI was trained based convolutional networks and a pixel-wise loss term. Performance improvements might be achieved with attention-based architectures to capture contextual features [91], and diffusion processes to increase reliability in model training [114]. Fourth, DEQ-MPI was demonstrated for reconstructing experimental phantoms at $2\times$ higher spatial resolution by bicubic SM upsampling. For higher upsampling factors, learning-based super-resolution methods can be adopted [20,55] and higher-resolution phantoms can be used for model training. Finally, imaging over large FOVs can be attained by performing patch-wise reconstructions with DEQ-MPI and fusing the multi-patch outputs [60].

4.7 Conclusion

Here, we introduced a novel deep equilibrium reconstruction for MPI with learned consistency. For improved performance and reliability, DEQ-MPI follows a physics-driven approach that integrates network blocks that regularize the image and enforce DC into an iterative algorithm. Simulated and experimental demonstrations indicate clear performance benefits and competitive efficiency over both traditional and recent learning-based methods. Thus, DEQ-MPI holds great promise for fast, high-fidelity image reconstruction in MPI.

Chapter 5

Conclusion

In this thesis, we have presented novel signal processing methods for magnetic particle imaging. To accelerate the calibration scan, we proposed a deep learning-based super-resolution of MPI system matrices. To improve image reconstruction quality, we proposed a deep equilibrium-based approach with learned consistency. We provided comprehensive evaluations on both simulated and experimental data for each case and showed that the proposed techniques outperform the current state of the art. Specific discussions and future directions for each study can be found below.

5.1 Discussions and Future Directions on Accelerated MPI System Matrix Calibration

We first described a deep learning-based super-resolution technique for accelerated MPI SM calibration. The technique included a vision transformer, convolutional and data consistency modules for improved and generalizable recovery of SMs. We have shown that it outperforms both compressed sensing and super-resolution-based techniques in terms of both SM recovery and image reconstruction performance. We believe that its performance may be improved in several

ways.

While previous compressed sensing approaches rely on sparsifying transformations and restricted isometry property [115] which require non-uniform undersampling, super-resolution techniques rely on the ability of the recovery technique to express the underlying high-resolution SM from uniformly undersampled data. Since they rely on different principles, the techniques may be combined to improve recovery performance. One such way may be to simply skip some of the pixels in the LR SM and consecutively first recover the LR SM through compressed sensing, followed by TranSMS super-resolution of the SM. Another possible direction may be to use structured MNP samples such that each collected sample is a hybrid combination of uniform and non-uniform undersampled data. In such a scenario, the data consistency module would need to be updated to reflect the shape of the structured MNP sample.

Moreover, recently proposed extrapolation techniques also show promising solutions [116]. Although the extrapolation provides modest acceleration (less than 2-fold) compared to very high acceleration factors of both compressed sensing (up to 27-fold for 3D imaging) and super-resolution (up to 64-fold for 2D imaging), these two techniques may also be combined such that SM calibration can be further accelerated.

5.2 Discussions and Future Directions on Deep Learning-based Image Reconstruction for MPI

Next, we described a deep learning enabled image reconstruction technique for MPI. The technique is inspired by the fixed point convergence training of the optimization algorithms through implicit differentiation. Furthermore, it included learned regularization and learned consistency blocks to better model both image

and data for improved reconstruction accuracy. We have shown that it outperforms conventional approaches, as well as other widely adopted techniques for computational imaging such as plug-and-play and unrolled techniques in terms of image reconstruction performance and computational time in both simulated and experimental phantoms. We believe that its performance may be improved in several ways.

First, since widely adopted training data for MPI is not available, we used extracted patches from magnetic resonance angiography images with the assumption that they have similar features to MPI. However, for non-vascular imaging this assumption may not be true and better datasets may be employed to improve the performance of DEQ-MPI [56]. Furthermore, techniques that learn without ground-truth may also help improve performance, since they can use experimental data directly for training [114, 117, 118].

Next, there is non-stationary background present in MPI data. While we simply subtracted background by gathering an empty measurement, other techniques that include background estimation during image reconstruction report improved performance [93]. While the joint estimation of background and the image requires additional selection of hyper-parameters and further validation of the technique on a wide dataset, it may also be included as a data-driven reconstruction within DEQ-MPI.

Next, while here we consider simpler architectures due to modest image sizes in MPI, attention-based architectures such as vision transformers may better capture the image statistics in MPI. Moreover, adversarial loss terms may better account for the ambiguity in the data and improve image reconstruction.

Finally, there are recent approaches in computational imaging that adopts Bayesian or probabilistic estimation of images that maps each data to multiple probable images [107, 114, 119–123]. These techniques may also provide uncertainty maps in MPI images, which can provide critical information given that the data may have very low SNR or very high background that may include estimation errors.

Bibliography

- [1] B. Gleich and J. Weizenecker, “Tomographic imaging using the nonlinear response of magnetic particles,” *Nature*, vol. 435, no. 7046, pp. 1214–1217, 2005.
- [2] E. U. Saritas, P. W. Goodwill, L. R. Croft, J. J. Konkle, K. Lu, B. Zheng, and S. M. Conolly, “Magnetic particle imaging (mpi) for nmr and mri researchers,” *J. Magn. Reson.*, vol. 229, pp. 116–126, 2013.
- [3] J. Weizenecker, B. Gleich, J. Rahmer, H. Dahnke, and J. Borgert, “Three dimensional real time in vivo magnetic particle imaging,” *Phys. Med. Biol.*, vol. 54, no. 5, pp. L1–L10, 2009.
- [4] B. Zheng, T. Vazin, P. W. Goodwill, A. Conway, A. Verma, E. U. Saritas, D. Schaffer, and S. M. Conolly, “Magnetic particle imaging tracks the long-term fate of in vivo neural cell implants with high image contrast,” *Sci. Rep.*, vol. 5, no. 1, pp. 1–9, 2015.
- [5] H. Arami, E. Teeman, A. Troksa, H. Bradshaw, K. Saatchi, A. Tomitaka, S. S. Gambhir, U. O. Häfeli, D. Liggitt, and K. M. Krishnan, “Tomographic magnetic particle imaging of cancer targeted nanoparticles,” *Nanoscale*, vol. 9, pp. 18723–18730, 2017.
- [6] G. Song, M. Chen, Y. Zhang, L. Cui, H. Qu, X. Zheng, M. Wintermark, Z. Liu, and J. Rao, “Janus iron oxides @ semiconducting polymer nanoparticle tracer for cell tracking by magnetic particle imaging,” *Nano Letters*, vol. 18, no. 1, pp. 182–189, 2018.

- [7] M. Utkur, Y. Muslu, and E. U. Saritas, “Relaxation-based viscosity mapping for magnetic particle imaging,” *Phys. Med. Biol.*, vol. 62, no. 9, pp. 3422–3439, 2017.
- [8] P. Ludewig, N. Gdaniec, J. Sedlacik, N. D. Forkert, P. Szwargulski, M. Graeser, G. Adam, M. G. Kaul, K. M. Krishnan, R. M. Ferguson, A. P. Khandhar, P. Walczak, J. Fiehler, G. Thomalla, C. Gerloff, T. Knopp, and T. Magnus, “Magnetic particle imaging for real-time perfusion imaging in acute stroke,” *ACS Nano*, vol. 11, no. 10, pp. 10480–10488, 2017.
- [9] C. Z. Cooley, J. B. Mandeville, E. E. Mason, E. T. Mandeville, and L. L. Wald, “Rodent cerebral blood volume changes during hypercapnia observed using magnetic particle imaging detection,” *NeuroImage*, vol. 178, pp. 713–720, 2018.
- [10] Z. W. Tay, P. Chandrasekharan, A. Chiu-Lam, D. W. Hensley, R. Dhavalikar, X. Y. Zhou, E. Y. Yu, P. W. Goodwill, B. Zheng, C. Rinaldi, and S. M. Conolly, “Magnetic particle imaging-guided heating in vivo using gradient fields for arbitrary localization of magnetic hyperthermia therapy,” *ACS Nano*, vol. 12, no. 4, pp. 3699–3713, 2018.
- [11] W. Tong, H. Hui, W. Shang, Y. Zhang, F. Tian, Q. Ma, X. Yang, J. Tian, and Y. Chen, “Highly sensitive magnetic particle imaging of vulnerable atherosclerotic plaque with active myeloperoxidase-targeted nanoparticles,” *Theranostics*, vol. 11, pp. 506–521, 2021.
- [12] P. W. Goodwill, E. U. Saritas, L. R. Croft, T. N. Kim, K. M. Krishnan, D. V. Schaffer, and S. M. Conolly, “X-space mpi: Magnetic nanoparticles for safe medical imaging,” *Adv. Mat.*, vol. 24, no. 28, pp. 3870–3877, 2012.
- [13] P. W. Goodwill and S. M. Conolly, “Multidimensional x-space magnetic particle imaging,” *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1581–1590, 2011.

- [14] A. A. Ozaslan, A. Alacaoglu, O. B. Demirel, T. Çukur, and E. U. Saritas, “Fully automated gridding reconstruction for non-cartesian x-space magnetic particle imaging,” *Physics in Medicine & Biology*, vol. 64, p. 165018, aug 2019.
- [15] S. Ilbey, C. B. Top, A. Güngör, T. Çukur, E. U. Saritas, and H. E. Güven, “Comparison of system-matrix-based and projection-based reconstructions for field free line magnetic particle imaging,” *Int. J. Magn. Part. Imaging*, vol. 3, no. 1, 2017.
- [16] M. Grüttner, T. Knopp, J. Franke, M. Heidenreich, J. Rahmer, A. Halkola, C. Kaethner, J. Borgert, and T. M. Buzug, “On the formulation of the image reconstruction problem in magnetic particle imaging,” *Biomed. Tech.*, vol. 58, no. 6, pp. 583–591, 2013.
- [17] A. Von Gladiss, M. Gräser, P. Szwargulski, T. Knopp, and T. M. Buzug, “Hybrid system calibration for multidimensional magnetic particle imaging,” *Phys. Med. Biol.*, vol. 62, no. 9, p. 3392, 2017.
- [18] J. Lampe, C. Bassoy, J. Rahmer, J. Weizenecker, H. Voss, B. Gleich, and J. Borgert, “Fast reconstruction in magnetic particle imaging,” *Phys. Med. Biol.*, vol. 57, no. 4, pp. 1113–1134, 2012.
- [19] A. Weber and T. Knopp, “Reconstruction of the magnetic particle imaging system matrix using symmetries and compressed sensing,” *Adv. Math. Phys.*, pp. 1687–9120, 2015.
- [20] I. M. Baltruschat, P. Szwargulski, F. Griesse, M. Grosser, R. Werner, and T. Knopp, “3d-smrnet: Achieving a new quality of mpi system matrix recovery by deep learning,” in *Med. Image Comput. Comput. Assist. Interv.*, pp. 74–82, 2020.
- [21] X. Chen, Z. Jiang, X. Han, X. Wang, and X. Tang, “The reconstruction of magnetic particle imaging: Current approaches based on the system matrix,” *Diagnostics*, vol. 11, no. 5, 2021.

- [22] T. Knopp, S. Biederer, T. F. Sattel, J. Rahmer, J. Weizenecker, B. Gleich, J. Borgert, and T. M. Buzug, “2D model-based reconstruction for magnetic particle imaging,” *Med. Phys.*, vol. 37, no. 2, pp. 485–491, 2010.
- [23] T. Kluth and B. Jin, “L1 data fitting for robust reconstruction in magnetic particle imaging: quantitative evaluation on Open MPI dataset,” *Int. J. Magn. Part. Imaging*, vol. 6, no. 2, pp. 1–23, 2020.
- [24] T. Knopp, J. Rahmer, T. F. Sattel, S. Biederer, J. Weizenecker, B. Gleich, J. Borgert, and T. M. Buzug, “Weighted iterative reconstruction for magnetic particle imaging,” *Phys. Med. Biol.*, vol. 55, no. 6, pp. 1577–1589, 2010.
- [25] J. J. Konkle, P. W. Goodwill, D. W. Hensley, R. D. Orendorff, M. Lustig, and S. M. Conolly, “A convex formulation for magnetic particle imaging X-space reconstruction,” *PLoS ONE*, vol. 10, no. 10, 2015.
- [26] M. Storath, C. Brandt, M. Hofmann, T. Knopp, J. Salamon, A. Weber, and A. Weinmann, “Edge Preserving and Noise Reducing Reconstruction for Magnetic Particle Imaging,” *IEEE Trans. Med. Imaging*, vol. 36, no. 1, pp. 74–85, 2017.
- [27] B. G. Chae, “Neural network image reconstruction for magnetic particle imaging,” *arXiv:1709.07560*, 2017.
- [28] P. Koch, M. Maass, M. Bruhns, C. Droigk, T. J. Parbs, and A. Mertins, “Neural network for reconstruction of MPI images,” in *Int. Workshop Magn. Part. Imaging*, pp. 39–40, 2019.
- [29] T. Hatsuda, S. Shimizu, H. Tsuchiya, T. Takagi, T. Noguchi, and Y. Ishihara, “A basic study of an image reconstruction method using neural networks for magnetic particle imaging,” in *Int. Workshop Magn. Part. Imaging*, 2015.
- [30] C. B. Top and A. Güngör, “Tomographic field free line magnetic particle imaging with an open-sided scanner configuration,” *IEEE Trans. Med. Imaging*, vol. 39, no. 12, pp. 4164–4173, 2020.

- [31] R. Ahmad, C. A. Bouman, G. T. Buzzard, S. Chan, S. Liu, E. T. Reehorst, and P. Schniter, “Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery,” *IEEE Sig. Process. Mag.*, vol. 37, no. 1, pp. 105–116, 2020.
- [32] S. U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, and T. Çukur, “Image synthesis in multi-contrast mri with conditional generative adversarial networks,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 10, pp. 2375–2388, 2019.
- [33] B. Gözcü, R. K. Mahabadi, Y.-H. Li, E. Ilıcak, T. Çukur, J. Scarlett, and V. Cevher, “Learning-based compressive mri,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1394–1406, 2018.
- [34] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems,” *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011.
- [35] M. Yurt, M. Özbey, S. U. Dar, B. Tinaz, K. K. Oguz, and T. Çukur, “Progressively volumetrized deep generative models for data-efficient contextual learning of MR image recovery,” *Medical Image Analysis*, vol. 78, p. 102429, 2022.
- [36] M. Yurt, S. U. Dar, A. Erdem, E. Erdem, K. K. Oguz, and T. Çukur, “mustGAN: multi-stream generative adversarial networks for MR image synthesis,” *Medical Image Analysis*, vol. 70, p. 101944, 2021.
- [37] O. Dalmaz, U. Mirza, G. Elmas, M. Özbey, S. U. Dar, E. Ceyani, S. Avestimehr, and T. Çukur, “One model to unite them all: Personalized federated learning of multi-contrast MRI synthesis,” *arXiv:2207.06509*, 2022.
- [38] M. B. Alver, A. Saleem, and M. Cetin, “Plug-and-play synthetic aperture radar image formation using deep priors,” *IEEE Transactions on Computational Imaging*, vol. 7, pp. 43–57, 2021.

- [39] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Now Foundations and Trends, 2011.
- [40] U. S. Kamilov, H. Mansour, and B. Wohlberg, “A plug-and-play priors approach for solving nonlinear imaging inverse problems,” *IEEE Sig. Process. Let.*, vol. 24, no. 12, pp. 1872–1876, 2017.
- [41] O. F. Kar, A. Gungor, and H. E. Guven, “Learning based regularization for spatial multiplexing cameras,” in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1–5, 2019.
- [42] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, “Plug-and-play priors for model based reconstruction,” in *IEEE Glob. Conf. Sig. Inf. Process.*, 2013.
- [43] H. K. Aggarwal, M. P. Mani, and M. Jacob, “Modl: Model-based deep learning architecture for inverse problems,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 394–405, 2019.
- [44] V. Monga, Y. Li, and Y. C. Eldar, “Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing,” *IEEE Sig. Process. Mag.*, vol. 38, no. 2, pp. 18–44, 2021.
- [45] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, “A transfer-learning approach for accelerated MRI using deep neural networks,” *Magn. Reson. Med.*, vol. 84, no. 2, pp. 663–685, 2020.
- [46] S. Shoushtari, J. Liu, Y. Hu, and U. S. Kamilov, “Deep model-based architectures for inverse problems under mismatched priors,” *IEEE Journal on Selected Areas in Information Theory*, pp. 1–1, 2022.
- [47] Y. Sun, B. Wohlberg, and U. S. Kamilov, “An online plug-and-play algorithm for regularized image reconstruction,” *IEEE Transactions on Computational Imaging*, vol. 5, no. 3, pp. 395–408, 2019.

- [48] J. Cheng, Z. Cui, W. Huang, Z. Ke, L. Ying, H. Wang, Y. Zhu, and D. Liang, “Learning data consistency and its application to dynamic mr imaging,” *IEEE Trans. Med. Imag.*, pp. 1–1, 2021.
- [49] S. Bai, J. Z. Kolter, and V. Koltun, “Deep equilibrium models,” in *Adv. Neural Inf. Process. Syst.*, vol. 32, pp. 1–12, 2019.
- [50] Z. Kolter, D. Duvenaud, and M. Johnson, “Deep implicit layers - neural odes, deep equilibrium models, and beyond.”
- [51] Z.-X. Cui, J. Cheng, Q. Zhu, Y. Liu, S. Jia, K. Zhao, Z. Ke, W. Huang, H. Wang, Y. Zhu, and D. Liang, “Equilibrated zeroth-order unrolled deep networks for accelerated mri,” *arXiv:2112.09891*, 2021.
- [52] J. Hu, S. Shoushtari, Z. Zou, J. Liu, Z. Sun, and U. S. Kamilov, “Robustness of deep equilibrium architectures to changes in the measurement model,” *arXiv:2211.00531*, 2022.
- [53] J. Weizenecker, B. Gleich, and J. Borgert, “Magnetic particle imaging using a field free line,” *Journal of Physics D: Applied Physics*, vol. 41, p. 105009, may 2008.
- [54] T. Knopp, P. Szwargulski, F. Griesse, and M. Gräser, “OpenMPIData: an initiative for freely accessible magnetic particle imaging data,” *Data in Brief*, vol. 28, p. 104971, 2020.
- [55] A. Güngör, B. Askin, D. A. Soydan, E. U. Saritas, C. B. Top, and T. Çukur, “TranSMS: Transformers for super-resolution calibration in magnetic particle imaging,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3562–3574, 2022.
- [56] B. Askin, A. Güngör, D. Alptekin Soydan, E. U. Saritas, C. B. Top, and T. Cukur, “PP-MPI: A deep plug-and-play prior for magnetic particle imaging reconstruction,” in *Mach. Learn. Med. Image Recon.*, pp. 105–114, 2022.
- [57] A. Güngör, B. Askin, D. A. Soydan, C. B. Top, E. U. Saritas, and T. Çukur, “DEQ-MPI: A deep equilibrium reconstruction with learned consistency for magnetic particle imaging,” *arXiv:2212.13233*, 2022.

- [58] B. Kilic, S. D. Alptekin, A. Gungor, and C. B. Top, “Inverse radon transform-based reconstruction with an open-sided magnetic particle imaging prototype,” *Signal Image Video Process.*, pp. 1–8, 2022.
- [59] S. Kaczmarz, “Angenaeherte aufloesung von systemen linearer gleichungen,” in *Bull. Int. Acad. Pol. Sci. Lett. A.*, vol. 35, pp. 355–357, 1937.
- [60] M. Boberg, T. Knopp, P. Szwargulski, and M. Möddel, “Generalized mpi multi-patch reconstruction using clusters of similar system matrices,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1347–1358, 2020.
- [61] T. Knopp and T. M. Buzug, *Magnetic particle imaging: An introduction to imaging principles and scanner instrumentation*. Springer, 2012.
- [62] T. Kluth and B. Jin, “Enhanced reconstruction in magnetic particle imaging by whitening and randomized SVD approximation,” *Phys. Med. Biol.*, vol. 64, no. 12, 2019.
- [63] S. Ilbey, C. B. Top, A. Güngör, T. Çukur, E. U. Saritas, and H. E. Güven, “Fast system calibration with coded calibration scenes for magnetic particle imaging,” *IEEE Trans. Med. Imag.*, vol. 38, no. 9, pp. 2070–2080, 2019.
- [64] T. Knopp and A. Weber, “Sparse reconstruction of the magnetic particle imaging system matrix,” *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1473–1480, 2013.
- [65] T. Kluth, C. Bathke, M. Jiang, and P. Maass, “Joint super-resolution image reconstruction and parameter identification in imaging operator: analysis of bilinear operator equations, numerical solution, and application to magnetic particle imaging,” *Inverse Probl.*, vol. 36, no. 12, p. 124006, 2020.
- [66] H. Paysen, O. Kosch, J. Wells, N. Loewa, and F. Wiekhorst, “Characterization of noise and background signals in a magnetic particle imaging system,” *Phys. Med. Biol.*, vol. 65, no. 23, p. 235031, 2020.
- [67] S. U. Dar, M. Yurt, M. Shahdloo, M. E. Ildiz, B. Tinaz, and T. Cukur, “Prior-guided image reconstruction for accelerated multi-contrast MRI via

- generative adversarial networks,” *IEEE J. Sel. Topics Sig. Process.*, vol. 14, no. 6, pp. 1072–1087, 2020.
- [68] L. K. Senel, T. Kilic, A. Gungor, E. Kopanoglu, H. E. Guven, E. U. Saritas, A. Koc, and T. Çukur, “Statistically segregated k-space sampling for accelerating multiple-acquisition mri,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 7, pp. 1701–1714, 2019.
- [69] E. Kopanoglu, A. Güngör, T. Kilic, E. U. Saritas, K. K. Oguz, T. Çukur, and H. E. Güven, “Simultaneous use of individual and joint regularization terms in compressive sensing: Joint reconstruction of multi-channel multi-contrast mri acquisitions,” *NMR in Biomed.*, vol. 33, no. 4, p. e4247, 2020.
- [70] H. E. Güven, A. Güngör, and M. Çetin, “An augmented lagrangian method for complex-valued compressed sar imaging,” *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 235–250, 2016.
- [71] A. Güngör, M. Çetin, and H. E. Güven, “Compressive synthetic aperture radar imaging and autofocusing by augmented lagrangian methods,” *IEEE Trans. Comput. Imaging*, vol. 8, pp. 273–285, 2022.
- [72] D. Ulyanov, A. Vedaldi, and V. Lempitsky, “Deep image prior,” in *IEEE Comput. Vis. Pattern. Recognit.*, 2018.
- [73] S. Dittmer, T. Kluth, D. O. Baguer, and P. Maass, “A deep prior approach to magnetic particle imaging,” in *Mach. Learn. Med. Imag. Recon.*, pp. 113–122, 2020.
- [74] T. Knopp and M. Grosser, “Warmstart approach for accelerating deep image prior reconstruction in dynamic tomography,” in *Medical Imaging with Deep Learning*, 2022.
- [75] M. E. Arican, O. Kara, G. Bredell, and E. Konukoglu, “ISNAS-DIP: image-specific neural architecture search for deep image prior,” in *IEEE Conf. Comput. Vis. Pat. Recognit.*, pp. 1960–1968, 2022.
- [76] C. Kaethner, W. Erb, M. Ahlborg, P. Szwargulski, T. Knopp, and T. M. Buzug, “Non-equispaced system matrix acquisition for magnetic particle

- imaging based on lissajous node points,” *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2476–2485, 2016.
- [77] O. A. Omer, H. Wojtczyk, and T. M. Buzug, “Simultaneous reconstruction and resolution enhancement for magnetic particle imaging,” *IEEE Trans. Magn.*, vol. 51, no. 2, pp. 1–4, 2015.
- [78] A. Gungor and C. B. Top, “Super-resolving reconstruction technique for mpi,” *Int. J. Magn. Part. Imaging*, vol. 6, no. 2, 2020.
- [79] A. Güngör, B. Askin, D. A. Soydan, C. B. Top, and T. Çukur, “Deep learned super resolution of system matrices for magnetic particle imaging,” in *IEEE. Eng. Med. Biol. Soc.*, pp. 1–4, 2021.
- [80] M. Grosser, M. Möddel, and T. Knopp, “Using low-rank tensors for the recovery of mpi system matrices,” *IEEE Trans. Comput. Imaging*, vol. 6, pp. 1389–1402, 2020.
- [81] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.
- [82] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” in *IEEE Comput. Vis. Pattern. Recognit.*, pp. 1646–1654, 2016.
- [83] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual dense network for image super-resolution,” in *IEEE Comput. Vis. Pattern. Recognit.*, pp. 2472–2481, 2018.
- [84] Y. Li, B. Sixou, and F. Peyrin, “A review of the deep learning methods for medical images super resolution problems,” *IRBM*, vol. 42, no. 2, pp. 120–133, 2021.
- [85] H. Wang and K. Jiang, “Research on image super-resolution reconstruction based on transformer,” in *Int. Conf. Artif. Intell. Ind. Design*, pp. 226–230, 2021.

- [86] X. Gu, F. Zhou, R. Chen, X. Ren, and W. Zhou, “Endoscopic single image super-resolution based on transformer and convolutional neural network,” in *Int. Conf. Intell. Comput. Sust. Energy Environ.*, Nov 2021.
- [87] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” *arXiv:2010.11929*, 2021.
- [88] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, “TransUNet: Transformers make strong encoders for medical image segmentation,” *arXiv:2102.04306*, 2021.
- [89] D. Karimi, S. Vasylechko, and A. Gholipour, “Convolution-free medical image segmentation using transformers,” *arXiv:2102.13645*, 2021.
- [90] Y. Korkmaz, S. U. Dar, M. Yurt, M. Özbey, and T. Çukur, “Unsupervised MRI reconstruction via zero-shot learned adversarial transformers,” *IEEE Trans. Med. Imaging*, vol. 41, no. 7, pp. 1747–1763, 2022.
- [91] O. Dalmaz, M. Yurt, and T. Çukur, “ResViT: Residual vision transformers for multi-modal medical image synthesis,” *IEEE Trans Med Imaging*, vol. 41, no. 10, pp. 2598–2614, 2022.
- [92] K. Them, M. G. Kaul, C. Jung, M. Hofmann, T. Mummert, F. Werner, and T. Knopp, “Sensitivity enhancement in magnetic particle imaging by background subtraction,” *IEEE Trans. Med. Imag.*, vol. 35, no. 3, pp. 893–900, 2016.
- [93] T. Knopp, M. Grosser, M. Graeser, T. Gerkmann, and M. Möddel, “Efficient joint estimation of tracer distribution and background signals in magnetic particle imaging using a dictionary approach,” *IEEE Trans. Med. Imag.*, pp. 1–1, 2021.
- [94] H. Wu, B. Xiao, N. Codella, M. Liu, X. Dai, L. Yuan, and L. Zhang, “CvT: Introducing convolutions to vision transformers,” *arXiv:2103.15808*, 2021.

- [95] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *IEEE Comput. Soc. Conf. Comput. Vis. Pattern. Recognit.*, pp. 1800–1807, 2017.
- [96] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *IEEE Comput. Vis. Pattern. Recognit.*, pp. 1874–1883, 2016.
- [97] S. M. A. Bashir, Y. Wang, M. Khan, and Y. Niu, “A comprehensive review of deep learning-based single image super-resolution,” *PeerJ Comp. Sci.*, vol. 7, p. e621, 2021.
- [98] F. Lieb and T. Knopp, “A wavelet-based sparse row-action method for image reconstruction in magnetic particle imaging,” *Med. Phys.*, vol. 48, no. 7, pp. 3893–3903, 2021.
- [99] A. Halkola, J. Rahmer, B. Gleich, J. Borgert, and T. Buzug, “System calibration unit for magnetic particle imaging: System matrix,” in *Int. Wkshp. Magn. Part. Imaging*, pp. 1–1, 2013.
- [100] N. Löwa, P. Radon, O. Kosch, and F. Wiekhorst, “Concentration dependent mpi tracer performance,” *Int. J. Magn. Part. Imaging*, vol. 2, no. 1, 2016.
- [101] T. Takagi, S. Shimizu, H. Tsuchiya, T. Hatsuda, T. Noguchi, and Y. Ishihara, “Image reconstruction method based on orthonormal basis of observation signal by singular value decomposition for magnetic particle imaging,” in *Int. Workshop Magn. Part. Imaging*, 2015.
- [102] L. Schmiester, M. Möddel, W. Erb, and T. Knopp, “Direct image reconstruction of lissajous-type magnetic particle imaging data using chebyshev-based matrix compression,” *IEEE Trans. Comput. Imaging*, vol. 3, no. 4, pp. 671–681, 2017.
- [103] S. Dittmer, T. Kluth, M. T. R. Henriksen, and P. Maass, “Deep image prior for 3d magnetic particle imaging: A quantitative comparison of regularization techniques on open mpi dataset,” in *Int. Workshop on Magn. Part. Imag.*, 2020.

- [104] A. von Gladiss, R. Memmesheimer, N. Theisen, A. C. Bakenecker, T. M. Buzug, and D. Paulus, “Reconstruction of 1d images with a neural network for magnetic particle imaging,” in *BV Med.*, pp. 247–252, 2022.
- [105] Y. Shang, J. Liu, L. Zhang, X. Wu, P. Zhang, L. Yin, H. Hui, and J. Tian, “Deep learning for improving the spatial resolution of magnetic particle imaging,” *Phys. Med. Biol.*, vol. 67, no. 12, p. 125012, 2022.
- [106] X. Wu, B. He, P. Gao, P. Zhang, Y. Shang, L. Zhang, J. Zhong, J. Jiang, H. Hui, and J. Tian, “PGNet: Projection generative network for sparse-view reconstruction of projection-based magnetic particle imaging,” *Med. Phys.*, 2022.
- [107] A. Güngör, S. U. Dar, S. Öztürk, Y. Korkmaz, G. Elmas, M. Özbey, and T. Çukur, “Adaptive diffusion priors for accelerated mri reconstruction,” *arXiv:2207.05876*, 2022.
- [108] D. Gilton, G. Ongie, and R. Willett, “Deep equilibrium architectures for inverse problems in imaging,” *IEEE Trans. Comput. Imaging*, vol. 7, pp. 1123–1133, 2021.
- [109] J. P. Haldar and J. Zhuo, “P-LORAKS: Low-Rank Modeling of Local k-Space Neighborhoods with Parallel Imaging Data,” *Magn. Reson. Med.*, vol. 75, no. 4, p. 1499, 2016.
- [110] H. F. Walker and P. Ni, “Anderson acceleration for fixed-point iterations,” *SIAM J. Num. Anal.*, vol. 49, no. 4, pp. 1715–1735, 2011.
- [111] “TubeTK - Magnetic Resonance Angiography Dataset.”
- [112] C. Bathke, T. Kluth, C. Brandt, and P. Maass, “Improved image reconstruction in magnetic particle imaging using structural a priori information,” *Int. J. Magn. Part. Imaging*, vol. 3, p. 10, 3 2017.
- [113] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

- [114] M. Özbey, O. Dalmaz, S. U. Dar, H. A. Bedel, S. Öztürk, A. Güngör, and T. Çukur, “Unsupervised medical image translation with adversarial diffusion models,” *arXiv:2207.08208*, 2022.
- [115] E. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [116] K. Scheffler, M. Boberg, and T. Knopp, “Extrapolation of system matrices in magnetic particle imaging,” *IEEE Transactions on Medical Imaging*, pp. 1–1, 2022.
- [117] M. Yurt, O. Dalmaz, S. Dar, M. Ozbey, B. Tınaz, K. Oguz, and T. Çukur, “Semi-supervised learning of mri synthesis without fully-sampled ground truths,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3895–3906, 2022.
- [118] J. Liu, Y. Sun, C. Eldeniz, W. Gan, H. An, and U. S. Kamilov, “RARE: Image reconstruction using deep priors learned without groundtruth,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1088–1099, 2020.
- [119] R. Laumont, V. D. Bortoli, A. Almansa, J. Delon, A. Durmus, and M. Pereyra, “Bayesian imaging using plug & play priors: When langevin meets tweedie,” *SIAM Journal on Imaging Sciences*, vol. 15, no. 2, pp. 701–737, 2022.
- [120] G. Elmas, S. U. Dar, Y. Korkmaz, E. Ceyani, B. Susam, M. Ozbey, S. Avestimehr, and T. Çukur, “Federated learning of generative image priors for mri reconstruction,” *IEEE Transactions on Medical Imaging*, pp. 1–1, 2022.
- [121] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” in *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.
- [122] H. Chung and J. C. Ye, “Score-based diffusion models for accelerated mri,” *Medical Image Analysis*, vol. 80, p. 102479, 2022.

- [123] J. Liu, R. Anirudh, J. J. Thiagarajan, S. He, K. A. Mohan, U. S. Kamilov, and H. Kim, “DOLCE: A model-based probabilistic diffusion framework for limited-angle ct reconstruction,” *arXiv:2211.12340*, 2022.



Appendix A

Code

Code for TranSMS is available at <https://github.com/icon-lab/TranSMS>.

Code for DEQ-MPI will be available at <https://github.com/icon-lab/DEQ-MPI>.