

Learned Discrepancy Reconstruction and Benchmark Dataset for Magnetic Particle Imaging

Meira Iske, Hannes Albers, Tobias Knopp, and Tobias Kluth

Abstract—Magnetic Particle Imaging (MPI) is an emerging imaging modality based on the magnetic response of superparamagnetic iron oxide nanoparticles to achieve high-resolution and real-time imaging without harmful radiation. One key challenge in the MPI image reconstruction task arises from its underlying noise model, which does not fulfill the implicit Gaussian assumptions that are made when applying traditional reconstruction approaches. To address this challenge, we introduce the Learned Discrepancy Approach, a novel learning-based reconstruction method for inverse problems that includes a learned discrepancy function. It enhances traditional techniques by incorporating an invertible neural network to explicitly model problem-specific noise distributions. This approach does not rely on implicit Gaussian noise assumptions, making it especially suited to handle the sophisticated noise model in MPI and also applicable to other inverse problems. To further advance MPI reconstruction techniques, we introduce the MPI-MNIST dataset — a large collection of simulated MPI measurements derived from the MNIST dataset of handwritten digits. The dataset includes noise-perturbed measurements generated from state-of-the-art model-based system matrices and measurements of a preclinical MPI scanner device. This provides a realistic and flexible environment for algorithm testing. Validated against the MPI-MNIST dataset, our method demonstrates significant improvements in reconstruction quality in terms of structural similarity, achieving up to 7.9% higher SSIM as well as 2.2 dB higher PSNR compared to classical reconstruction techniques across varying noise levels, underscoring its robustness in high-noise scenarios.

Index Terms—Magnetic particle imaging, Bayesian inverse problems, dataset, deep learning, invertible neural networks

I. INTRODUCTION

MAGNETIC Particle Imaging (MPI) is a promising imaging modality that exploits the magnetization behaviour of super paramagnetic nanoparticles [1]. The non-linear magnetization response of the particles to externally applied magnetic fields induces a voltage, which is measured by receive coils. The acquired signal can be translated to the spatial distribution of the particle concentration, thereby enabling image reconstruction [2]. MPI offers high spatial resolution without the use of harmful radiation. Additionally, it exhibits high sensitivity, with the signal remaining unaffected by background noise from surrounding tissues [3]. Due to

its rapid data acquisition capabilities, this imaging technique enables real-time imaging [4], [5]. These characteristics make MPI advantageous for a range of clinical applications, such as vascular applications [6] including blood flow imaging [7], [8] and instrument tracking [9]. For instance, it has applications in the precise positioning of stents [10] or catheters [11], [12]. Additionally, by labeling cells with super paramagnetic particles, MPI can be applied for cell tracking [13]–[15], which is particularly relevant for cancer detection and cancer therapy [16], [17].

The reconstruction of the particle concentration requires solving a linear ill-posed inverse problem [18], [19] by the means of the system matrix, which constitutes the forward operator. However, appropriate reconstruction quality is concerned with major aspects of the modeling chain such as a careful determination of the system matrix and accurate modeling of noise conditions. The former aspect can be accomplished through a data-driven calibration approach. Although this method accounts for all relevant physical properties, it introduces noise and its measurement process is time-consuming. These limitations have led to the development of model-based system matrices, that determine the MPI system matrix based on physical models. The latter allows for noise-free system matrices and potentially much faster generation times. Initial model-based methods [20], [21] employed the simplified equilibrium model, which assumes no relaxation time for the magnetic nanoparticles. More advanced approaches include the effect of Brownian [22], [23] and Néel rotation [24], [25] of the magnetic nanoparticles on its magnetic moment by solving the Fokker-Planck equation. For additional details on model-based system matrices, we refer to [23], [26] and references therein.

The second essential aspect of the modeling chain focuses on the reduction of unknown noise introduced by the MPI scanner. While in theory the noise is primarily caused by the resistive components in the receive chain of the MPI scanner and follows a standard normal distribution, in practice thermal effects in the electronic components lead to temporal drifts that are difficult to model [27]. The noise is also often correlated across frequencies due to non-linear mixing in certain scanner components, which is difficult to avoid completely in practice. Previous approaches [27], [28] have addressed this component by applying a background correction from scanner measurements prior to the image reconstruction procedure. The conventional reconstruction objective involves minimizing the Tikhonov functional, typically achieved through the application of the regularized Kaczmarz algorithm [5], [7], [24]. Subsequent research [18], [29] adapts the formulation of the common objective by introducing a linear transform

M. Iske is with the Center for Industrial Mathematics, University of Bremen, Germany. (Email: iskem@uni-bremen.de).

H. Albers is with the Center for Industrial Mathematics, University of Bremen, Germany.

T. Knopp is with the Section for Biomedical Imaging, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, with the Institute for Biomedical Imaging, Hamburg University of Technology, Hamburg, Germany, and with the Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering IMTE, Lübeck, Germany.

T. Kluth is with the Center for Industrial Mathematics, University of Bremen, Germany.

to the data discrepancy term of the Tikhonov-type framework. While the linear transform in [18] is defined by the row-energy of the system matrix, [29] expands upon this by applying a whitening operation that directly includes the scanner noise characteristics in a data-based fashion. A derivation of the Tikhonov objective from a Bayesian perspective by estimating the maximum a posteriori (MAP) amounts to making a standard Gaussian assumption on the noise distribution. Similarly, for the adapted objective in [18], [29] one assumes Gaussian distributed noise with a covariance matrix depending on the transformation matrix.

Beyond noise-specific methods, further reconstruction approaches address system matrix deviations and data structure priors. For instance, total least squares methods [30] and the RESESOP algorithm [31], [32] focus on system matrix inaccuracies. Prior data knowledge can be introduced by applying a TV penalty term [33], [34] or by using an ℓ^1 -data discrepancy term [35] in the formulation of the objective. In recent years, there has been significant interest in the development of learning-based methods for MPI, particularly those leveraging deep learning techniques. These data-driven methods seem extremely promising due to their ability of capturing problem specific data structures and have been applied by following various strategies. Post-processing tools utilize neural networks to refine images reconstructed by traditional methods [36]. End-to-end learning approaches have gained attention for their ability to directly learn the inverse of the system matrix [37]–[39], where plug-and-play methods integrate neural networks into an iterative reconstruction framework [40]–[42]. In this context, the neural network is typically responsible for denoising in image space. Similarly, deep equilibrium approaches apply iterative schemes as in [43], [44], by employing a fixed-point model where the output of the neural network represents the equilibrium point of the iterative process. The unsupervised Deep Image Prior (DIP) approach [45] has also been applied in MPI, where the neural network itself acts as a prior, capturing image structures without the need for extensive training data. Some of these learning-based approaches explicitly target the crucial aspect of the data discrepancy, which is determined by the noise behaviour, for instance by incorporating iterative denoising steps [40]–[42]. However, these approaches rely on simplified Gaussian assumptions regarding the MPI noise model. The approaches in [43], [44] integrate a learned discrepancy term into a deep equilibrium framework, enabling the model to implicitly adapt to noise characteristics without requiring an explicit noise model. Although this provides flexibility, it also means that the implicitly learned iterative correction is difficult to interpret from a statistical perspective. To the best of our knowledge, no existing method explicitly models and leverages the true statistical noise characteristics of MPI within the reconstruction process.

The application of suitable reconstruction schemes is crucial in order to obtain reasonable reconstructions. The development of effective reconstruction techniques, in turn, relies on the availability of validation data. However, the limited number of MPI scanners poses a challenge in providing the research community with sufficient access to reference data. So far, the

publicly available OpenMPI [46] dataset has been widely used for the evaluation of reconstruction algorithms [29], [35], [45], [47]. Despite representing an initial step toward establishing a data foundation for the evaluation of image reconstruction approaches, the number of available MPI data samples is still limited. Moreover, there is a general lack of paired ground truth measurements in MPI. This presents a significant challenge when it comes to the conception of learning-based approaches, which typically require a large amount of data.

In this work we revisit two key challenges in MPI reconstruction. First, we propose the Learned Discrepancy Approach (LDA), a data-based reconstruction approach that directly accounts for complex and problem-specific noise models, thereby addressing one aspect of the modeling chain. Unlike previous methods that use simplified assumptions on the noise model, leading to standard discrepancy terms within the Tikhonov framework, the LDA incorporates a learned discrepancy term. This term relies on an invertible neural network trained to explicitly learn the unknown noise distribution. This results in an objective consistent with the variational regularization framework of Tikhonov-type, which is tailored to the underlying noise characteristics that were captured during training. While related learning-based approaches in other contexts, such as [48], [49], have focused on learning a parametrization of discrepancy functions for preselected noise distribution classes, our method derives the entire discrepancy functional from a Bayesian perspective and thus requires no prior assumption on the considered noise distribution class. This results from a direct integration of the learned noise distribution into the variational reconstruction framework. In contrast to existing reconstruction methods, our approach generalizes the classical Tikhonov objective by explicitly integrating the problem-specific noise distribution. Thus, the approach is capable of including correlations across different components of the noise model and is transferable to a wide range of inverse problems.

Second, we introduce the MPI-MNIST dataset, based on the MNIST dataset of handwritten digits [50]. It includes both simulated MPI measurements obtained by using a model-based system matrix in combination with real scanner noise from the preclinical MPI system (Bruker, Ettlingen, Germany). This enables a realistic yet verifiable testing environment for reconstruction algorithms. Moreover, the large set of contained noise measurements forms the foundation of our proposed data-driven reconstruction method, allowing for a realistic modeling of the underlying noise distribution in MPI. With the MPI-MNIST dataset, we intend to make a further step toward closing the existing data gap in MPI.

This work is structured as follows: In Section II we cover different reconstruction methods, where Section II-A focuses on common approaches in the context of MPI image reconstruction. In Section II-B we introduce the Learned Discrepancy Approach. Section III details the generation of the model-based system matrix, which is followed by the process of measuring noise samples in Section IV that both build the foundation of the MPI-MNIST dataset. The dataset's structure is thoroughly described in Section V, which is applied to our reconstruction approach in Section VI. In this section,

we present a qualitative and quantitative comparison of our approach with the methods from Section II-A.

II. METHODS

We consider reconstruction methods with broad applicability to linear inverse problems, where $X \subseteq \mathbb{R}^N$, $Y \subseteq \mathbb{R}^M$ and $A : X \rightarrow Y$ a linear operator satisfying

$$Ax = y \quad (1)$$

for $x \in X$ and $y \in Y$. In this setting, we aim to reconstruct x from noise-perturbed data $y^\delta = y + \eta$ with $\|\eta\| \leq \delta$.

A. Standard Reconstruction Approaches

Typically, inverse problems of type (1) are solved by the concept of variational regularization, namely by minimizing

$$\mathcal{T}(x) = \mathcal{D}(Ax, y^\delta) + \mathcal{R}(x),$$

where $\mathcal{D} : Y \times Y \rightarrow [0, \infty]$ measures the data discrepancy and $\mathcal{R} : X \rightarrow (-\infty, \infty]$ denotes the regularization function. Specific choices for both functions lead to the definition of the *Tikhonov functional*

$$\mathcal{T}_\alpha^{y^\delta}(x) = \|y^\delta - Ax\|_Y^2 + \alpha \|x\|_X^2, \quad (2)$$

with $\alpha > 0$. The analytical solution of (2) is obtained by

$$\hat{x}_{\text{Tikhonov}} = (A^H A + \alpha I)^{-1} A^H y^\delta,$$

where A^H denotes the conjugate transpose of A .

A frequently applied reconstruction approach in MPI image reconstruction is the so-called *regularized Kaczmarz* (RK) algorithm, as proposed in [51]. This well-established reconstruction method iteratively converges to the minimizer of (2) and is often chosen due to its rapid convergence while still delivering reconstructions of notably high quality as demonstrated in [24], [29].

Beyond the conventional regularized Kaczmarz algorithm, [29] have shown the success of the whitened variant of this algorithm. Its objective is expressed as

$$\hat{x} = \arg \min_x \frac{1}{2} \|W(y^\delta - Ax)\|^2 + \frac{\alpha}{2} \|x\|^2, \quad (3)$$

where $W \in \mathbb{R}^{K \times K}$ a diagonal matrix with entries $w_{j,j} = \min_k \text{std}(\eta_k) / \text{std}(\eta_j)$ for $j = 1, \dots, K$. Here, $\text{std}(\eta_j)$ is computed for each noise component j over the additive measurement noise on the train set. The minimizer in (3) is again determined via the regularized Kaczmarz algorithm, by defining $A_W := WA$ and $y_W^\delta = W y^\delta$ in (2). This approach balances the impact of different noise features within the optimization process. Following the terminology of [29], we refer to the latter reconstruction method as *whitened regularized Kaczmarz* (WRK).

B. Learned Discrepancy Approach

In this section, we introduce the Learned Discrepancy Approach (LDA), a learning-based reconstruction method, which is broadly applicable to inverse problems following an operator equation as in (1). The proposed method is derived from a Bayesian perspective and aims to encode the intrinsic noise distribution of the underlying measurement noise. Since MPI noise characteristics can at this point only be inferred through measurement data, the approach seems especially promising for MPI image reconstruction tasks. It is particularly effective with the data collection contained in the MPI-MNIST dataset as proposed in Section V. By leveraging MPI-specific noise characteristics, our approach holds significant promise for enhancing reconstruction accuracy.

We assume that η follows a probability density function (pdf) $p_H : Y \rightarrow \mathbb{R}_0^+$ and x follows the prior pdf $p_X : X \rightarrow \mathbb{R}_0^+$. In the following, we will further assume that the random variables $x \sim p_X$ and $\eta \sim p_H$ are stochastically independent.

When performing reconstruction tasks, from a probabilistic point of view, we are interested in the posterior distribution $p(x|y^\delta)$ in order to identify the value of x that maximizes the probability given the observed data y^δ . According to Bayes' Theorem, the posterior distribution is given by

$$p(x|y^\delta) = \frac{p(y^\delta|x)p_X(x)}{p_Y(y^\delta)} \propto p(y^\delta|x)p_X(x), \quad (4)$$

where p_Y describes the pdf of $y \in Y$. The maximum a posteriori (MAP) estimate $\hat{x} \in X$ is obtained by maximizing the posterior distribution, i.e.,

$$\begin{aligned} \hat{x} &= \arg \max_x p(x|y^\delta) = \arg \max_x \log p(x|y^\delta) \\ &= \arg \min_x -\log p(x|y^\delta). \end{aligned} \quad (5)$$

Combining (4) and (5) implies

$$\hat{x} = \arg \min_x -\log p(y^\delta|x) - \log p_X(x). \quad (6)$$

Due to the independence of x and η , the likelihood can be expressed as $p(y^\delta|x) = p_H(y^\delta - Ax)$. Commonly, η is assumed to follow a Gaussian distribution $p_H = \mathcal{N}(0, I)$, which, together with (6) leads to the objective

$$\hat{x} = \arg \min_x \frac{1}{2} \|y^\delta - Ax\|^2 + \mathcal{R}(x), \quad (7)$$

as known from the Tikhonov regularization with regularizing functional $\mathcal{R}(x) = -\log p_X(x)$. By making a Gaussian assumption on the prior, i.e., assuming $p_X = \mathcal{N}(0, \frac{1}{\alpha} I)$ for $\alpha > 0$, we obtain $\mathcal{R}(x) = \frac{\alpha}{2} \|x\|^2$ as in (2).

However, in practical scenarios, where the noise follows some sophisticated structures due to given physical circumstances or external influences, it can be restrictive to assume Gaussian distributed noise. Our aim is to reformulate and generalize (2) by allowing for arbitrary pdfs p_H . We exploit the features of invertible neural networks (INNs), which can learn approximations of point-wise evaluations of an unknown p_H . We adapt the specific choice for the discrepancy term of $\mathcal{D}(Ax, y^\delta) = \|y^\delta - Ax\|_Y^2$ in (2), by tailoring the definition of

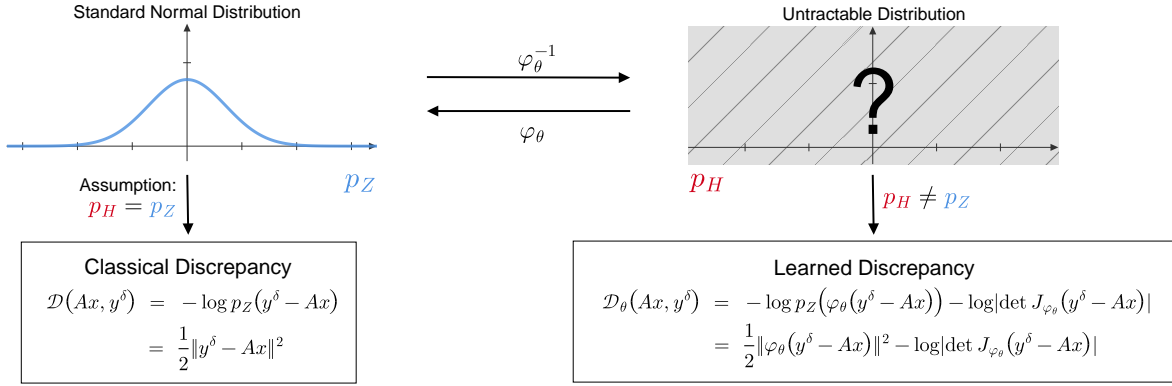


Fig. 1: Illustration of the LDA. Oftentimes, the noise distribution is assumed to be standard Gaussian, leading to a fixed, analytical form of the discrepancy term \mathcal{D} . In contrast, the LDA models the noise distribution in a data-driven, problem-specific manner using an invertible neural network φ_θ . This allows for a flexible discrepancy term depending on φ_θ that reflects the true, potentially complicated noise structure.

\mathcal{D} to the noise distribution p_H . Figure 1 provides an illustration of this concept.

Let $z \in \mathbb{R}^M$ be a random variable following a known and tractable probability density function p_Z , which we will refer to as the latent distribution. Let $g : \mathbb{R}^M \rightarrow \mathbb{R}^M$ be a diffeomorphism, i.e., an invertible and differentiable mapping. Furthermore, let g be such that $z = g(\eta)$. By the change-of-variable theorem it holds

$$p_H(\eta) = p_Z(g(\eta)) |\det J_g(\eta)|, \quad (8)$$

where $J_g(\eta) := \frac{\partial g}{\partial \eta}$ denotes the Jacobian of g . Therefore, the underlying probability density function p_H can be expressed as a transformation of the latent distribution p_Z through the diffeomorphism g .

To approximate the true underlying distribution p_H , we train an invertible neural network (INN) $\varphi_\theta : \mathbb{R}^M \rightarrow \mathbb{R}^M$ parameterized by $\theta \in \Theta$, where Θ denotes the parameter space of the neural network in order to model the diffeomorphism g given in (8). The goal is to find a transformation φ_θ that maps an input η to a latent variable z such that $p_\theta(\eta) \approx p_H(\eta)$ can be approximated, where $p_\theta(\eta)$ is the model distribution defined by

$$p_\theta(\eta) = p_Z(\varphi_\theta(\eta)) |\det(J_{\varphi_\theta}(\eta))|.$$

The network is trained on a dataset $\{\eta^{(i)}\}_{i=1}^L$ by maximizing the log-likelihood

$$\begin{aligned} & \max_{\theta \in \Theta} \mathcal{L}(\theta) \\ & \text{with } \mathcal{L}(\theta) = \sum_{i=1}^L \log(p_\theta(\eta^{(i)})) \\ & = \sum_{i=1}^L \left(\log p_Z(\varphi_\theta(\eta^{(i)})) + \log |\det J_{\varphi_\theta}(\eta^{(i)})| \right), \end{aligned} \quad (9)$$

which is equivalent to minimizing the Kullback-Leibler divergence $D_{\text{KL}}(p_H \| p_\theta)$. By choosing our latent distribution p_Z to

be of Gaussian nature, together with (6) we have

$$\begin{aligned} \hat{x} &= \arg \min_x -\log p_H(y^\delta - Ax) - \log p_X(x) \\ &= \arg \min_x -\log p_Z(\varphi_\theta(y^\delta - Ax)) \\ &\quad - \log |\det J_{\varphi_\theta}(y^\delta - Ax)| - \log p_X(x) \\ &= \arg \min_x \frac{1}{2} \|\varphi_\theta(y^\delta - Ax)\|^2 - \log |\det J_{\varphi_\theta}(y^\delta - Ax)| \\ &\quad - \log p_X(x), \end{aligned}$$

which is a modified version of the typical MAP estimation that incorporates a normalizing flow to account for a more complex noise distribution beyond the standard Gaussian assumption. We define

$$\mathcal{D}_\theta(Ax, y^\delta) := \frac{1}{2} \|\varphi_\theta(y^\delta - Ax)\|^2 - \log |\det J_{\varphi_\theta}(y^\delta - Ax)|,$$

so that our objectives writes as

$$x_{\text{LDA}} = \arg \min_x \mathcal{D}_\theta(Ax, y^\delta) - \mathcal{R}(x), \quad (10)$$

where x_{LDA} is the minimizer of the Learned Discrepancy Approach (LDA). The LDA incorporates a learned discrepancy functional \mathcal{D}_θ and extends the objective in (2) by the log-determinant of the trained network and an evaluation of the network within the discrepancy term. Therefore, we refer to φ_θ as the *discrepancy network*.

III. PHYSICAL MODEL

As a preparatory step for data generation and for testing the proposed algorithm in Section II-B in the context of MPI, we simulate system matrices in a number of different settings. This simulation concept is in line with the one given in [24], which we extend by using a linear combination of simulations in order to obtain a better approximation of a measured system matrix. One such system matrix was simulated to provide, to the best of our knowledge, the closest approximation to a measured system matrix in the same setting. We therefore treat this reference simulation as the best possible approximation to a noise-free system matrix. The remaining simulated system

matrices are used for numerical experiments, where we assume that only a system matrix of lower quality than the real-life physical system is available. This allows for testing in contexts such as operator correction or post-processing of artifacts resulting from incomplete or inaccurate modeling.

A. System Matrix Simulation Framework

The voltage induced in a receive coil of an MPI scanner with sensitivity profile $p : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ can be described by

$$\tilde{v}(t) = -\mu_0 \int_{\Omega} c(x)p(x)^\top \frac{\partial}{\partial t} \bar{m}(x, t) dx$$

in V , where $c : \Omega \rightarrow \mathbb{R}_0^+$ in mol/L is the concentration of the magnetic nanoparticles, $\Omega \subset \mathbb{R}^3$ the imaging volume and $\bar{m} : \mathbb{R}^3 \times [0, T] \rightarrow \mathbb{R}^3$ the mean magnetic moment in $10^{-3} \text{Am}^2 \text{mol}^{-1}$. The direct feedthrough of the signal is filtered by an analog filter that can be expressed by a filter kernel $a : \mathbb{R} \rightarrow \mathbb{R}$, where $v = a * \tilde{v}$ the filtered signal. The T -periodic signal v can consequently be expanded into a Fourier series with coefficients

$$\hat{v}_k = -a_k \frac{\mu_0}{T} \int_{\Omega} c(x)p(x)^\top \int_0^T \left(\frac{\partial}{\partial t} \bar{m}(x, t) \right) \exp\left(\frac{-2\pi i t k}{T}\right) dt dx$$

for $k \in \mathbb{N}_0$. By defining the system function

$$s_k(x) = -a_k \frac{\mu_0}{T} \int_0^T p(x)^\top \left(\frac{\partial}{\partial t} \bar{m}(x, t) \right) \exp\left(\frac{-2\pi i t k}{T}\right) dt,$$

the Fourier coefficients write as $\hat{v}_k = \int_{\Omega} c(x)s_k(x) dx$. Sampling from the system function then yields the system matrix $A = (s_k(x_\ell))_{k=0, \dots, K-1; \ell=1, \dots, N} \in \mathbb{C}^{K \times N}$. A crucial part of the computation of the components s_k requires access to the mean magnetic moment, depending on the T -periodic external magnetic field $H_{\text{app}} : \mathbb{R}^3 \times [0, T] \rightarrow \mathbb{R}^3$ in Am^{-1} . The applied magnetic field combines the selection field $H_S : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ with the drive field $H_D : \mathbb{R}^3 \times [0, T] \rightarrow \mathbb{R}^3$ by $H_{\text{app}}(x, t) = H_S(x) + H_D(x, t)$. The magnetic mean writes as

$$\bar{m}(x, t) = m_0 \int_{S^2} m f(m, x, t) dm, \quad (11)$$

where $m_0 = V_C M_S$ with $V_C = \frac{1}{6} \pi D_{\text{core}}^2$ and f the solution to the Fokker-Planck equation

$$\frac{\partial}{\partial t} f = \text{div}_{S^2} \left(\frac{1}{2\tau} \Delta_{S^2} f \right) - \text{div}_{S^2} (b f),$$

with $\tau > 0$ the relaxation time constant and (velocity) field $b : S^2 \times \mathbb{R}^3 \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$, where

$$b(m, H_{\text{app}}, n) = p_1 \times m + p_2 (m \times H_{\text{app}}) \times m + p_3 (n \cdot m) n \times m + p_4 (n \cdot m) (m \times n) \times m,$$

with $n \in S^2$ the easy axis of the magnetic particles. In this setting, we consider a pure Néel rotation of the magnetic particles including anisotropy, defining the physical constants p_i , $i = 1, \dots, 4$ by $p_1 = \tilde{\gamma} \mu_0$, $p_2 = \tilde{\gamma} \alpha \mu_0$, $p_3 = p_4 = 2\tilde{\gamma} \frac{K_{\text{anis}}}{M_S}$ and $\tau = \frac{V_C M_S}{2k_B T_B \tilde{\gamma} \alpha}$ with $\tilde{\gamma} = \frac{\gamma}{1 + \alpha^2}$. To model fluid particles, the easy axis and the anisotropy constant are selected in analogy to Model B3 of [24]: The easy axis $n(x) = \frac{H_S(x)}{|H_S(x)|}$ is chosen in the direction of the selection field and the anisotropy constant $K_{\text{anis}}(x) = K_{\text{anis}}^{\max} \frac{|H_S(x)|}{h}$, with constant K_{anis}^{\max} and

$h := \max_{x \in \Omega} |H_S(x)|$, is chosen proportional to the selection field.

The system matrix simulation was performed by solving the Fokker-Planck equation for each pixel as described in [52]. Different system matrices were created for a number of different assumed maximum Néel anisotropy constants K_{anis}^{\max} and particle core diameters D_{core} . In particular, system matrices were simulated for the following equidistant parameters:

$$D_{\text{core}} \in \mathcal{D} = \{16, 18, \dots, 24 \text{ nm}\},$$

$$K_{\text{anis}}^{\max} \in \mathcal{K} = \{500, 800, \dots, 3800 \text{ J/m}^3\}.$$

We denote a system matrix corresponding to the combination of these parameters by $A_{\mathcal{D}(i), \mathcal{K}(j)}$ with $\mathcal{D}(i)$ the i -th element of \mathcal{D} and $\mathcal{K}(j)$ the j -th element of \mathcal{K} .

To model the most accurate polydisperse system matrix, the system matrices corresponding to the different parameter combinations were equipped with weights $w_{i,j} \geq 0$, which were estimated in a least-squares sense by

$$W^* = \arg \min_W \|A_{\text{calib}} - \sum_{i,j} w_{i,j} A_{\mathcal{D}(i), \mathcal{K}(j)}\|^2, \quad (12)$$

where $W = (w_{i,j})_{i=1, \dots, 5; j=1, \dots, 12}$ and A_{calib} is a suitable measured reference matrix. The resulting matrix $A_{\text{opt}} = \sum_{i,j} w_{i,j}^* A_{\mathcal{D}(i), \mathcal{K}(j)}$ with $W^* = (w_{i,j}^*)_{i=1, \dots, 5; j=1, \dots, 12}$ will be treated as the most accurate available simulated matrix. In addition, the two system matrices with the lowest and with the highest anisotropy and diameter were included in the dataset to include mismatched or unrealistic system matrices. These are denoted $A_{\text{SP}} := A_{\mathcal{D}(1), \mathcal{K}(1)}$ for the small parameters and $A_{\text{LP}} := A_{\mathcal{D}(5), \mathcal{K}(12)}$ for the large parameters. Lastly, a standard equilibrium model was evaluated for $D_{\text{core}} = 20 \text{ nm}$ to serve as a crude approximation of the true system matrix, denoted by A_{eq} .

Finally, all system matrices were scaled and time-shifted to fit A_{calib} . The measured transfer function [53] was added to each system matrix for its comparability to measurements. In total, we result in a set of system matrices where A_{opt} is as close as possible to the real-world situation, while the others contain simplifications or model errors of varying magnitude.

B. Experimental Setup for A_{calib}

The preclinical MPI device (Bruker, Ettlingen, Germany), which was used for measuring A_{calib} has three send and three receive channels with excitation fields of $12 \text{ mT} \mu_0^{-1}$ in both the x - and y -direction. The selected frequencies of the drive field are $f_x = 2.5/102 \text{ MHz}$ in the x -direction and $f_y = 2.5/96 \text{ MHz}$ in the y -direction. This results in a trajectory cycle of $652.8 \mu\text{s}$. The selection field gradient has a strength of $-1.0, -1.0$ and $2.0 \text{ T} \mu_0^{-1} \text{m}^{-1}$ in x -, y -, and z -direction, respectively. The Lissajous trajectory is of cosine type, specifically cosine in x -direction and negative cosine in the y -direction. Signal reception uses the built-in 3D receive coils. The analog signal undergoes band-pass filtering and is digitized at a sampling rate of 2.5 MHz , where 1632 sampling points were captured for one Lissajous period, thus resulting in $K = 817$ frequency components. A block average

over 10 drive-field periods is applied during data acquisition. Spatial sampling of the system matrix takes place on a grid of size 17×15 with $N_{\text{coarse}} = 255$ equally spaced measurement points, covering a $34 \times 34 \text{ mm}^2$ field of view. The used delta probe of size $2 \times 2 \times 2 \text{ mm}^3$ consists of Perimag. This tracer material contains an iron concentration of $10 \text{ mg}_{\text{Fe}}/\text{mL}$. Altogether, this configuration yields the system matrix $A_{\text{calib}} \in \mathbb{C}^{3 \times K \times N}$, the first dimension referring to the different receive coils.

The simulations adapt to the experimental setup of the measured system matrix, while refining the spatial grid to a size of 85×75 . The simulated system matrices were additionally subsampled for two grids of size 51×45 and 17×15 . Each system matrix includes three receive channels, corresponding to the three receive coils integrated into the preclinical MPI system. This results in $A_{\text{opt}}, A_{\text{SP}}, A_{\text{LP}}, A_{\text{eq}} \in \mathbb{C}^{3 \times K \times N}$ with $N \in \{N_{\text{coarse}} = 255, N_{\text{int}} = 2295, N_{\text{fine}} = 6375\}$.

IV. NOISE MEASUREMENTS

To obtain a sufficient amount of noise measurements, the MPI scanner was left to run empty for a duration of approximately one hour while capturing the received voltage signals within this time period. The acquisition setup coincides with the setup of the reference system matrix as described in the previous section. In contrast to the measured system matrix, no block averaging was applied so that each individual acquisition within the Lissajous cycle resulted in a distinct noise frame $\bar{\eta}^{(j)} \in \mathbb{C}^{3 \times K}$ for $j = 1, \dots, 4694400$. Subsequent to the acquisition process, sequences of ten consecutive measurements were averaged by computing

$$\eta^{(i)} = \frac{1}{10} \sum_{j=1}^{10} \bar{\eta}^{(10(i-1)+j)} \in \mathbb{C}^{3 \times K} \quad \text{for } i = 1, \dots, L_{\text{noise}}$$

with $L_{\text{noise}} = 469440$ to produce a single noise sample $\eta^{(i)}$. This is in line with the common measurement practice in MPI of performing multiple MPI measurements with subsequent averaging. The resulting noise samples are partitioned into five groups for the integration of the noise samples to multiple parts of the dataset. Primarily, the captured samples serve as additive measurement noise of the simulated MPI measurements, as described in Section V-B. In addition, these samples can be used to simulate pixelwise deviations of the system matrix as in (14). Both additive measurement noise and system matrix deviations might require background corrections within the reconstruction process. Therefore, we further provide noise samples for background correction of both cases. To account for environmental factors such as temperature fluctuations during the measurement process, we employ alternating sequences of noise acquisitions for the system matrix and its background correction. Specifically, each sequence of noise measurements corresponding to a pixel in the system matrix is followed by an acquisition sequence of equal length reserved for background correction of the respective pixel including a short pause between the two sequences. Moreover, we provide a large amount of additional noise samples. These can be used e.g. for gaining statistical insights on the noise behaviour and are especially suitable for learning-based approaches as in

Section II-B. The organization of these five components is detailed in Section V.

V. MPI-MNIST: DATASET GENERATION AND STRUCTURE

To address the existing data gap in MPI, we introduce the MPI-MNIST dataset containing data tuples of ground truth phantoms along with simulations of corresponding MPI measurements. An overview of the dataset is illustrated in Figure 2. The measurement simulations result from the multi-coil system matrices described in Section III as well as the noise measurements captured by the preclinical MPI scanner, detailed in Section IV. The MPI-MNIST dataset comprises processed versions of all images included in the MNIST dataset of handwritten digits [50]. The MNIST dataset itself contains 70 000 labeled images of handwritten digits between zero and nine, thereby defining the size of the MPI-MNIST dataset. The images were chosen as a suitable ground truth dataset for MPI-MNIST due to the structural similarities of the images compared to common MPI phantoms such as blood vessels, particularly in terms of their sparsity. Consequently, we will refer to the two-dimensional MNIST images as *phantoms* in the following discussion. The system matrices described in Section III are utilized for the simulation of MPI measurements in combination with the scanner data from Section IV, which serve as measurement noise. Further system matrices underlying different physical models at various resolutions are provided. This increases the compatibility of the data with a broad range of reconstruction settings containing potential operator deviations, as illustrated on the right-hand side in Figure 2. Additionally, different system matrices can be selected for data generation, as detailed in Appendix A. The dataset further includes a large collection of noise data, separated into distinct components. These can serve for multiple applications, which are listed in Section V-D2.

The full dataset is publicly accessible via Zenodo (<https://doi.org/10.5281/zenodo.12799417>) at a total size of 273 GB. This iteration of the dataset features a specific setting, for which several parameters, e.g. particle concentration and resolution, had to be fixed. In this section we provide a comprehensive overview of this dataset's structure and data generation together with a comprehensive list of the provided files on the Zenodo website. We discuss potential modifications and extensions of the dataset in the first section of the appendix.

A. Preprocessing of MNIST Phantoms

The creation of data tuples in the MPI-MNIST dataset involves several adaptations to the original MNIST images, serving as the foundation for generating both noise-free and noise-perturbed simulated MPI measurements. The MNIST phantoms are originally given in [50] as images with a size of $28 \times 28 \text{ px}$. According to the adaptation of the simulated system matrix to the measured reference system matrix in Section III, we aim to reconstruct on the same scale, i.e., reconstruct images of $17 \times 15 \text{ px}$ including an additional overscan of the system matrix. For this reason, the MNIST phantoms are first downsampled to the size of $11 \times 11 \text{ px}$ by applying nearest neighbour interpolation and subsequently

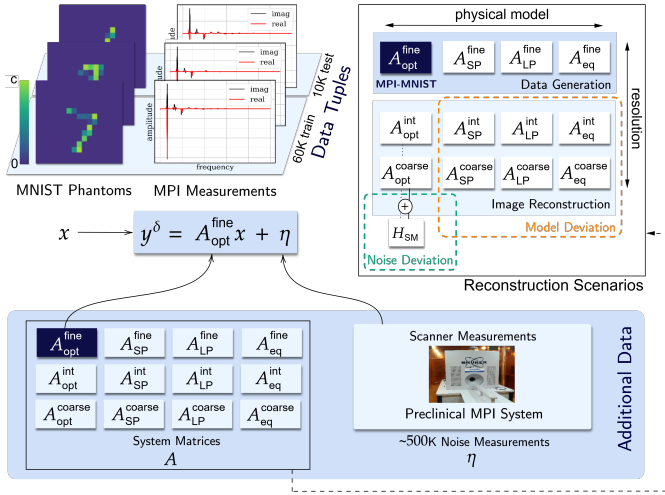


Fig. 2: Overview of the MPI-MNIST dataset and included reconstruction settings.

framed by zero valued pixels. The resulting image is denoted by $x_{\text{coarse}}^{(i)} \in \mathbb{R}^{W_{\text{coarse}} \times H_{\text{coarse}}}$ with $W_{\text{coarse}} = 17$, $H_{\text{coarse}} = 15$ for $i = 1, \dots, 70\,000$. The pixel values of each processed phantom are scaled to the desired pixel range $[0, c]$, determined by the concentration $c = 10$ in $\text{mg}_{\text{Fe}}\text{mL}^{-1}$. Therefore, $(x_{\text{coarse}}^{(i)})_{j,k} \in [0, c]$ for $j = 1, \dots, W_{\text{coarse}}$, $k = 1, \dots, H_{\text{coarse}}$. The samples $x_{\text{coarse}}^{(i)}$ will serve as ground truth images of our dataset. However, an upsampling of the phantoms is required for generation of the MPI data in order to circumvent an inverse crime [54]. We apply nearest neighbour interpolation to $x_{\text{coarse}}^{(i)}$ by scaling each spatial dimension by a factor of $s = 5$. More precisely, we obtain $x_{\text{fine}}^{(i)} \in \mathbb{R}^{W_{\text{fine}} \times H_{\text{fine}}}$ with $W_{\text{fine}} = s \cdot W_{\text{coarse}}$, $H_{\text{fine}} = s \cdot H_{\text{coarse}}$, where

$$(x_{\text{fine}}^{(i)})_{sj+u, sk+v} = (x_{\text{coarse}}^{(i)})_{j,k} \quad \text{for } u, v = 1, \dots, s$$

and $j = 1, \dots, W_{\text{coarse}}$, $k = 1, \dots, H_{\text{coarse}}$ defines the upsampling operation. In preparation for a matrix multiplication of the upsampled phantoms with the system matrix, we flatten the images to vector shape, resulting in $\bar{x}_{\text{fine}}^{(i)} \in \mathbb{R}^{N_{\text{fine}}}$, $N_{\text{fine}} = 6375$, where

$$(\bar{x}_{\text{fine}}^{(i)})_{\ell} = (x_{\text{fine}}^{(i)})_{j,k} \quad \text{with } \ell = j \cdot W_{\text{fine}} + k,$$

$j = 1, \dots, W_{\text{fine}}$, $k = 1, \dots, H_{\text{fine}}$. An analogous transformation is applied in order to obtain $\bar{x}_{\text{coarse}}^{(i)} \in \mathbb{R}^{N_{\text{coarse}}}$ from $x_{\text{coarse}}^{(i)}$ with $N_{\text{coarse}} = 255$.

B. Simulation of MPI-MNIST Measurements

The data samples can be organized as tuples $(x^{(i)}, y^{(i)}, y^{\delta, (i)})$ for $i = 1, \dots, 70\,000$, where

- I $x^{(i)} = \bar{x}_{\text{coarse}}^{(i)} \in \mathbb{R}^{N_{\text{coarse}}}$ as in Section V-A,
- II $y^{(i)} = [A_1 \bar{x}_{\text{fine}}^{(i)} \mid A_2 \bar{x}_{\text{fine}}^{(i)} \mid A_3 \bar{x}_{\text{fine}}^{(i)}]^T \in \mathbb{C}^{3 \times K}$ the measurement of $x^{(i)}$ by channelwise multiplication with the system matrix $A \in \mathbb{C}^{3 \times K \times N_{\text{fine}}}$ from Section III, where $A_j \in \mathbb{C}^{K \times N_{\text{fine}}}$ corresponds to the j -th receive channel of A ,

- III $y^{\delta, (i)} = y^{(i)} + \eta^{(i)} \in \mathbb{C}^{3 \times K}$ the noise-perturbed measurement of $x^{(i)}$ with noise $\eta^{(i)} \in \mathbb{C}^{3 \times K}$, as in Section IV.

For the generation of $y^{(i)}$ we select $A = A_{\text{opt}}$, giving the best possible approximation of the measured system matrix. Therefore, each sample $y^{(i)}$ simulates a multi-coil MPI measurement stemming from a simulated system Matrix with three receive coils.

In compliance with MNIST, we propose a split of the dataset into a training set of $L_{\text{train}} = 60\,000$ samples and a test set of $L_{\text{test}} = 10\,000$ samples. This results in the two sets

- $(x_{\text{train}}^{(i)}, y_{\text{train}}^{(i)}, y_{\text{train}}^{\delta, (i)})_{i=1, \dots, L_{\text{train}}}$,
- $(x_{\text{test}}^{(i)}, y_{\text{test}}^{(i)}, y_{\text{test}}^{\delta, (i)})_{i=1, \dots, L_{\text{test}}}$.

C. Reconstruction System Matrix

To reconstruct $x^{(i)}$ from a given $y^{\delta, (i)}$, a reconstruction system matrix $A_{\text{rec}} \in \mathbb{C}^{3 \times K \times N}$ needs to be selected. To prevent an inverse crime, this matrix is typically of lower spatial resolution compared to the data generation matrix $A \in \mathbb{C}^{3 \times K \times N_{\text{fine}}}$, i.e., $N < N_{\text{fine}}$. In the following we denote the spatially downsampled version of A following the same physical model by \tilde{A} , so that $\tilde{A} \in \mathbb{C}^{3 \times K \times N}$. If the system matrix \tilde{A} differs from system matrix A_{rec} for data reconstruction, e.g. in terms of its underlying physical model, we have an operator deviation $\varepsilon: \mathbb{R}^N \rightarrow \mathbb{C}^{3 \times K}$, where

$$\varepsilon(x) = \tilde{A}x - A_{\text{rec}}x \quad \text{for } x \in \mathbb{R}^N. \quad (13)$$

To address the operator error ε , oftentimes, this requires an appropriate operator correction strategy incorporated by the reconstruction method. This setting is especially relevant in the context of MPI since the system matrix of actual MPI measurements cannot be accessed. We therefore expect some kind of operator deviation when selecting a system matrix for our reconstruction methods.

With the MPI-MNIST dataset, two different types of operator deviations are covered: A *model deviation* and a *noise deviation*. The former case is covered when the physical models of data generation system matrix and reconstruction system matrix differ. The noise deviation can be simulated by applying

$$A_{\text{rec}} = \tilde{A} + H_{\text{SM}} \quad (14)$$

with some noise matrix $H_{\text{SM}} \in \mathbb{C}^{3 \times K \times N}$. For all $j = 1, \dots, N$ the matrix entries $(H_{\text{SM}})_j \in \mathbb{C}^{3 \times K}$ result from averaging over ten empty scanner measurements of the Bruker MPI scanner as described in Section IV. Thus, by adding H_{SM} , we create a pixelwise deviation of the operators given by the noise model of the MPI scanner. Both types of deviations are illustrated in Figure 2. A combination of both deviation types will increase the operator error. Our setting aims to simulate not only the case in which a calibrated system matrix is used for reconstructing from real MPI measurements (noise deviation) but also the case where a model-based system matrix is used (model deviation). A simulation of these scenarios thus might enable the validation of operator correcting reconstruction schemes.

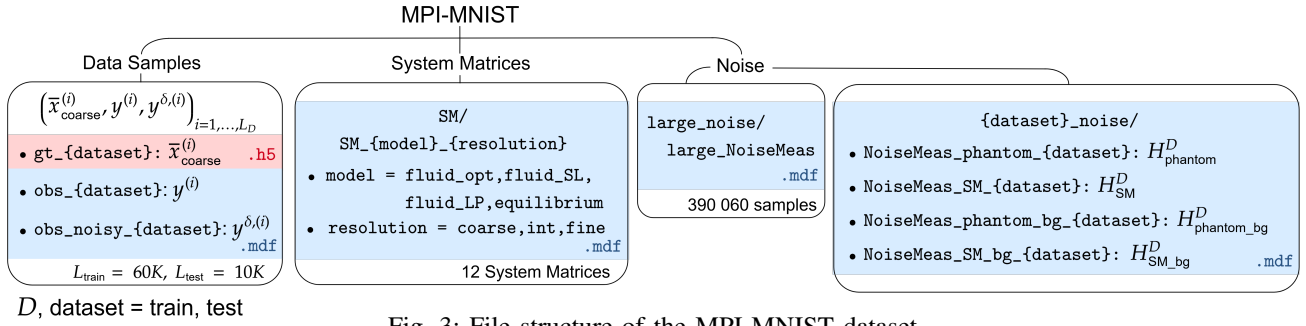


Fig. 3: File structure of the MPI-MNIST dataset.

D. File Structure

The MPI-MNIST data is provided in different files of the MDF data format [55], specifically tailored to MPI measurements. All data, except for the ground truth phantoms, are stored in this file format, while the ground truth phantoms are available as HDF5 files. An overview of the dataset is illustrated in Figure 3, while this section provides a detailed description of all components.

1) *System Matrices:* The file `SM.tar.gz` archives twelve system matrices, covering four physical models at three different resolutions. Each matrix is stored as `/SM/SM_{physical model}_{resolution}.mdf`.

The physical model of the system matrix can be selected by choosing one of the contained cases for physical model:

- `fluid_opt`: $A_{opt} \in \mathbb{C}^{3 \times K \times N}$
- `fluid_small_params`: $A_{SP} \in \mathbb{C}^{3 \times K \times N}$
- `fluid_large_params`: $A_{LP} \in \mathbb{C}^{3 \times K \times N}$
- `equilibrium`: $A_{eq} \in \mathbb{C}^{3 \times K \times N}$

The option resolution covers three resolutions of the spatial system matrix dimension N :

- `fine`: $N = N_{fine} = 6375$ (85×75 px),
- `int(ermediate)`: $N = N_{int} = 2295$ (51×45 px),
- `coarse`: $N = N_{coarse} = 255$ (17×15 px).

The fine resolution is used for generating the MPI measurements together with the physical model `fluid_opt`, while `coarse` equals the selected ground truth dimension of the MNIST phantoms. The additionally provided resolution `int` can be used in a customized setting, as discussed in the first section of the appendix. Furthermore, all provided system matrices can be selected for image reconstruction. In the following, we denote the system matrices by A_p^r with $r = \text{resolution}$ and $p = \text{physical model}$.

2) *Noise Measurements:* The noise measurements as described in Section IV serve distinct purposes and are thus stored in separate files. Therefore, the noise samples $\eta^{(i)}$ for $i = 1, \dots, L_{noise}$ are grouped and stored as matrices $H_g^D \in \mathbb{C}^{L_{noise}^{g(D)} \times 1 \times 3 \times K}$, where $L_{noise}^{g(D)}$ denotes the number of noise samples contained in group $g(D)$ on dataset D with $L_{noise} = \sum_{g,D} L_{noise}^{g(D)}$. Here, $D = \text{dataset}$ either takes the value `train` or `test`. The data is organized as follows:

- `{dataset}_noise.tar.gz` contains
 1. `NoiseMeas_phantom_{dataset}.mdf` with tensor $H_{phantom}^D \in \mathbb{C}^{L_D \times 1 \times 3 \times K}$, L_D noise samples reserved for the additive measurement noise in III,

2. `NoiseMeas_SM_{dataset}.mdf` with tensor $H_{SM}^D \in \mathbb{C}^{N_{int} \times 1 \times 3 \times K}$, N_{int} noise samples as system matrix noise as in (14),
3. `NoiseMeas_phantom_bg_{dataset}.mdf` with tensor $H_{phantom_bg}^D \in \mathbb{C}^{100 \times 1 \times 3 \times K}$, 100 noise samples reserved for background correction of 1,
4. `NoiseMeas_SM_bg_{dataset}.mdf` with tensor $H_{SM_bg}^D \in \mathbb{C}^{N_{int} \times 1 \times 3 \times K}$, N_{int} noise samples reserved for background correction of 2.

- `large_noise.tar.gz` contains `large_NoiseMeas.mdf` with tensor $H_{train} \in \mathbb{C}^{390\,060 \times 1 \times 3 \times K}$: This file contains 390 060 noise samples.

The shape of the listed tensors adapts to the conventions of the MDF file structure. The spatial dimensionality N_{int} of 2. and 4. allows for flexibility when using the intermediate resolution. Nevertheless, only $N_{coarse} = 255$ spatial points are required at the coarsest resolution.

3) *Measurements:* For each part of the dataset ($D = \text{dataset} = \text{train}$ or $D = \text{dataset} = \text{test}$), we store the MNIST phantoms and phantom measurements with the following structure:

- `{dataset}_obs.tar.gz` contains `{dataset}_obs.mdf`: Set of noise-free MPI measurements $y = [y^{(1)} | \dots | y^{(L_D)}] \in \mathbb{R}^{L_D \times 1 \times 3 \times K}$ of all phantoms resulting from $y^{(i)} = A_{opt}^{fine} \bar{x}_{fine}^{(i)}$ as in II,
- `{dataset}_obsnoisy.tar.gz` contains `{dataset}_obsnoisy.mdf`: Set of noise-perturbed MPI measurements $y^\delta = [y^{\delta,(1)} | \dots | y^{\delta,(L_D)}] \in \mathbb{R}^{L_D \times 1 \times 3 \times K}$ of all phantoms resulting from $y^\delta = y + H_{phantom}^D$ as in III.

4) *Phantoms:* The phantom files are stored in `{dataset}_gt.tar.gz` as `{dataset}_gt.hdf5` for each part of the dataset (`dataset = train` or `dataset = test`). These contain the set of ground truth preprocessed and flattened MNIST phantoms $\bar{x}_{coarse} \in \mathbb{R}^{L_D \times N_{coarse}}$ with the train and test split.

E. Metadata

The MDF file format is designed for storing merged MPI data together with all relevant meta information of the experimental setup. We adhered to the conventions established in [55] for our data and made necessary adaptations.

The provided system matrices of simulated nature were adapted to a chosen calibrated system matrix. Therefore, all experimental parameters align with those of A_{calib} . This also includes the signal-to-noise ratio (SNR) stored in `/calibration/snr` of all system matrix files. Although the simulated system matrix does not contain noise, the SNR serves as an orientation for frequency selection during the reconstruction process, since its threshold can be seen as a regularization parameter, for further details see e.g. [56]. In addition to the conventional meta information, we added two parameters to the `/tracer` group of `SM_fluid_opt_{resolution}.mdf`, `SM_fluid_small_params_{resolution}.mdf`, `SM_fluid_large_params_{resolution}.mdf` at all three resolutions:

- `_K_anis`: Anisotropy constants $K_{\text{anis}}^{\text{max}}$ used, in J/m^3 ,
- `_diameters`: Particle diameters D_{core} used, in nm.

Since `SM_fluid_opt_{resolution}.mdf` contains multiple anisotropy constants and particle diameters, the additional parameter `_weights` was added to the `/tracer` group and captures the weight matrix W^* as in (12) of all anisotropy constants and particle diameters. For the equilibrium model, only the parameter `_diameters` was added to the file's meta information, since its model does not depend on any anisotropy constant.

VI. NUMERICAL RESULTS

In this section we evaluate reconstructions of MPI measurements from the MPI-MNIST dataset by applying the LDA image reconstruction approach. This is complemented by a comparison of our learning-based method to standard reconstruction techniques listed in Section II-A.

For the reconstruction of MNIST phantoms from simulated MPI measurements we choose the default MPI-MNIST setting, which selects $A = A_{\text{opt}}^{\text{fine}}$ for data generation of the measurements $(y^{(i)})_i$ together with the noise matrices $H_{\text{phantom}}^{\text{train}}, H_{\text{phantom}}^{\text{test}}$ for computing $(y^{\delta,(i)})_i$. The reconstruction of the phantoms $(x^{(i)})_i$ is performed at the coarse resolution N_{coarse} . Therefore, at a concentration of $c = 10 \text{ mg}_{\text{Fe}}/\text{mL}$ we can simply use the data tuples $(x^{(i)}, y^{(i)}, y^{\delta,(i)})_i$ as provided on the Zenodo website, detailed in Section V. For our reconstruction scenario we choose the same physical model at coarse resolution, i.e., we select $A_{\text{rec}} = A_{\text{opt}}^{\text{coarse}}$ for reconstruction. For all involved components in the frequency domain, we apply a selection of frequency entries $I = \{50, \dots, 813\} \subset I_K$, where $I_L = \{0, \dots, L-1\}$ the index set of length $L \in \mathbb{N}$. This corresponds to applying a band-pass filter with a frequency range of [77 kHz, 1245 kHz]. Subsequently, we stack the selected frequency entries along the channel dimension, resulting in $K' = 2292$ frequency components. For $A_j \in \mathbb{R}^{K \times N_{\text{fine}}}$ the j -th channel of A , we redefine $A = [A_1|_{I \times I_{N_{\text{fine}}}}, A_2|_{I \times I_{N_{\text{fine}}}}, A_3|_{I \times I_{N_{\text{fine}}}}] \in \mathbb{C}^{K' \times N_{\text{fine}}}$, analogously for $A_{\text{rec}} \in \mathbb{C}^{K' \times N_{\text{coarse}}}$. Similarly, we define $y^{(i)} = [y_1^{(i)}|_I, y_2^{(i)}|_I, y_3^{(i)}|_I] \in \mathbb{C}^{K'}$, analogously for $y^{\delta,(i)}, \eta^{(i)} \in \mathbb{C}^{K'}$. For evaluation, the ground truth phantoms are retransformed to image shape, i.e., $x^{(i)} = x_{\text{coarse}}^{(i)} \in \mathbb{R}^{17 \times 15}$. In addition to the selected concentration, we consider the

same setting for concentrations $c = 2, 5, 20, 50 \text{ mg}_{\text{Fe}}/\text{mL}$. This creates five versions of the dataset. Since the noise data remains unchanged for simulating the noise-perturbed measurements $y^{\delta,(i)}$ at all concentrations, the noise level is inherently increased at a decreasing particle concentration.

A. Reconstruction Methods

We employ four different reconstruction approaches on the noise-perturbed MPI measurements for the reconstruction of each of the MNIST phantoms.

In addition to conventional methods introduced in Section II-A, we apply the LDA from Section II-B by minimizing the objective in (10) via gradient descent. In analogy to the standard methods, for the LDA we choose the regularization functional $\mathcal{R}(x) = \frac{\alpha}{2} \|x\|^2$. Furthermore, the initial value for the minimization of (10) is chosen as the reconstruction result obtained from the regularized Kaczmarz algorithm. Therefore, the approach can be seen as a post-processing tool.

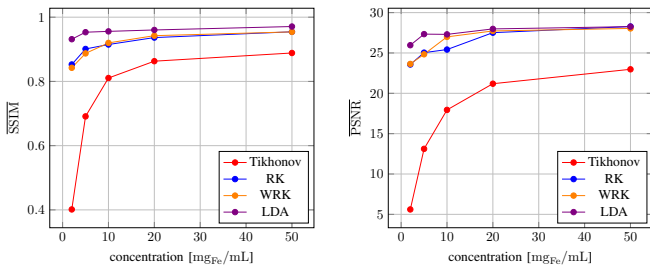
The regularization parameters of all reconstruction methods are selected according to an extensive grid search over the parameter spaces corresponding to the structural similarity index measure (SSIM) as defined in [57]. The grid search is performed separately with respect to each concentration level over the first 100 images of the test dataset for the Tikhonov approach, the regularized Kaczmarz approach and the whitened regularized Kaczmarz approach for a fair comparison and representative reconstruction quality. For the LDA, we apply the grid search to the first five images of the test dataset. All selected regularization parameters are listed in the appendix, Table V.

B. Discrepancy Network

Applying the LDA requires the training of a discrepancy network. To this end, we train an invertible neural network $\varphi_{\theta} : \mathbb{R}^{2 \times K'} \rightarrow \mathbb{R}^{2 \times K'}$ on noise samples $(\eta^{(i)})_i$. The network processes the real and imaginary components of the signal separately, resulting in input and output dimensions of $\mathbb{R}^{2 \times K'}$, where the first dimension describes the real and complex channels, respectively. The noise data from the MPI-MNIST dataset used for training originates from the samples included in the file `large_NoiseMeas.mdf` as described in Section V, where approximately 300 000 samples were used for training the network. The network follows a multi-scale architecture based on the idea to capture information at different scales in order to improve the overall model's ability to represent complex data distributions. The architecture is based on affine coupling layers. A detailed description of the network architecture containing approximately 13.9 million trainable parameters can be found in the appendix. For approximating the true noise distribution, the loss function for the training of our INN is chosen as in (9) and the network weights were optimized using Adam [58] on a batch size $b = 256$ over 25 epochs of training. Note that, according to the usage of identical noise samples for the generation of all five versions of the dataset, the same trained discrepancy network can be utilized at all levels of concentration.

		c = 2	c = 5	c = 10	c = 20	c = 50
$\overline{\text{SSIM}}$	Tikhonov	0.5087 ± 0.092	0.7455 ± 0.082	0.8000 ± 0.0818	0.8538 ± 0.068	0.8887 ± 0.048
	RK	0.8527 ± 0.082	0.9010 ± 0.050	0.9149 ± 0.036	0.9365 ± 0.039	0.9545 ± 0.029
	WRK	0.8424 ± 0.082	0.8872 ± 0.058	0.9199 ± 0.051	0.9424 ± 0.035	0.9542 ± 0.026
	LDA	0.9314 ± 0.039	0.9532 ± 0.027	0.9558 ± 0.023	0.9602 ± 0.022	0.9709 ± 0.017
$\overline{\text{PSNR}}$	Tikhonov	8.633 ± 0.451	15.587 ± 0.657	16.276 ± 1.553	19.406 ± 1.700	22.983 ± 1.779
	RK	23.569 ± 1.129	25.040 ± 1.703	25.422 ± 1.978	27.525 ± 1.469	28.277 ± 1.459
	WRK	23.641 ± 1.092	24.836 ± 1.671	26.995 ± 1.296	27.728 ± 1.527	28.041 ± 1.629
	LDA	25.962 ± 1.312	27.335 ± 1.551	27.305 ± 1.666	27.981 ± 1.596	28.262 ± 1.539

TABLE I: Average errors over particle concentrations (in $\text{mg}_{\text{Fe}}/\text{mL}$) of given reconstruction method on MNIST test set. Higher particle concentration implies higher relative measurement noise.



(a) Mean SSIM values.

(b) Mean PSNR values.

Fig. 4: SSIM (left) and PSNR (right) values of the reconstruction methods evaluated over the test dataset for concentrations $c = 2, 5, 10, 20, 50 \text{ mg}_{\text{Fe}}/\text{mL}$ with the unchanging added noise measurements (i.e., lower concentrations correspond to higher relative noise levels).

C. Results

Figure 5 presents reconstructed images corresponding to selected MPI-MNIST test dataset phantoms, obtained using four distinct reconstruction methods across varying concentrations. The analytical Tikhonov approach generally yields relatively poor reconstruction quality, failing to clearly resolve the phantom structures. In contrast, the other methods demonstrate significant improvements in terms of reconstruction quality. The regularized Kaczmarz algorithm, however, introduces artifacts near the upper left corner of the images and close to the digit in the center of the image, which become increasingly visible as the concentration decreases, coinciding with higher noise levels. The whitened regularized Kaczmarz method mitigates the corner artifacts, so that they become only slightly visible, while the centered artifacts are maintained or even increased. The LDA effectively eliminates these artifacts, leaving only high-frequency background noise noticeable at concentrations $c = 2, 5 \text{ mg}_{\text{Fe}}/\text{mL}$. This behaviour is also reflected by both error measures, which severely differ from the standard methods especially at a low particle concentration. Among all presented methods, LDA consistently delivers superior reconstruction quality, especially at increasing noise level. This aligns with the method's design, where the noise structure learned by the discrepancy network becomes increasingly advantageous in high-noise scenarios.

To investigate the general reconstruction quality on the MPI-MNIST dataset on a larger amount of phantoms, we compute

the mean SSIM and PSNR values

$$\overline{\text{SSIM}} = \frac{1}{L_{\text{test}}} \sum_{i=1}^{L_{\text{test}}} \text{SSIM}(x^{(i)}, \hat{x}^{(i)}),$$

$$\overline{\text{PSNR}} = \frac{1}{L_{\text{test}}} \sum_{i=1}^{L_{\text{test}}} \text{PSNR}(x^{(i)}, \hat{x}^{(i)}),$$

over the test dataset containing $L_{\text{test}} = 10000$ images for different concentrations. The error measures for $c = 2, 5, 10, 20, 50 \text{ mg}_{\text{Fe}}/\text{mL}$ are listed in Table I and show the same effect as observed in Figure 5: The higher the noise level, the more significant the improvement of reconstruction quality of the LDA. The LDA yields the highest mean SSIM values for all concentrations, whereas the regularized Kaczmarz algorithm performs slightly better at $c = 50 \text{ mg}_{\text{Fe}}/\text{mL}$ in terms of the PSNR measure. These quantitative measures are visualized in Figure 4, where the significance of the varying reconstruction quality over the methods at high noise levels becomes visible. All methods have the tendency to converge to a high reconstruction quality at increasing concentration, whereas the LDA shows the steepest increase towards a high reconstruction quality at low particle concentrations. However, the LDA requires a severely higher number of gradient descent iterations for a single reconstruction in comparison to the required iterations per reconstruction with other approaches, which in turn results in an increased computation time. Note that the definition of a single iteration varies across the different reconstruction approaches. While one iteration of the LDA is considered as a single gradient descent step, the RK and WRK use iterations of the Kaczmarz algorithm. We consider the analytical computation of the Tikhonov function minimizer as a single-iteration approach. Despite of varying iteration procedures, the iteration steps provide a consistent quantitative measure for algorithm comparison with potential to indicate computation duration. Table II shows the iteration numbers for image reconstruction of the approaches derived from our grid search together with the corresponding computation times in Table III at $c = 2, 5, 10, 20, 50 \text{ mg}_{\text{Fe}}/\text{mL}$. All computations were performed on a system with dual Intel(R) Xeon(R) Silver 4210 CPU. The computation times were averaged over the reconstruction times of the first 100 images from the MPI-MNIST dataset. Note that we constrained our grid search for the LDA to a maximum number of 2000 iterations, although allowing a higher iteration count for the LDA might lead to an improved reconstruction quality. To prioritize reconstruction

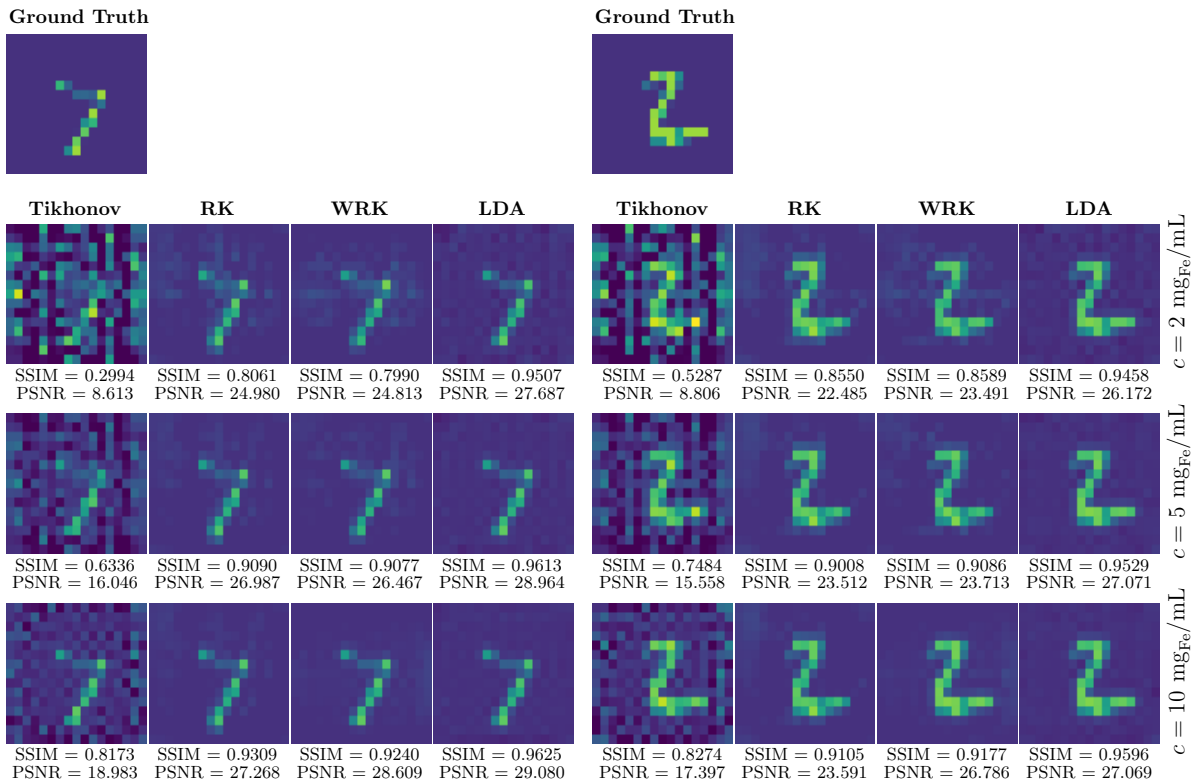


Fig. 5: Reconstructions of selected phantoms of the MPI-MNIST test dataset from different reconstruction methods over concentrations $c = 2, 5, 10$ mg_{Fe}/mL with corresponding error measures SSIM and PSNR compared to the ground truth.

c [mg _{Fe} /mL]	Tikhonov	RK	WRK	LDA
2	1	1	1	2000
5	1	1	1	2000
10	1	1	20	2000
20	1	10	20	2000
50	1	200	20	1000

TABLE II: Number of iterations for a single reconstruction of different reconstruction algorithms and concentration levels averaged over the first 100 samples of the MPI-MNIST test dataset.

c [mg _{Fe} /mL]	Tikhonov	RK	WRK	LDA
2	0.0079	0.0665	0.0691	89.5331
5	0.0078	0.0684	0.0701	82.9576
10	0.0081	0.0512	0.9854	84.8896
20	0.0080	0.4830	0.9655	87.0539
50	0.0070	10.0687	1.0601	42.2373

TABLE III: Time in seconds for a single reconstruction of different reconstruction algorithms and concentration levels averaged over the first 100 samples of the MPI-MNIST test dataset.

time, the number of iterations was limited accordingly.

VII. CONCLUSION

In the present work, we addressed two challenges in MPI image reconstruction: The formulation of a reconstruction algorithm that accurately captures the structure of MPI scanner noise and the scarcity of available MPI datasets.

From an algorithmic perspective, we introduced the LDA, which leverages the features and capabilities of the MPI-MNIST dataset. Unlike conventional Tikhonov regularization and existing deep learning approaches, our approach relaxes the Gaussian noise assumption and integrates the problem-specific noise distribution. The applied MAP estimation within the LDA includes a formulation of the likelihood of the data under the model which comprises a learned version of the true noise distribution by an invertible neural network. Therefore, the method is capable of taking complex noise structures into account and generalizes the Tikhonov objective. This incorporation of problem-specific noise features directly addresses the crucial noise aspect of the MPI modeling chain. The method is generally applicable to linear inverse problems and appears particularly promising for problems characterized by multifaceted or unknown noise distributions. This also concerns applications where assumptions about the noise distribution are made, but could be inaccurate or only approximately satisfied in practice.

Addressing the second challenge, we presented the publicly available MPI-MNIST dataset, which is beneficial for the development and testing of reconstruction methods. The extensive collection of ground truth phantoms and corresponding simulated MPI measurements as well as a substantial number of additional noise samples captured by the preclinical MPI scanner allow for a detailed analysis of complex data structures. The available data scope facilitates the application of learning-based approaches, providing a large number of samples suitable for both noise-specific methods as well as

supervised end-to-end reconstruction frameworks. All components can be accessed individually, allowing for targeted studies such as operator correction or, in particular, system matrix denoising. In addition, the combination of available system matrices and noise samples offers the possibility to generate further data classes in order to analyze desired phantom structures, including the simulation of different noise levels. Finally, the structure and variety of the dataset make it particularly well-suited for benchmarking and comparative evaluation of different reconstruction techniques.

In our numerical experiments we demonstrated the effectiveness of our reconstruction approach by training the discrepancy network on the noise data provided with the MPI-MNIST dataset. Compared to classical reconstruction methods, our learned approach showed significant improvements in reconstruction quality, particularly at higher noise levels, which occurs with a reduction of the particle concentration. However, a current limitation is its extended reconstruction duration in comparison to classical methods. Motivated by [59], we currently apply a warm start to our LDA, initializing it with the reconstruction obtained from the regularized Kaczmarz method, which already reduces the number of required iterations. Additional strategies to decrease computation time could include a reduction of model parameters, thereby shortening the time needed to compute gradients and accelerating each iteration by reducing model capacity. A similar effect can be achieved by carefully selecting the band-pass filter to limit the number of significant frequency components, thus reducing the input dimension. Moreover, a potential trade-off analysis could explore reducing the number of iterations at the expense of slight losses in reconstruction quality, while still achieving superior performance compared to traditional methods.

These findings serve as a basis for several future directions. In addition to the optimization of computation time, it is essential to evaluate the performance of the LDA on real MPI measurements by using the trained model from our numerical experiments. This step is crucial for validating the method's generalizability in practical scenarios. Moreover, the regularization properties of the LDA remain unexplored. Examining the conditions of the discrepancy network, including specific network architecture choices, under which the approach maintains existence, stability, and convergence of a regularization scheme will be an essential focus for future research. In this work, we focus on the discrepancy functional within the Tikhonov objective by addressing the noise distribution. However, the approach could be extended by additionally relaxing the Gaussian assumption on the prior. In such cases, we could introduce an additional network to learn the prior distribution, potentially leading to a more accurate penalty term and bringing the method even closer to a realistic MPI setting.

ACKNOWLEDGEMENTS

The authors would like to thank Florian Thieben and Konrad Scheffler for conducting the experiments. Furthermore, the authors would like to thank Marija Boberg for her useful assistance with data preparation. Meira Iske ac-

knowledges the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 281474342/GRK2224/2. Hannes Albers acknowledges financial support from the KIWi project funded by the German Federal Ministry of Education and Research (BMBF) - Project number 05M22LBA.

REFERENCES

- [1] B. Gleich and J. Weizenecker, "Tomographic imaging using the non-linear response of magnetic particles," *Nature*, vol. 435, pp. 1214–7, 2005.
- [2] T. Knopp, N. Gdaniec, and M. Möddel, "Magnetic particle imaging: From proof of principle to preclinical applications," *Physics in Medicine & Biology*, vol. 62, no. 14, p. R124, 2017.
- [3] T. Buzug, G. Bringout, M. Erbe, K. Gräfe, M. Graeser, M. Grüttner, A. Halkola, T. Sattel, W. Tenner, H. Wojtczyk, J. Haegele, F. Vogt, J. Barkhausen, and K. Lüdtke-Buzug, "Magnetic particle imaging: Introduction to imaging and hardware realization," *Zeitschrift für Medizinische Physik*, vol. 22, no. 4, pp. 323–334, 2012.
- [4] B. Gleich, J. Weizenecker, and J. Borgert, "Experimental results on fast 2D-encoded magnetic particle imaging," *Physics in Medicine & Biology*, vol. 53, no. 6, p. N81, 2008.
- [5] T. Knopp and M. Hofmann, "Online reconstruction of 3D magnetic particle imaging data," *Physics in Medicine & Biology*, vol. 61, no. 11, p. N257, 2016.
- [6] S. Herz, P. Vogel, T. Kampf, M. Rückert, S. Veldhoen, V. Behr, and T. Bley, "Magnetic particle imaging for quantification of vascular stenoses: A phantom study," *IEEE Transactions on Medical Imaging*, vol. 37, no. 1, pp. 61–67, 2018.
- [7] J. Weizenecker, B. Gleich, J. Rahmer, H. Dahnke, and J. Borgert, "Three-dimensional real-time in vivo magnetic particle imaging," *Physics in Medicine & Biology*, vol. 54, no. 5, p. L1, 2009.
- [8] M. G. Kaul, J. Salamon, T. Knopp, H. Itrich, G. Adam, H. Weller, and C. Jung, "Magnetic particle imaging for in vivo blood flow velocity measurements in mice," *Physics in Medicine & Biology*, vol. 63, no. 6, p. 064001, 2018.
- [9] J. Haegele, J. Rahmer, B. Gleich, J. Borgert, H. Wojtczyk, N. Panagiotopoulos, T. Buzug, J. Barkhausen, and F. Vogt, "Magnetic particle imaging: Visualization of instruments for cardiovascular intervention," *Radiology*, vol. 265, no. 3, pp. 933–938, 2012.
- [10] S. Herz, P. Vogel, T. Kampf, P. Dietrich, S. Veldhoen, M. A. Rückert, R. Kickuth, V. Behr, and T. Bley, "Magnetic particle imaging-guided stenting," *Journal of Endovascular Therapy*, vol. 26, no. 4, pp. 512–519, 2019.
- [11] J. Rahmer, D. Wirtz, C. Bontus, J. Borgert, and B. Gleich, "Interactive magnetic catheter steering with 3-D real-time feedback using multi-color magnetic particle imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1449–1456, 2017.
- [12] M. Ahlborg, T. Friedrich, T. Götsche, V. Scheitenberger, R. Linemann, M. Wattenberg, A. T. Buessen, T. Knopp, P. Szwargulski, M. G. Kaul *et al.*, "First dedicated balloon catheter for magnetic particle imaging," *IEEE transactions on medical imaging*, vol. 41, no. 11, pp. 3301–3308, 2022.
- [13] J. Bulte, P. Walczak, M. Janowski, K. Krishnan, H. Arami, A. Halkola, B. Gleich, and J. Rahmer, "Quantitative "hot spot" imaging of transplanted stem cells using superparamagnetic tracers and magnetic particle imaging (MPI)," *Tomography*, vol. 1, no. 2, pp. 91–97, 2015.
- [14] O. Sehl, J. Gevaert, K. Meli, N. Knie, and P. Foster, "A perspective on cell tracking with magnetic particle imaging," *Tomography*, vol. 6, no. 4, pp. 315–324, 2020.
- [15] J. W. Bulte, C. Wang, and A. Shakeri-Zadeh, "In vivo cellular magnetic imaging: Labeled versus unlabeled cells," *Advanced functional materials*, vol. 32, no. 50, p. 2207626, 2022.
- [16] E. Yu, M. Bishop, B. Zheng, R. Ferguson, A. Khandhar, S. Kemp, K. Krishnan, P. Goodwill, and S. Conolly, "Magnetic particle imaging: A novel in vivo imaging platform for cancer detection," *Nano Letters*, vol. 17, no. 3, pp. 1648–1654, 2017.
- [17] L. M. Bauer, S. F. Situ, M. A. Griswold, and A. C. S. Samia, "High-performance iron oxide nanoparticles for magnetic particle imaging-guided hyperthermia (hMPI)," *Nanoscale*, vol. 8, no. 24, pp. 12162–12169, 2016.

- [18] T. Knopp, J. Rahmer, T. Sattel, S. Biederer, J. Weizenecker, B. Gleich, J. Borgert, and T. Buzug, "Weighted iterative reconstruction for magnetic particle imaging," *Physics in Medicine & Biology*, vol. 55, no. 6, p. 1577, 2010.
- [19] T. Knopp, S. Biederer, T. Sattel, and T. Buzug, "Singular value analysis for magnetic particle imaging," in *2008 IEEE Nuclear Science Symposium Conference Record*, 2008, pp. 4525–4529.
- [20] T. Knopp, T. Sattel, S. Biederer, J. Rahmer, J. Weizenecker, B. Gleich, J. Borgert, and T. Buzug, "Model-based reconstruction for magnetic particle imaging," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 12–18, 2010.
- [21] T. Knopp, S. Biederer, T. Sattel, J. Rahmer, J. Weizenecker, B. Gleich, J. Borgert, and T. Buzug, "2D model-based reconstruction for magnetic particle imaging," *Medical Physics*, vol. 37, no. 2, pp. 485–491, 2010.
- [22] T. Yoshida and K. Enpuku, "Nonlinear behavior of magnetic fluid in brownian relaxation: Numerical simulation and derivation of empirical model," in *Magnetic Particle Imaging*, T. M. Buzug and J. Borgert, Eds. Springer Berlin Heidelberg, 2012.
- [23] T. Kluth, "Mathematical models for magnetic particle imaging," *Inverse Problems*, vol. 34, no. 8, p. 083001, 2018.
- [24] T. Kluth, P. Szwargulski, and T. Knopp, "Towards accurate modeling of the multidimensional magnetic particle imaging physics," *New Journal of Physics*, vol. 21, no. 10, p. 103032, 2019.
- [25] H. Albers, T. Knopp, M. Möddel, M. Boberg, and T. Kluth, "Modeling the magnetization dynamics for large ensembles of immobilized magnetic nanoparticles in multi-dimensional magnetic particle imaging," *Journal of Magnetism and Magnetic Materials*, vol. 543, p. 168534, 2022.
- [26] M. Maass, T. Kluth, C. Droigk, H. Albers, K. Scheffler, A. Mertins, and T. Knopp, "Equilibrium model with anisotropy for model-based reconstruction in magnetic particle imaging," *IEEE Transactions on Computational Imaging*, vol. 10, pp. 1588–1601, 2024.
- [27] T. Knopp, N. Gdaniec, R. Rehr, M. Graeser, and T. Gerkmann, "Correction of linear system drifts in magnetic particle imaging," *Physics in Medicine & Biology*, vol. 64, no. 12, p. 125013, 2019.
- [28] K. Them, M. Kaul, C. Jung, M. Hofmann, T. Mummert, F. Werner, and T. Knopp, "Sensitivity enhancement in magnetic particle imaging by background subtraction," *IEEE Transactions on Medical Imaging*, vol. 35, no. 3, pp. 893–900, 2015.
- [29] T. Kluth and B. Jin, "Enhanced reconstruction in magnetic particle imaging by whitening and randomized SVD approximation," *Physics in Medicine & Biology*, vol. 64, no. 12, p. 125026, 2019.
- [30] T. Kluth and P. Maass, "Model uncertainty in magnetic particle imaging: Nonlinear problem formulation and model-based sparse reconstruction," *International Journal of Magnetic Particle Imaging*, vol. 3, no. 2, 2017.
- [31] S. Blanke, B. Hahn, and A. Wald, "Inverse problems with inexact forward operator: Iterative regularization and application in dynamic imaging," *Inverse Problems*, vol. 36, no. 12, p. 124001, 2020.
- [32] M. Nitzsche, H. Albers, B. Hahn, and T. Kluth, "Compensating model imperfections during image reconstruction via RESESOP," *International Journal of Magnetic Particle Imaging*, vol. 8, no. 1, 2022.
- [33] C. Bathke, T. Kluth, C. Brandt, and P. Maass, "Improved image reconstruction in magnetic particle imaging using structural a priori information," *International Journal of Magnetic Particle Imaging*, vol. 3, no. 1, 2017.
- [34] 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, 2017.
- [35] T. Kluth, "L1 data fitting for robust reconstruction in magnetic particle imaging: Quantitative evaluation on Open MPI dataset," *International Journal on Magnetic particle Imaging*, vol. 6, no. 2, 2020.
- [36] T. Knopp, P. Jürß, and M. Grosser, "A deep learning approach for automatic image reconstruction in MPI," *International Journal of Magnetic Particle Imaging*, vol. 9, no. 1, 2023.
- [37] A. von Gladiss, I. Kramer, N. Theisen, R. Memmesheimer, A. Bakenecker, T. Buzug, and D. Paulus, "Data augmentation for training a neural network for image reconstruction in MPI," *International Journal of Magnetic Particle Imaging*, vol. 8, no. 1, 2022.
- [38] A. von Gladiss, R. Memmesheimer, N. Theisen, A. Bakenecker, T. Buzug, and D. Paulus, "Reconstruction of 1D images with a neural network for magnetic particle imaging," in *Bildverarbeitung für die Medizin 2022*, K. Maier-Hein, T. M. Deserno, H. Handels, A. Maier, C. Palm, and T. Tolxdorff, Eds. Springer Fachmedien Wiesbaden, 2022, pp. 247–252.
- [39] P. Koch, M. Maass, M. Bruhns, C. Droigk, T. Parbs, and A. Mertins, "Neural network for reconstruction of MPI images," in *9th International Workshop on Magnetic Particle Imaging (IWMPI)*, 2019.
- [40] B. Askin, A. Güngör, D. Alptekin Soydan, E. Saritas, C. Top, and T. Cukur, "PP-MPI: A deep plug-and-play prior for magnetic particle imaging reconstruction," in *Machine Learning for Medical Image Reconstruction*. Springer International Publishing, 2022, pp. 105–114.
- [41] A. Güngör, B. Askin, D. Soydan, E. Saritas, C. Top, and T. Çukur, "A denoiser scaling technique for plug-and-play MPI reconstruction," in *International Journal of Magnetic Particle Imaging*, vol. 9, 2023.
- [42] A. Tsanda, P. Jürß, N. Hackelberg, M. Grosser, M. Möddel, and T. Knopp, "Extension of the kaczmarz algorithm with a deep plug-and-play regularizer," in *International Journal of Magnetic Particle Imaging*, vol. 10, 2024.
- [43] 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," *IEEE Transactions on Medical Imaging*, vol. 43, no. 1, pp. 321–334, 2023.
- [44] A. Güngör, E. Saritas, and T. Çukur, "A deep equilibrium technique for 3D MPI reconstruction," in *International Journal of Magnetic Particle Imaging*, vol. 10, no. 1, 2024.
- [45] S. Dittmer, T. Kluth, M. Henriksen, and P. Maass, "Deep image prior for 3D magnetic particle imaging: A quantitative comparison of regularization techniques on open MPI dataset," *International Journal on Magnetic Particle Imaging*, vol. 7, no. 1, 2021.
- [46] T. Knopp, P. Szwargulski, F. Griese, and M. Gräser, "OpenMPIData: An initiative for freely accessible magnetic particle imaging data," *Data in Brief*, vol. 28, p. 104971, 2020.
- [47] 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," *Physics in Medicine & Biology*, vol. 67, no. 12, p. 125012, 2022.
- [48] T. Pinetz, E. Kobler, T. Pock, and A. Effland, "Shared prior learning of energy-based models for image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 14, no. 4, pp. 1706–1748, 2021.
- [49] L. Calatroni, C. Cao, J. C. D. los Reyes, C.-B. Schönlieb, and T. Valkonen, *Bilevel approaches for learning of variational imaging models*. Berlin, Boston: De Gruyter, 2017, pp. 252–290.
- [50] Y. LeCun, "The MNIST database of handwritten digits," 1998. [Online]. Available: <http://yann.lecun.com/exdb/mnist/>
- [51] A. Dax, "On row relaxation methods for large constrained least squares problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 3, pp. 570–584, 1993.
- [52] H. Albers, T. Kluth, and T. Knopp, "Simulating magnetization dynamics of large ensembles of single domain nanoparticles: Numerical study of brown/néel dynamics and parameter identification problems in magnetic particle imaging," *Journal of Magnetism and Magnetic Materials*, vol. 541, p. 168508, 2022.
- [53] F. Thieben, H. Albers, F. Mohn, F. Foerger, M. Boberg, K. Scheffler, M. Möddel, T. Kluth, and T. Knopp, "Experimental parameter calibration of the scanner model for model-based MPI," *International Journal on Magnetic Particle Imaging IJMPI*, 2024.
- [54] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*. Springer New York, NY, 2004.
- [55] T. Knopp, T. Viereck, G. Bringout, M. Ahlborg, J. Rahmer, and M. Hofmann, "MDF: Magnetic particle imaging data format," 2020. [Online]. Available: <https://arxiv.org/abs/1602.06072>
- [56] K. Scheffler, M. Boberg, and T. Knopp, "Solving the MPI reconstruction problem with automatically tuned regularization parameters," *Physics in Medicine & Biology*, vol. 69, no. 4, p. 045024, 2024.
- [57] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [58] D. Kingma and P. Dhariwal, "Glow: Generative flow with invertible 1x1 convolutions," in *Advances in neural information processing systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31. Curran Associates, Inc., 2018.
- [59] T. Knopp and M. Grosser, "Warmstart approach for accelerating deep image prior reconstruction in dynamic tomography," in *International Conference on Medical Imaging with Deep Learning*. PMLR, 2022, pp. 713–725.
- [60] L. Dinh, J. Sohl-Dickstein, and S. Bengio, "Density estimation using real NVP," in *ICLR*, 2017.

APPENDIX MODIFICATIONS OF MPI-MNIST

The fully provided dataset described in Section V contains preselected parameters such as a fixed particle concentration, reconstruction resolution, as well as a data-generating system matrix. However, together with the provided data, this framework can be customized.

Particle Concentration

For changing the particle concentration to a concentration \tilde{c} , all steps of the preprocessing in Section V-A can be applied to the phantoms, whereas the pixel values are scaled by the desired concentration \tilde{c} , so that $x_{j,k} \in [0, \tilde{c}]$ for all $j = 1, \dots, W_{\text{coarse}}$ and $k = 1, \dots, H_{\text{coarse}}$. The measurements y can either be scaled by the factor \tilde{c}/c or y and y^δ can be computed as in I. and II., respectively. By changing the concentration, the noise level of y^δ will inherently change – a higher tracer concentration will lead to a larger signal and thus a higher SNR when the noise is kept fixed.

Resolution and physical model

The resolution and underlying physical model can be adapted both for the data generation and data reconstruction. In both cases, the system matrix is selected as A_p^r for $r = \text{coarse}, \text{int}, \text{fine}$ and $p = \text{fluid_opt}, \text{fluid_small_params}, \text{fluid_large_params}, \text{equilibrium}$ from the provided system matrix files listed in the previous section. In order to obtain either the corresponding ground truth or data generation phantoms x_r, \bar{x}_r of the desired resolution r , the following steps are required:

- (i) $r = \text{fine}$: Apply preprocessing as in Section V-A to obtain $x_{\text{fine}}^{(i)}$ and $\bar{x}_{\text{fine}}^{(i)}$.
- (ii) $r = \text{int}$: Change the upsampling factor s in Section V-A to $s = 3$ so that $x_{\text{int}}^{(i)} \in \mathbb{R}^{51 \times 45}$. Subsequently flatten analogously to obtain $\bar{x}_{\text{int}}^{(i)} \in \mathbb{R}^{N_{\text{int}}}$.
- (iii) $r = \text{coarse}$: Set $\bar{x}_r^{(i)} = \bar{x}_{\text{coarse}}^{(i)}$ and obtain $x_{\text{coarse}}^{(i)} \in \mathbb{R}^{17 \times 15}$ by reversing the matrix to vector transformation described in Section V-A.

For the data generation process, compute $y^{(i)} = A_p^r \bar{x}_r^{(i)}$ and $y^{\delta, (i)} = y^{(i)} + \eta^{(i)}$, where $\eta^{(i)}$ results from H_{phantom}^D .

MODEL ARCHITECTURE

The model architecture used in Section VI follows a multi-scale architecture inspired by [60] that comprises affine coupling layers with shape transforming layers. The real part η^{real} and imaginary part η^{imag} of our network input were considered channelwise, so that $\eta = [\eta^{\text{real}} \mid \eta^{\text{imag}}] \in \mathbb{R}^{p \times K'}$ with $p = 2$ channels. We define by $\eta_j := [\eta_j^{\text{real}} \mid \eta_j^{\text{imag}}]$ the j -th component of each channel with $j = 0, \dots, K' - 1$. By treating each channel separately, an affine coupling layer $f_{\text{ac}} : \mathbb{R}^{p \times K'} \rightarrow \mathbb{R}^{p \times K'}$ is defined by

$$(f_{\text{ac}}(\eta))_j = \begin{cases} \eta_j, & \text{if } j \in I_{d, K'}, \\ \eta_j \cdot \exp(s((\eta_j)_{j \in I_{d, K'}})) \\ \quad + t((\eta_j)_{j \in I_{d, K'}}), & \text{if } j \in I_{d, K'}^C, \end{cases}$$

for $0 < d \leq K' - 1$ and an index set $I_{d, K'} \subseteq \{j : 0 \leq j \leq K' - 1\}$ with $|I_{d, K'}| = d$ and $I_{d, K'}^C = I_{K'} \setminus I_{d, K'}$. The functions $s, t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ are neural networks, referred to as *scaling* and *shifting networks*. The inverse of $y = f_{\text{ac}}(\eta)$ can be computed by

$$\eta_j = \begin{cases} y_j, & \text{if } j \in I_{d, K'}, \\ (y_j - t((y_j)_{j \in I_{d, K'}})) \\ \quad \cdot \exp(-s((y_j)_{j \in I_{d, K'}})), & \text{if } j \in I_{d, K'}^C, \end{cases}$$

without the necessity to compute the inverse of s, t . Its Jacobian is a lower triangular matrix, with a tractable Jacobian determinant. To operate on different scales, we incorporate squeeze layers, split layers and concatenation layers in addition to the coupling layers. These are responsible for transforming the data shapes. The squeeze layers $f_{\text{squeeze}} : \mathbb{R}^{p \times K'} \rightarrow \mathbb{R}^{2p \times K'/2}$ compress the frequency dimensions while expanding the number of channels, transforming the input w by

$$(f_{\text{squeeze}}(w))_{j,k} = w_{\lceil j/p \rceil}, \quad k \bmod p,$$

for $j = 0, \dots, 2p - 1, k = 0, \dots, K'/2 - 1$, as illustrated in Figure 6. Split layers $f_{\text{split}} : \mathbb{R}^{p \times K'} \rightarrow \mathbb{R}^{p/2 \times K'} \times \mathbb{R}^{p/2 \times K'}$ with $f_{\text{split}}(w) = (w_{0:p/2-1, 0:K'-1}, w_{p/2:p-1, 0:K'-1})$, on the other hand, partition an input along its channel dimension. This enables separate processing pathways that are later recombined to restore the original dimensionality (see Figure 6). The concatenation layer $f_{\text{concat}} : \mathbb{R}^{p/2 \times K'} \times \mathbb{R}^{p/2 \times K'} \rightarrow \mathbb{R}^{p \times K'}$, used as an output layer of the network, is responsible for this recombination by concatenating two inputs along the channel dimension, where

$$(f_{\text{concat}}(v, w))_{j,k} = \begin{cases} v_{j,k} & \text{if } j \leq p/2 - 1 \\ w_{j,k} & \text{otherwise.} \end{cases}$$

Figure 7 contains the detailed overall layer architecture, which extends to a third dimension referring to the batch size b . Table IV shows the chosen identical architectures of the shifting and the scaling networks within each affine coupling layer $\ell = 1, \dots, 8$. The width of the fully connected layers depends on the scale of the coupling block with

$$W_\ell = \begin{cases} 512 & \text{for } \ell = 1, 2 \\ 256 & \text{for } \ell = 3, 4 \\ 128 & \text{for } \ell = 5, \dots, 8. \end{cases}$$

This network design allows for efficient transformation of the input across multiple scales, with the affine coupling layers ensuring invertibility and computational efficiency of the network. For further reading on the layer types we refer to [60].

In our setting, the index set $I_{d, K'}$ was implemented by defining $I_{d, K'} := \{j : j \bmod 2 = 0, 0 \leq j \leq K' - 1\}$ with $d = \lfloor K'/2 \rfloor$, therefore separating into odd and even entries, where the entries of the index set $I_{d, K'}^C$, containing odd indices are passed through the shifting and scaling network of the coupling blocks. The definition of the squeeze and split operations within our network leads to a separation of real and complex components within the split layer. Therefore, our network processes the complex part of the noise signal

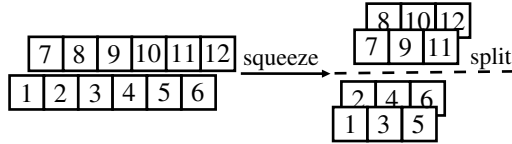


Fig. 6: Illustration of a squeeze and a split layer. A squeeze transformation, indicated by the right arrow, transforms an input of shape $p \times K'$ to $2p \times K'/2$. The dashed line indicates a split transformation on the output of the squeeze transformation. This results in two outputs of shape $p \times K'/2$.

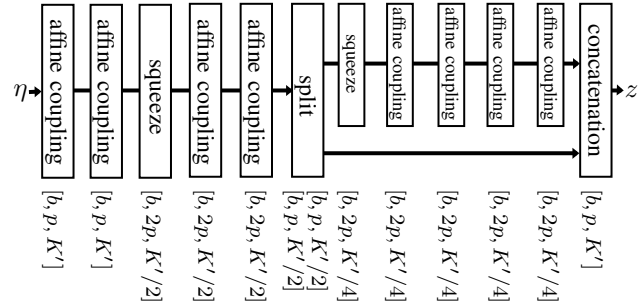


Fig. 7: Network architecture of our multi-scale discrepancy network used in Section VI with the layer types and corresponding output dimensions for each layer.

Block	Layer Type	Width
1	Fully Connected Layer	W_ℓ
	Layer Normalization	
	ReLU Activation	
⋮	⋮	⋮
6	Fully Connected Layer	W_ℓ
	Layer Normalization	
	ReLU Activation	
7	Fully Connected Layer	W_ℓ
	Tanh Activation	

TABLE IV: Network architecture of the scaling and shifting networks s, t of each affine coupling block $\ell = 1, \dots, 8$ utilized in Figure 7.

		$c = 2$	$c = 5$	$c = 10$	$c = 20$	$c = 50$
Tikhonov	α	$6.14 \cdot 10^{-11}$	$6.14 \cdot 10^{-11}$	$4.33 \cdot 10^{-12}$	$4.33 \cdot 10^{-12}$	$2.55 \cdot 10^{-11}$
	iter	1	1	1	10	200
WRK	α	$4.52 \cdot 10^{-3}$	$4.52 \cdot 10^{-4}$	$1.08 \cdot 10^{-3}$	$1.08 \cdot 10^{-3}$	$8.53 \cdot 10^{-4}$
	iter	1	1	20	20	20
LDA	α	50	10	10	10	50000
	iter	2000	2000	2000	2000	1000
	lr	$1.39 \cdot 10^{-2}$	$3.73 \cdot 10^{-2}$	$1.00 \cdot 10^{-1}$	$1.39 \cdot 10^{-2}$	$7.20 \cdot 10^{-4}$

TABLE V: Regularization parameters for the reconstruction approaches at different concentration levels used for the experiments in Section VI. The parameter α corresponds to the regularization weight in (2), with iter denoting the number of iterations used in each method. The table also lists the learning rate lr applied in each gradient descent step of the LDA.

on a single scale, whereas the real part is processed on two different scales before both signal parts are concatenated in the output layer. Compared to a single-scale network of comparable size we observed a faster convergence behaviour using this architecture, terminating at a number of 25 epochs of training at $b = 256$. Moreover, we found an improvement of reconstruction results with respect to both error measures considered in Section VI when using the multi-scale architecture in Figure 7.