

Bilevel Optimized Implicit Neural Representation for Scan-Specific Accelerated MRI Reconstruction

Hongze Yu¹, Jeffrey A. Fessler¹, *Fellow, IEEE*, and Yun Jiang¹

Abstract—Deep learning (DL) methods can reconstruct highly accelerated magnetic resonance imaging (MRI) scans, but they rely on application-specific large training datasets and often generalize poorly to out-of-distribution data. Self-supervised deep learning algorithms perform scan-specific reconstructions, but still require complicated hyperparameter tuning based on the acquisition and often offer limited acceleration. This work develops a bilevel-optimized implicit neural representation (INR) approach for scan-specific MRI reconstruction. The method explicitly formulates the undersampled MRI reconstruction problem as a bilevel optimization problem and automatically optimizes the multidimensional hyperparameters of the reconstruction method for a given acquisition protocol, enabling a tailored reconstruction without training data. The proposed algorithm uses Gaussian process regression to optimize INR hyperparameters, accommodating various acquisitions. The INR includes a trainable positional encoder for high-dimensional feature embedding and a small multi-layer perceptron for decoding. The bilevel optimization is computationally efficient, requiring only a few minutes per typical 2D Cartesian scan. On the scanner hardware, the subsequent scan-specific reconstruction—using offline-optimized hyperparameters—is completed within seconds, while achieving comparable or improved image quality compared to previous model-based and self-supervised learning methods.

Index Terms—Implicit neural representation, Bayesian optimization, self-supervised deep learning, undersampled MRI reconstruction, hyperparameter optimization.

Received 20 December 2025; revised 25 February 2026 and 4 April 2026; accepted 20 April 2026. Date of publication 22 April 2026; date of current version 9 July 2026. This work was supported in part by NIH under Grant R37CA263583 and Grant R01CA284172 and in part by the Siemens Healthineers and Cook Medical. Recommended by Associate Editor G. Wang. (*Corresponding author: Yun Jiang.*)

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Michigan Medical School Institutional Review Board (IRBMED) under Application No. HUM00175599. The study was HIPAA-compliant, and informed written consent was obtained from each participant.

Hongze Yu and Jeffrey A. Fessler are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: hongze@umich.edu; fessler@umich.edu).

Yun Jiang is with the Department of Radiology and the Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: yunjiang@umich.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TMI.2026.3686724>, provided by the authors.

Digital Object Identifier 10.1109/TMI.2026.3686724

I. INTRODUCTION

MAGNETIC resonance imaging (MRI) is a diagnostic imaging technique that provides excellent soft-tissue contrast without ionizing radiation. However, its prolonged acquisition times can pose a challenge in clinical practice. To address this drawback, a common strategy is to develop reconstruction algorithms that can recover artifact-free images from sub-Nyquist sampled data. Parallel imaging techniques [1], [2] exploit redundancy among multiple receiver coils. Classical methods, however, are constrained by the number and geometry of the coils, leading to potential residual artifacts and noise amplification at high acceleration rates. Compressed sensing [3], [4] and related model-based methods [5], [6], [7] further mitigate aliasing artifacts by enforcing sparsity or low-rank constraints and can be combined with parallel imaging to achieve higher acceleration rates in dynamic and quantitative MRI [8], [9], [10], [11].

Supervised deep learning-based methods can offer efficient reconstruction once they are trained. Typically, neural networks are trained using a large fully-sampled dataset [12]. Common approaches include training the neural network to learn a direct mapping from undersampled images to fully-sampled images [13], [14], or employing the network as a regularizer in model-based iterative reconstruction [15], [16]. However, these approaches require curated, fully-sampled training datasets [17] that are often unavailable for many MRI acquisitions, such as dynamic or diffusion-weighted MR imaging. Moreover, these pre-trained models often struggle to generalize effectively when applied to out-of-distribution data [18].

Various semi- and self-supervised deep learning-based methods have been proposed to overcome the need for fully-sampled training data in supervised methods, focusing on scan-specific reconstruction. One category exploits the shift-invariant autoregressive structure of k-space. Algorithms, such as RAKI, LORAKI, and SPARK [19], [20], [21], use convolutional or recurrent networks trained on fully sampled center k-space lines to predict missing samples. Other approaches enforce data consistency by perturbing or splitting the measurement. For example, Noise2Recon [22] enforces consistency between reconstructions of the original undersampled data and their noise-augmented counterparts, whereas data-splitting approaches [23], [24] achieve

self-supervision by further partitioning the undersampled measurement into disjoint training and validation sets. A different line of work relies on implicit network architectural regularization, as in DIP [25] and DNLIN [26].

Recent advances in novel-view synthesis and volume rendering [27] offer a new approach to reconstruct undersampled MRI data. Implicit neural representation (INR) models target signals as continuous coordinate-based functions, reformulating the voxel-based reconstruction into optimization of the weights of a compact neural representation. In novel-view synthesis, INR learns from multiple projection views to reconstruct a continuous target scene. Similarly, in MRI, INR can leverage the spatially varying coil sensitivity to reconstruct alias-free images from undersampled MRI data. Previous works using INR have achieved patient-specific image reconstruction by employing fully-sampled patient history images as priors [28] or adding explicit, hand-crafted image-domain regularizers such as total variation and low rank [29], [30], [31].

However, like other self-supervised reconstruction techniques, INR-based methods are sensitive to the choice of hyperparameters. Because the final reconstructed image parameterized by INR is a joint function of both network weights and hyperparameters, suboptimal settings can noticeably degrade image quality. Consequently, these methods typically require dedicated hyperparameter tuning, such as grid or random searches, tailored to the anatomy, contrast, and sampling pattern. In practice, such tuning often relies on population-based, fully-sampled reference data. To reduce this dependence, various automatic regularization parameter selection methods have been proposed, including the discrepancy principle [32], [33], the L-curve [34], [35], [36], generalized cross-validation (GCV) [37], [38], [39], [40], and *Stein's unbiased risk estimate* (SURE) [40], [41], [42], [43]. However, these methods primarily optimize the regularization strength for a fixed model and do not account for the joint optimization of the regularization, model capacity, and optimization-related hyperparameters that govern the final INR-based reconstruction.

In this work, we propose a *bilevel-optimized INR* for hyperparameter-optimized, scan-specific MRI reconstruction. We introduce a weight-domain regularized INR that achieves high-quality reconstructions without requiring external hand-crafted regularizers, by systematically designing the key components, such as the multiresolution positional encoder, the decoder multilayer perceptron (MLP), the loss function weighting, and the weight-domain regularization. Inspired by prior automatic parameter selection methods [40], [43], we explicitly formulate reconstruction as a bilevel optimization problem [44], where the lower level is a scan-specific multiresolution INR network and the upper level handles multidimensional hyperparameter optimization over regularization, multiresolution scaling, and optimization-related hyperparameters, solved using Bayesian optimization [45]. We divide the undersampled single-scan k-space data into mutually exclusive training and validation sets to enable self-supervision [23], [24]. For each candidate hyperparameter

vector, the INR is trained on the training set, and the validation loss is computed using unseen data.

We adopt Bayesian optimization for its derivative-free characteristic and computational efficiency, which make it suitable for this global optimization task. The proposed bilevel-optimized INR can adapt hyperparameters to various acquisition protocols, such as anatomy, sampling pattern, and contrast. We also demonstrate hyperparameter transferability: settings learned on one scan remain effective for other subjects acquired with the same protocol, showing only modest degradation. Nevertheless, a direct scan-specific bilevel optimization still outperforms both these transferred settings and bilevel optimization using population data from the same protocol, highlighting the value of per-scan tuning.

Our results demonstrate that the proposed method achieved comparable or improved image quality relative to state-of-the-art self-supervised deep-learning and model-based reconstruction methods. With offline-optimized hyperparameters, our approach can achieve scan-specific MRI reconstruction in less than 5 seconds for a 2D Cartesian scan with 16 coils with matrix size 384×384 on a single NVIDIA A40 GPU.

We also validate hyperparameter transferability across the same anatomies using dynamic images acquired in a real-time interventional MRI (iMRI) during a simulated biopsy procedure with residual learning. Specifically, Our INR method, hyperparameter-optimized on the same anatomy and trained on the fully-sampled first frame of an MR-guided biopsy scan, reconstructs subsequent undersampled frames with minimal fine-tuning (under 180 milliseconds per frame with a matrix size of 128×128). This residual learning strategy potentially enables real-time interventional MRI reconstruction with highly accelerated acquisition ($6\times$).

This work makes the following contributions:

- We introduce an efficient, scan-specific bilevel optimization framework for MRI reconstruction. This framework uses k-space data splitting within a single undersampled scan to define an outer validation loss and jointly optimize multidimensional hyperparameters, including regularization strengths, multiresolution scaling, and optimization parameters in nonlinear reconstruction models.
- We design a weight-regularized, hash-encoded INR with a self-weighted k-space loss. Through systematic ablations, we demonstrate that, with appropriately optimized hyperparameters, INR can function as a stand-alone continuous parametric reconstruction model for highly accelerated MRI without relying on additional handcrafted priors.
- We empirically characterize the differences between scan-specific and population-based, or cross-subject hyperparameter optimization, demonstrating consistent improvements from per-scan bilevel tuning, and practically useful transferability of optimized hyperparameters across subjects and anatomies acquired using the same protocol.

The remaining materials are organized as follows. Sections II and III detail the theoretical background and the proposed method. Section IV describes experiment settings. Sections V and VI present results and discussion.

II. BACKGROUND

A. MR Image Reconstruction

We denote scalars by italic letters, e.g., t , and vectors and matrices by bold letters, e.g., $\theta, \mathbf{A}$. Spatial coordinates are denoted by letters with arrows, e.g., $\vec{r}$.

In an MRI system with multiple receiver coils, the signal from the l th coil can be modeled as

$$s_l(t) = \int c_l(\vec{r}) f(\vec{r}) e^{-i2\pi \vec{k}(t) \cdot \vec{r}} d\vec{r}, \quad (1)$$

where $c_l(\vec{r})$ is the coil sensitivity, $f(\vec{r})$ denotes the finite-support continuous function that represents magnetization, and $\vec{k}(t)$ is the sampling location in k-space at time t . During the acquisition, $s_l(t)$ is sampled at discrete time points, resulting in the measurements $y_l \in \mathbb{C}^M$. Let $\mathbf{x} \in \mathbb{C}^N$ represent the discretized form of f on a Cartesian grid $\mathbf{z} \in \mathbb{R}^{N \times d}$ of spatial sampling locations, where each row of $\mathbf{z}$ stores the coordinate of one spatial location $\vec{r}$. By stacking the measurements from all L coils, we represent the multi-coil forward model [46] as:

$$\mathbf{y} \triangleq \begin{bmatrix} y_1 \\ \vdots \\ y_L \end{bmatrix} = \underbrace{(\mathbf{I}_L \otimes \mathbf{F}) \mathbf{C}}_{\mathbf{A}} \mathbf{x} + \boldsymbol{\epsilon}, \quad \mathbf{C} \triangleq \begin{bmatrix} \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_L \end{bmatrix}. \quad (2)$$

Here, $\mathbf{y} \in \mathbb{C}^{LM}$ is the stacked measurement vector, $\mathbf{F} \in \mathbb{C}^{M \times N}$ represents the Fourier encoding operator, $\mathbf{C}_l \in \mathbb{C}^{N \times N}$ represents the sensitivity of the l th coil, and $\boldsymbol{\epsilon} \in \mathbb{C}^{LM}$ denotes measurement noise. The Kronecker product with identity matrix $\mathbf{I}_L$ duplicates the Fourier encoding into the multi-coil acquisition.

When the image is fully-sampled ($M \geq N$), the model can be fully determined with proper coil arrangement. However, to accelerate MRI acquisitions, fewer k-space samples ($M < N$) are acquired, leading to an underdetermined forward operator $\mathbf{A}$. Consequently, the naive least-squares solution is non-unique and introduces undersampling artifacts in images, necessitating a regularizer $\mathcal{R}(\mathbf{x})$ in the image reconstruction problem that can be formulated as:

$$\mathbf{x}^* \triangleq \arg \min_{\mathbf{x}} \frac{1}{LM} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}). \quad (3)$$

Common choices for $\mathcal{R}(\cdot)$ include sparsity- or low-rank-based constraints [3], [47].

B. Implicit Neural Representation

Instead of solving the discrete pixel-based optimization problem in (3), classical continuous parametric models (e.g., Fourier or spline finite-series expansions [43], [48]) parameterize the image as a continuous function using a predetermined fixed basis, in which the image is expressed as a linear weighted sum of basis functions such as sinusoids or piecewise-polynomials. Estimating these expansion coefficients aligns naturally with the continuous physics of MRI magnetization, and the resulting function class is mathematically well-characterized.

An INR offers a different approach to continuous image parameterization. A positional encoder maps spatial coordinates into a higher-dimensional feature space, with the encoding either predetermined (e.g., frequency encoding [27] and Fourier feature encoding [49]) or task-adaptive with

additional optimizable parameters (e.g., multiresolution grid-based encoders [50], [51]). An MLP then combines these features nonlinearly, rather than as a linear weighted sum. In the task-adaptive case, the effective basis is optimized with respect to a specific target jointly with the MLP weights. Therefore, instead of using a neural network as an explicit regularizer [15], an INR models the target image $f_\theta(\vec{r})$ as a continuous function of $\vec{r}$, parameterized by network weights θ , including both the positional encoder weights θ_{Enc} and decoder MLP weights θ_{MLP} . In an INR-based approach, the optimization in (3) is reformulated as an optimization over the INR weights θ .

Let $f_\theta(\cdot) : \mathbb{R}^d \rightarrow \mathbb{C}$ denote the continuous function representation that maps a spatial coordinate $\vec{r}$ to its complex image value, and let $f_\theta(\mathbf{z}) \in \mathbb{C}^N$ denote its evaluation on the Cartesian grid $\mathbf{z} \in \mathbb{R}^{N \times d}$ introduced above. A basic (unregularized) INR-based reconstruction problem can then be formulated as:

$$\mathbf{x}_\theta = f_\theta(\mathbf{z}), \quad \hat{\theta} = \arg \min_{\theta} \frac{1}{LM} \|\mathbf{y} - \mathbf{A}f_\theta(\mathbf{z})\|_2^2, \quad (4)$$

$$f_\theta(\mathbf{z}) = M_{\theta_{\text{MLP}}}(\gamma_{\theta_{\text{Enc}}}(\mathbf{z})), \quad (5)$$

where $\hat{\theta}$ denotes the optimized network parameters under a selected set of hyperparameters, $M_{\theta_{\text{MLP}}}$ represents the MLP, and $\gamma_{\theta_{\text{Enc}}}$ is the positional encoding function applied pointwise to each spatial location in $\mathbf{z}$ (subscript “.” indicates pointwise application). The “training” process during the reconstruction time updates the network weights θ iteratively to minimize the data-fidelity term in (4), while the small MLP implicitly promotes continuity and constrains the reconstructed image to its function class. Note that the formulation in (4)–(5) represents purely scan-specific nonlinear function fitting using a continuous parametric representation, rather than a generalizable population-trained learning model. While we adopt standard deep learning terminology (e.g., “learning”, “training”, “self-supervised”) to maintain consistency with prior INR literature, we emphasize that “learning” here refers to optimizing the network weights θ to fit a single undersampled scan, without generalizing across subjects or acquisition protocols.

Compared with classical finite-series expansion models, INR does not have an established theoretical advantage for accelerated MRI reconstruction. Both provide continuous image models and permit straightforward differentiation through the MRI forward model. Finite-series expansions are more interpretable, as their basis functions and approximation space are analytically defined, and are generally more parameter-efficient. By contrast, the effective basis of an INR is implicitly determined by the encoder, decoder, regularization, and nonconvex optimization, making its approximation space less transparent and harder to characterize theoretically.

INR offers a possible advantage because its task-adaptive multiresolution encoder can yield a scan-adaptive multiscale representation with optimizable, data-dependent parameters, rather than relying on a predetermined basis. Coarsely encoded features can represent global image structure, while finer features can capture localized detail, and the decoder combines these features nonlinearly. For piecewise-smooth MR images, an INR with this flexibility may better capture anatomical boundaries than a fixed smooth basis such as cubic B-splines. Such flexibility comes at the cost of more unknown parameters

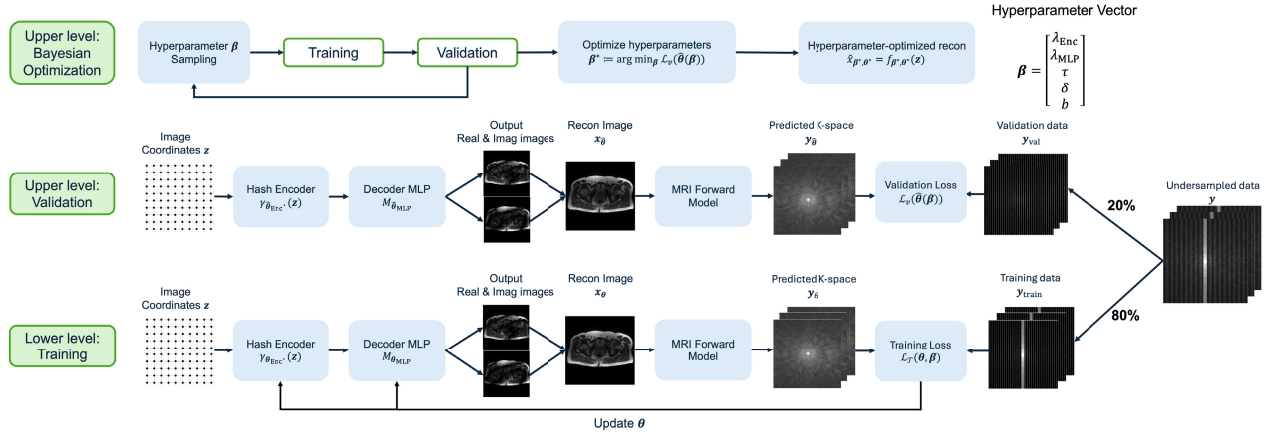


Fig. 1. Proposed bilevel-optimized INR framework. The undersampled data y is split into training and validation sets. In each upper-level Bayesian optimization iteration, a hyperparameter vector β is selected and used to train the INR on y_{train} . The trained INR is then validated on y_{val} using a weighted ℓ_2 loss. The final INR reconstruction and hyperparameter set are chosen based on the smallest validation loss.

and potential identifiability issue in ill-conditioned inverse problems. Accordingly, existing INR-based MRI methods often include handcrafted image-domain regularizers [29], [30], [31]. Such handcrafted regularizers could also be applied to single- or multi-resolution finite-series expansions. In this work, we instead use a simple Tikhonov regularizer of the weights of the representation. (See Section V-F, Fig. 13, and [52] and empirical comparison.) We emphasize that this discussion aims to provide intuition, rather than a general theoretical claim that INR outperforms finite-series models.

III. METHODOLOGY

A. Overall Framework

Fig. 1 shows the proposed bilevel optimized INR framework for MR image reconstruction, which has two levels: (1) lower-level INR-based image reconstruction, and (2) upper-level hyperparameter optimization using Bayesian optimization [45] with data split from undersampled k-space [23].

We formulate this bilevel optimization problem as:

$$\begin{aligned} \beta^* &\triangleq \arg \min_{\beta} \mathcal{L}_{\mathcal{V}}(\hat{\theta}(\beta)) \\ \text{subject to } \hat{\theta}(\beta) &\triangleq \arg \min_{\theta} \mathcal{L}_{\mathcal{T}}(\beta, \theta), \end{aligned} \quad (6)$$

where

$$\mathcal{L}_{\mathcal{V}}(\hat{\theta}(\beta)) \triangleq \left\| \mathbf{W}(\hat{\theta}(\beta), \delta_0) (\mathbf{y}_{\text{val}} - \mathbf{A}_{\mathcal{V}} f_{\hat{\theta}(\beta)}(\mathbf{z})) \right\|_2^2, \quad (7a)$$

$$\begin{aligned} \mathcal{L}_{\mathcal{T}}(\beta, \theta) &\triangleq \frac{1}{|\mathcal{T}|} \left\| \mathbf{W}(\theta, \delta) (\mathbf{y}_{\text{train}} - \mathbf{A}_{\mathcal{T}} f_{\theta}(\mathbf{z})) \right\|_2^2 \\ &\quad + \lambda_{\text{Enc}} \|\boldsymbol{\theta}_{\text{Enc}}\|_2^2 + \lambda_{\text{MLP}} \|\boldsymbol{\theta}_{\text{MLP}}\|_2^2. \end{aligned} \quad (7b)$$

Here, $\mathcal{L}_{\mathcal{T}}(\beta, \theta)$ denotes the lower-level training loss and $\mathcal{L}_{\mathcal{V}}(\hat{\theta}(\beta))$ is the upper-level validation loss. $\mathbf{W}(\theta, \delta) \in \mathbb{C}^{LM \times LM}$ represents a k-space weighting operator, emphasizing higher frequency components. β is a hyperparameter vector, including ℓ_2 regularization strengths λ_{Enc} and λ_{MLP} , the learning rate τ , the loss-weight bias δ , and the cross-level resolution scale factor b of the hash encoder. For the upper-level validation loss, we analogously use $\mathbf{W}(\hat{\theta}(\beta), \delta_0)$ to emphasize high-frequency residuals, where $\hat{\theta}(\beta)$ is the final estimate obtained from the

lower-level optimization for the current hyperparameter choice β , and δ_0 is a fixed numerical-stability constant (not included in β).

Following [23], we randomly split the undersampled k-space data into a training set $\mathcal{T}$ and a validation set $\mathcal{V}$ (e.g., an 80%/20% split). Let $\mathcal{K}_{\mathcal{T}}$ and $\mathcal{K}_{\mathcal{V}}$ denote the mutually exclusive sets of k-space sampling locations used for training and validation, respectively. The forward operators $\mathbf{A}_{\mathcal{T}}$ and $\mathbf{A}_{\mathcal{V}}$ apply the multi-coil Fourier encoding in (2) restricted to the sampling locations in $\mathcal{K}_{\mathcal{T}}$ and $\mathcal{K}_{\mathcal{V}}$, respectively. In each upper-level iteration, the algorithm selects a hyperparameter vector β and trains an INR by iteratively passing the vectorized image coordinates into a multi-resolution hash encoder $\gamma_{\theta_{\text{Enc}}}(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^{KF}$ that encodes each spatial location into a higher-dimensional feature space. We normalize the lower-level loss by $1/|\mathcal{T}|$ so that the optimized regularization strengths do not depend on the training-set size and can be reused when reconstructing from the full undersampled data.

An MLP $M_{\theta_{\text{MLP}}}(\cdot) : \mathbb{R}^{KF} \rightarrow \mathbb{R}^2$ decodes these features into the real and imaginary parts of the target image $\mathbf{x}_{\theta}$. The reconstructed image is then passed through the MR forward model to produce $\mathbf{y}_{\theta}$. The parameters of the encoder and decoder are updated by calculating a weighted ℓ_2 loss between $\mathbf{y}_{\theta}$ and the acquired undersampled k-space $\mathbf{y}_{\text{train}}$. Once training converges, the validation loss compares the predicted k-space with the unseen measurements $\mathbf{y}_{\text{val}}$. After a predefined number of iterations in the upper-level Bayesian optimization, the method selects the hyperparameters β^* and the associated reconstruction $\hat{\mathbf{x}}_{\beta^*, \theta^*}$ with the lowest validation loss.

B. INR for MRI Reconstruction

For a given hyperparameter vector β , the proposed INR approach for MRI reconstruction extends (4) as follows:

$$\begin{aligned} \hat{\theta}(\beta) &= \arg \min_{\theta} \frac{1}{LM} \left\| \mathbf{W}(\theta, \delta) (\mathbf{y} - \mathbf{A}_{\mathcal{T}} f_{\theta}(\mathbf{z})) \right\|_2^2 \\ &\quad + \lambda_{\text{Enc}} \|\boldsymbol{\theta}_{\text{Enc}}\|_2^2 + \lambda_{\text{MLP}} \|\boldsymbol{\theta}_{\text{MLP}}\|_2^2, \end{aligned} \quad (8)$$

$$\hat{\mathbf{x}}(\beta) = f_{\hat{\theta}(\beta)}(\mathbf{z}). \quad (9)$$

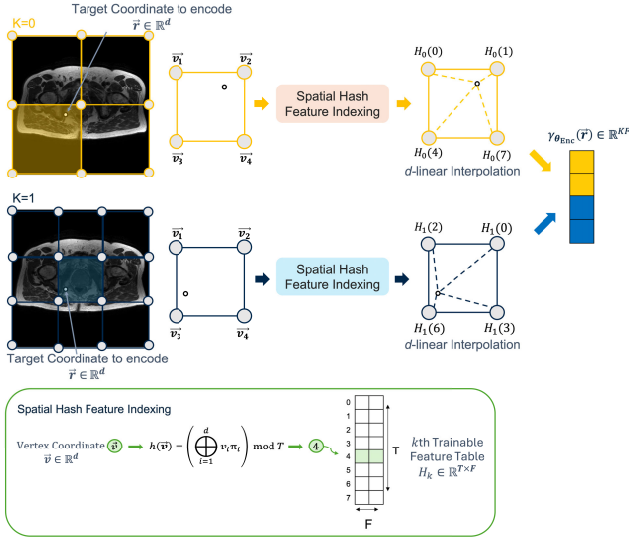


Fig. 2. Multiresolution hash grid encoder $\gamma_{\theta_{\text{Enc}}}(\vec{r})$ for $K=2$ and $d=2$.

Here we treat the optimized INR weights θ as a function of the selected hyperparameters β .

The INR architecture involves three key components: a positional encoder, a decoder MLP, and the loss function. We outline each component below.

1) Positional Encoder: Positional encoding maps each spatial coordinate $\vec{r} \in \mathbb{R}^d$ to a higher-dimensional latent feature vector $\gamma_{\theta_{\text{Enc}}}(\vec{r}) \in \mathbb{R}^{KF}$, so that a compact decoder MLP can represent the high-frequency image content. We adopt a multiresolution hash grid encoder [51] that stores trainable feature vectors of dimension F on a cascade of K Cartesian grids with increasing resolution and accesses them through pre-allocated tables. This multi-scale representation encodes both global and local features of the target image. Intuitively, each grid level supplies a learnable multi-scale basis: coarse levels model smooth, global structure, while finer levels add local detail.

As illustrated in Fig. 2, grid k ($0 \leq k < K$) has resolution $a b^k$, where a is the base resolution and $b \in (1, 2]$ is the cross-level resolution scale factor. For any query point $\vec{r}$, the encoder locates the 2^d vertices of the enclosing patch on each grid, retrieves their feature vectors, and combines them by d -linear interpolation.

Each level is associated with a trainable feature table H_k . For the coarse resolution levels, where the total number of vertices $(\lfloor a b^k \rfloor)^d$ is smaller than the maximum table length T , every vertex maps to a unique entry of H_k like dense grid encoding [50]. For the finer levels, a vertex coordinate $\vec{v}$ is hashed into H_k by

$$h(\vec{v}) = \left(\bigoplus_{i=1}^d v_i \pi_i \right) \bmod T, \quad (10)$$

where π_i are large prime numbers and $\oplus$ denotes bit-wise XOR. Because only the finest grids collide, the coarser levels remain injective and supply a collision-free low-resolution scaffold, while collisions at finer levels are pseudo-random and infrequent for any given pair of points across different resolution levels. In practice, samples that generate larger

loss gradients (e.g., points on high-signal structures) dominate the shared updates, so these collisions act as an implicit average-based regularizer during optimization [51]. Stacking the interpolated vectors from all K grids yields the encoding $\gamma_{\theta_{\text{Enc}}}(\vec{r})$. All feature vectors are initialized with small random weights.

During back-propagation only the $K \times 2^d$ vertex features adjacent to each query are updated (64 vectors when $d=2$ and $K=16$), whereas a dense MLP-only decoder architecture would update its entire weight matrix. The hash grid encoder therefore provides adaptive resolution, constant memory, and sparse parameter updates. We optimize the scale factor b in the outer bilevel loop, since it controls the number of learnable parameters per level and effectively the model capacity, enabling a tailored scan-specific multi-resolution feature representation.

2) Multilayer Perceptron Decoder: The MLP functions as a decoder that maps the positionally encoded features to the target image intensities. Its learnable function class is determined by its activation function, width and depth. In this work, we use an 8-layer MLP with 64 neurons per layer and ReLU activation. The network outputs two channels corresponding to the real and imaginary parts of the complex image, avoiding issues including phase periodicity, magnitude-coupled gradient and sign ambiguity that are associated with representing the image in terms of magnitude and phase.

3) Loss Function: Due to a large magnitude difference between the lower-frequency components at the center and the higher-frequency components at the periphery of k -space, we propose a modified self-weighting loss that weights each network-predicted k -space sample by its magnitude, rather than treating real and imaginary channels separately [31], as is commonly done with RGB channels in recovering images [53]. The loss function weighting is defined as:

$$\mathbf{W}(\theta, \delta) = \text{Diag}(|A f_{\theta}(z)| + \delta \mathbf{1})^{-1}, \quad (11)$$

where here the absolute value is applied element-wise. This self-weighted loss emphasizes higher frequency components and attenuates lower ones, thus avoiding an overly smooth reconstruction. It can also be interpreted as a learnable density compensation function that evolves over iterations. Although no explicit regularizer is applied, the loss function includes ℓ_2 regularization terms for the encoder and MLP weights. Their strengths are controlled by λ_{Enc} and λ_{MLP} , respectively.

C. Bayesian Optimization

The upper-level hyperparameter optimization in (6) is computationally expensive because each evaluation requires training a new INR. Using gradient-based methods would be challenging due to the computational complexity, potential optimization difficulties, and the presence of discrete hyperparameters. To address these issues, we adopt a zeroth-order method called Bayesian optimization [45] that solves this hyperparameter optimization by building a Gaussian process (GP) regression model over the objective function $f(\beta)$. After each evaluation of β (i.e., training an INR and measuring the validation loss), the GP posterior $p(f(\beta) | \mathcal{D})$ is updated, where $\mathcal{D}$ denotes the accumulated data from all previous evaluations.

This surrogate model yields both a predicted mean $\mu(\beta)$ and uncertainty $\sigma(\beta)$. To select the next β , we use the upper confidence bound (UCB) acquisition function

$$\alpha_{\text{UCB}}(\beta) = \mu(\beta) + \kappa \sigma(\beta), \quad (12)$$

where κ balances the trade-off between exploration (high- σ regions) versus exploitation (low predicted loss). By iteratively updating the GP and evaluating α_{UCB} , this method efficiently converges toward near-optimal hyperparameters with far fewer total network trainings compared to grid or random searches.

D. Implementation Details

The proposed bilevel optimized INR tunes five hyperparameters in the upper-level problem: the learning rate τ , the loss weighting bias δ , the weight decay terms λ_{Enc} and λ_{MLP} for the encoder and MLP, and the resolution scale b for adjacent hash grid. We used 60 upper-level iterations, each consisting of 4000 lower-level iterations, to ensure convergence of the INR reconstruction. All image coordinates were normalized to (0, 1) in each dimension.

The hash encoder included 16 levels with a maximum of $2^{18} - 2^{21}$ parameters per level for 2D image reconstruction based on the image matrix size, and base resolution $a = 16$. The decoder MLP had eight hidden layers, each with 64 ReLU neurons, while the input and output layers used no activation. For reconstructing an image with a matrix size of 256×256 , the hash-encoded INR has ~ 2.2 million parameters in the encoder and 27 thousand parameters in the MLP.

Coil sensitivity maps were estimated using E-SPiRiT [54] from a separate gradient-echo scan or from fully-sampled center k-space lines. Virtual coil compression [55] and noise prewhitening [56] were applied before the reconstruction. Nonuniform fast Fourier Transforms [57] for non-Cartesian acquisitions were implemented via the `torchkbnufft` and `cufinufft` toolboxes [58], [59]. For the compressed sensing methods used in the comparison, we solved the ℓ_1 -wavelet using the proximal optimized gradient method (POGM) [60], and the ℓ_1 -total-variation (TV) using the forward-backward primal-dual algorithm [61], both implemented based on the MIRTorch toolbox [62].

The study was implemented in PyTorch 1.12.1 (Python 3.9) on a high-performance computing cluster using an NVIDIA A40 GPU and a 2.9 GHz Intel Xeon Gold 6226R CPU with 16 GB of memory. The source code is available at https://github.com/MIITT-MRI-Jianglab/bilevel_optimized_INR

IV. EXPERIMENTS

A. Experiments

1) *Comparison With Prior Art*: We compared the image quality of bilevel optimized INR with several other model-based and zero-shot self-supervised reconstruction methods, such as CG-SENSE [63], ℓ_1 -wavelet [3], SENSE-LORAKS [64], IMJENSE [29], and ZS-SSL [24]. Comparisons were performed using data acquired across various anatomies, sampling patterns, and field strengths. Specifically, we used datasets acquired with balanced SSFP (bSSFP) or turbo spin echo (TSE) sequences using either a Cartesian or Poisson undersampling pattern. Additionally, we evaluated reconstruction results

TABLE I
ACQUISITION PROTOCOLS AND PARAMETERS

Imaging Type	Coil No.	Sequence	Matrix Size	Resolution (mm ² / mm ³)
Brain (3T)	16	bSSFP / T2-weighted TSE	256 × 256	1 × 1
Cardiac (0.55T)	18	bSSFP	225 × 160	1.6 × 1.6
Prostate (3T)	12	T2-weighted TSE	384 × 384 × 32	1 × 1 × 3
iMRI (0.55T)	24	bSSFP	26 × 128 × 128	2.3 × 2.3
fastMRI Brain (3T)	20	T2-weighted TSE	768 × 396	0.6 × 0.6

for cardiac imaging using a bSSFP sequence on a 0.55T scanner. To further demonstrate the capabilities of our technique with non-Cartesian sampling, we simulated a spiral acquisition from a fully-sampled T2-weighted TSE prostate dataset at 3T. The spiral trajectory, requiring 48 interleaves to fully sample k-space, was designed using minimum-time gradient design [65] with zeroth moment compensation. A comparison using public data was also performed using axial T2-weighted brain scans from the fastMRI data [12]. Table I summarizes scan protocols and acquisition parameters.

For fairness, the regularization strengths of the model-based baselines (ℓ_1 -wavelet, SENSE-LORAKS) and the SENSE-LORAKS rank parameter were selected using the same bilevel optimization framework employed for the proposed INR. For SENSE-LORAKS, we used the public LORAKS V2.1 implementation with S-type matrices and a neighborhood radius of 3. For the ℓ_1 -wavelet method, we used an orthogonal Daubechies-4 wavelet with a single global regularization parameter for all subbands. IMJENSE and ZS-SSL were run using their own recommended settings for fastMRI brain data.

2) *Ablation Study of INR Design*: We performed experiments to determine the most effective formulation of INR for MRI reconstruction and to validate the choices in Section III-B. We examined the impact of different decoders, including different numbers of MLP layers and the use of a linear layer, as well as different positional encoders, such as no encoder, frequency encoding [27], dense grid encoding [50], and hash encoding [51]. For the loss function comparison, we tested several weighting strategies, such as identity weighting, density compensation function weighting [66], acquired data weighting, and self-weighting as described in (11). We also conducted experiments on activation function, comparing ReLU with sinusoidal activation [67]. Additionally, we assessed the effect of weight domain regularization by reconstructing images both with and without the ℓ_2 penalties on the encoder and MLP weights in (9). This comparison isolates the specific contribution of weight regularization to noise suppression and artifact reduction.

3) *Hyperparameter Optimization*: We first examined three hypotheses on hyperparameter optimization in bilevel-optimized INR for image reconstruction. 1. Comparable Image Quality to an Oracle. We hypothesized that our bilevel optimization approach would yield image quality comparable to those of an oracle that selects near-optimal hyperparameters using fully-sampled data for validation loss. To evaluate this hypothesis, we compared our reconstructions against this oracle standard using identical datasets and acquisition parameters. 2. Protocol-Specific Optimization. We hypothesized that the optimal hyperparameters are protocol specific. To validate this, we applied hyperparameters optimized for bSSFP to T2-weighted TSE data from the same subject, assessing

performance consistency, and compared with scan-specifically optimized reconstructions for each protocol. 3. Transferability Across Subjects/Anatomies. We hypothesized that hyperparameters optimized for one subject or anatomy under the same protocol can be transferred to another with modest performance loss. To test this, we first conducted bSSFP acquisitions on two volunteers, comparing reconstructions using each subject's optimized hyperparameters as well as those optimized for the other subject. We then extended this evaluation to multislice prostate T2-weighted imaging, performing a similar comparison across different slices.

Additionally, we conducted experiments comparing population-based and scan-specific bilevel optimization. We selected 30 axial T2-weighted TSE brain subjects from the fastMRI dataset [12]. Slices from 20 subjects with similar anatomical positions were used to derive population hyperparameters, $\hat{\beta}_{\text{pop}}$, using the oracle loss (i.e., the fully-sampled upper-level validation loss). These population hyperparameters were then applied to INR reconstructions of similar slices from the remaining 10 subjects. Separately, each of these 10 slices was reconstructed with scan-specific bilevel optimization, using the undersampled k-space for the upper-level loss. To ensure a fair comparison, we monitored the lower-level optimization for population-based reconstructions and reported the best observed metrics based on NRMSE, i.e., with early stopping. For the scan-specific setting, such early stopping was unnecessary because the metrics typically plateau, so we report the values at the final iteration.

4) *Real-Time Interventional MRI*: We validated the proposed algorithm's capability for real-time residual learning and hyperparameter transferability for the same anatomy using images acquired during a simulated liver biopsy procedure using a phantom. We used the fully-sampled, first frame to optimize the hyperparameters and pre-train the INR. We then retrospectively undersampled the rest time frames $6\times$ with 4% fully-sampled center k-space lines and fine-tuned the INR for 50 iterations for each frame. The temporal resolution for the interventional MRI acquisition is 870 milliseconds per frame when fully sampled.

5) *Function Fitting Comparison With B-Spline*: To motivate the use of INRs and compare them against classical finite-series models, which also serve as continuous parametric representations, we conducted a preliminary 2D function fitting experiment comparing a cubic B-spline model against the hash-encoded INR. To isolate representation difference without handcrafted image-domain regularizers, we applied only Tikhonov regularization to the model weights. The target was a 2D rect function defined on the domain $[0, 3] \times [0, 3]$ and sampled at $N = 10000$ locations drawn uniformly at random. Both models were evaluated on a dense 501×501 Cartesian grid covering $[-0.3, 3.3] \times [-0.3, 3.3]$.

For the classical baseline, we used a standard tensor-product cubic B-spline model with uniformly spaced knots in each spatial dimension. The corresponding design matrix was constructed by evaluating the basis functions at the sample locations, and the spline coefficients were determined by solving a Tikhonov-regularized least-squares problem using Cholesky decomposition. To ensure a fair comparison, we performed an oracle grid search over the number of knots

per dimension $M \in \{5, 10, 15, \dots, 100\}$ and the regularization strength $\log_{10} \lambda \in [-5, 5]$ with step size 0.1, assuming access to the ground truth on the dense evaluation grid. For the INR, we used the same hash-encoded architecture as in our MRI experiments and split the same random samples into 80%/20% training and validation sets for bilevel optimization. The INR weights were trained by minimizing a mean-squared-error loss with Tikhonov penalties applied on the encoder and MLP weights. The hyperparameters $\beta = (\lambda_{\text{Enc}}, \lambda_{\text{MLP}}, \tau, b)$ were selected using our proposed bilevel optimization framework.

B. Performance Measurement

Performance in this study was mainly quantified using three metrics: normalized root-mean-squared error (NRMSE), structural similarity index (SSIM), and peak signal-to-noise Ratio (PSNR). Resolution analysis was performed using the Fourier error radial spectrum plot [68]. Reconstructions were compared against the ground truth from fully-sampled data in both retrospective and simulated studies. For all metric calculations, the coil map estimated by E-SPiRiT was used as the region-of-interest (ROI) mask. Code implementation was based on the scikit-image package.¹

To assess sensitivity to small image features, we also computed local perturbation responses (LPRs) [69]. Checkerboard perturbations (10×10 pixels) with amplitudes of 0.5% and 5% of the maximum reference image magnitude were added to the undersampled k-space. For ZS-SSL, IMJENSE, and the proposed method, separate models were trained on unperturbed and perturbed data. The hyperparameters were re-optimized by bilevel learning for the proposed method.

V. RESULTS

A. Comparisons With Other Self-Supervised Reconstruction Methods

Fig. 3 compares our proposed approach with other model-based iterative and scan-specific INR methods.

For $8\times$ undersampled Cartesian bSSFP brain experiments (Fig. 3(a)), CG-SENSE introduced notable noise across the brain, and ℓ_1 -wavelet, SENSE-LORAKS, and IMJENSE exhibited residual artifacts. In contrast, the proposed method showed minor residual artifacts and reduced the NRMSE by at least 0.07 lower compared to the other approaches.

Fig. 3(b) presents results from experiments using $20\times$ undersampled Poisson-disc sampling pattern with a central 10×10 pixel fully-sampled region. Both images reconstructed using CG-SENSE and SENSE-LORAKS showed visible, unstructured noise. In contrast, ℓ_1 -wavelet, IMJENSE, and the proposed bilevel INR successfully suppressed noise in the reconstructions, with the proposed method outperforming the others as supported by lowest NRMSE of 0.04 and tying for highest SSIM of 0.99 with ℓ_1 -wavelet.

Fig. 3(c) compares images reconstructed by various methods for a single frame of cardiac cine imaging acquired with bSSFP at 0.55T. Visually, CG-SENSE and SENSE-LORAKS retained noise throughout the heart, whereas ℓ_1 -wavelet, IMJENSE, and our method introduced slight smoothing in

¹<https://scikit-image.org/>

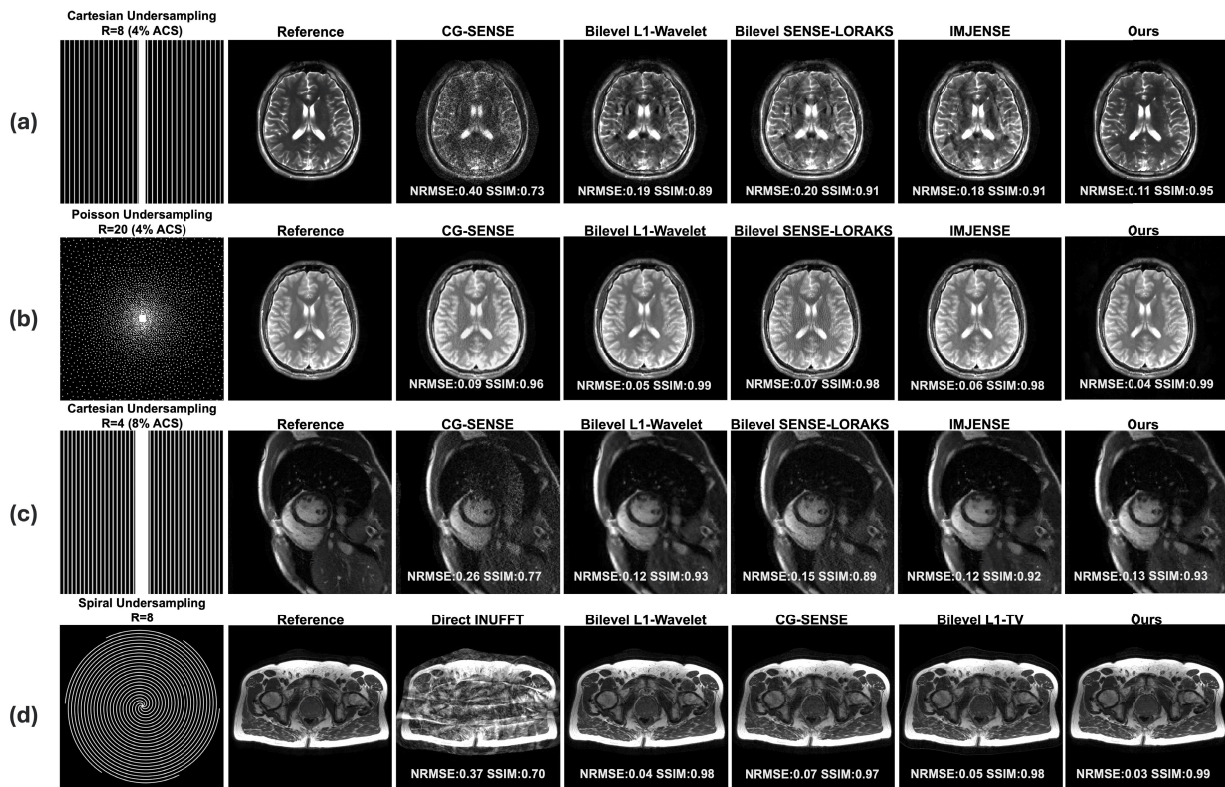


Fig. 3. Method comparisons. (a) Brain volunteer at 3T using bSSFP of $1 \times 1 \text{ mm}^2$ resolution. (b) Brain volunteer at 3T using T2-weighted TSE of $1 \times 1 \text{ mm}^2$ resolution. (c) Cardiac volunteer at 0.55T using bSSFP of $1.6 \times 1.6 \text{ mm}^2$ resolution. (d) Prostate volunteer at 3T using T2-weighted TSE of $1 \times 1 \text{ mm}^2$ resolution. Detailed acquisition protocols are shown in Table I.

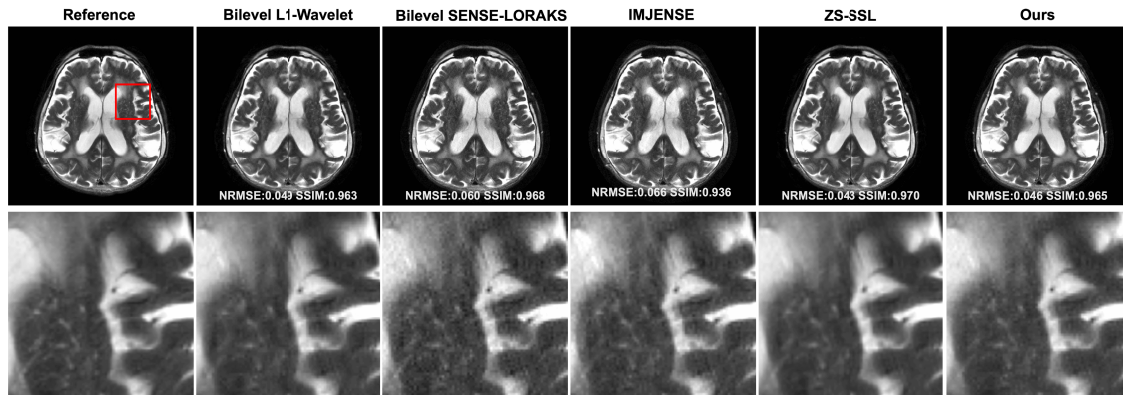


Fig. 4. Comparison of reconstructions on $4 \times$ Cartesian undersampled with 6% fully-sampled center k-space lines using fastMRI Axial T2-weighted brain dataset.

the left ventricle. Quantitatively, ℓ_1 -wavelet and IMJENSE achieved the lowest NRMSE of 0.12; our approach was only 0.01 higher while attaining the highest SSIM of 0.93, matching ℓ_1 -wavelet.

Finally, for the $8 \times$ undersampled spiral acquisition (Fig. 3(d)), direct inverse NUFFT and CG-SENSE reconstruction showed undersampling artifacts. The proposed method achieved better metrics and retained the least amount of residual artifacts, leading the second best method, bilevel-optimized ℓ_1 -wavelet, by 0.01 in NRMSE and 0.01 in SSIM. IMJENSE was omitted here because it only supports Cartesian sampling patterns.

Figs. 4–6 further compares images reconstructed by the bilevel optimized INR with ones of bilevel optimized model-based methods and two scan-specific deep-learning approaches: IMJENSE and ZS-SSL.

Fig. 4 shows results of a $4 \times$ undersampled Cartesian, T2-weighted TSE brain dataset. ZS-SSL achieved the best quantitative metrics (NRMSE/SSIM = 0.043/0.970), with the proposed INR a close second (0.046/0.965). Bilevel optimized ℓ_1 -wavelet ranked third (0.049/0.963) and exhibited blur in the white-matter region shown in the zoomed-in view. The proposed INR and bilevel SENSE-LORAKS (0.060/0.968) appeared sharper in this region, but

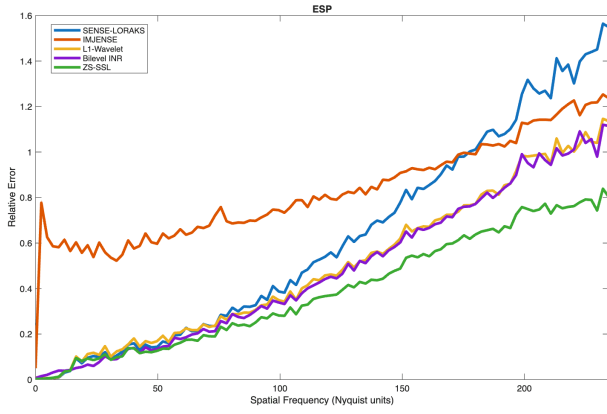


Fig. 5. Comparison of the Fourier radial error spectrum plot (ESP) on $4\times$ Cartesian undersampled brain dataset in Fig. 4.

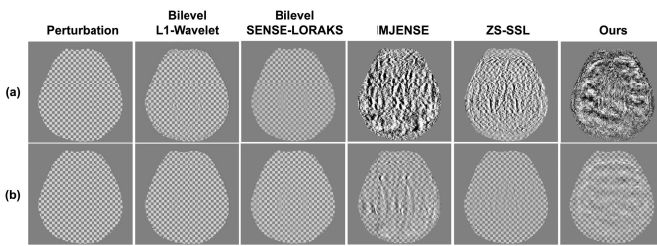


Fig. 6. Comparison of Local Perturbation Responses (LPRs) using the same $4\times$ Cartesian undersampled brain dataset in Fig. 4. Checkerboard perturbations with amplitudes of (a) 0.5% and (b) 5% relative to the maximum reference image magnitude were added to the k-space. Difference images are windowed to $\pm 1.5\times$ the perturbation amplitude (real part). Note that ZS-SSL, IMJENSE, and the proposed method were retrained on the perturbed data. For the proposed method, hyperparameters were also re-optimized by the bilevel framework.

SENSE-LORAKS had worse NRMSE and more visible aliasing artifacts. IMJENSE recorded the highest NRMSE and the lowest SSIM (0.066/0.936).

Fig. 5 presents a frequency-resolved view of the reconstruction error. ZS-SSL showed the lowest relative error over most of the spatial-frequency range, except at the lowest frequencies, where bilevel ℓ_1 -wavelet and the INR were slightly better. These two methods remained close in the high-frequency band. Bilevel SENSE-LORAKS had the largest high-frequency error, matching its residual aliasing, whereas IMJENSE showed the largest error at low frequencies.

Fig. 6 compares LPRs across methods. At the 0.5% perturbation level, all three scan-specific methods (ZS-SSL, IMJENSE, and the proposed method) largely suppressed the checkerboard pattern, with the two INR-based methods showing stronger suppression. When the perturbation was raised to 5%, IMJENSE still exhibited strong residual aliasing, while ZS-SSL showed blurring with central aliasing and the proposed bilevel INR showed mild signal leakage. In contrast, bilevel SENSE-LORAKS preserved the perturbation at both magnitudes, and bilevel ℓ_1 -wavelet retained the pattern with slight blurring near residual aliasing regions.

For the reconstruction time comparison in Fig. 4, all methods were run with their hyperparameters already optimized or tuned. With the same image of matrix size 768×396 , SENSE-LORAKS (CPU implementation) completed in 89 seconds,

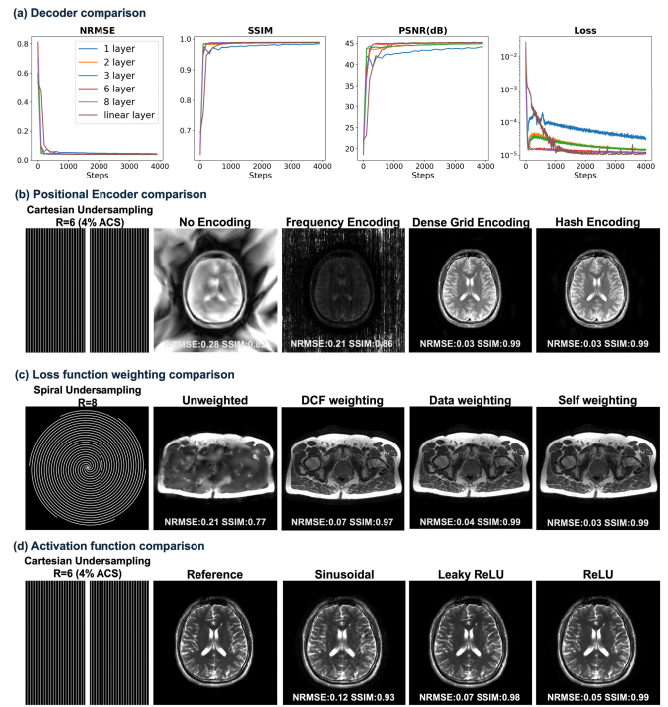


Fig. 7. Ablation study. (a) Decoder comparison with fixed Hash Encoder (i.e., Linear layer vs $1/2/3/6/8$ layer MLP). (b) Encoder comparison with fixed Decoder MLP. (c) Loss function weighting comparison. (d) Activation function comparison.

ℓ_1 -wavelet (GPU) in 4.4 seconds, IMJENSE in 6.8 seconds, the proposed bilevel-optimized INR in 9.7 seconds, and ZS-SSL required 133.3 minutes. Across the experiments in Figs. 3 and 4, the proposed INR achieved either the lowest NRMSE and/or highest SSIM, or its NRMSE and SSIM were within 0.01 of the best competing method, while maintaining reconstruction times less than 10 seconds for 2D scans.

B. Methods Ablation Studies

Fig. 7 presents ablation experiments on key components of the bilevel-optimized INR. Fig. 7(a) compares decoders of varying depths against a linear layer under identical hyperparameters. Except for the single-layer MLP, all decoders converged to similar performance based on quantitative metrics. Even a linear layer can decode the positionally encoded features effectively, indicating the encoder's dominant role in the reconstruction pipeline.

Fig. 7(b) evaluates different positional encoders on a $6\times$ undersampled Cartesian T2-weighted TSE image. Without encoding or with frequency encoding, the model struggled to capture fine details in the brain, whereas dense grid and hash grid encodings yielded comparable metrics. However, dense grid encoding requires $41\times$ more memory and $15\times$ longer training time compared to hash grid encoding, when $b = 1.50$ and $T = 2^{18}$. Fig. 7(c) focuses on the loss-function weighting, evaluated on a prostate image undersampled by a spiral trajectory. The proposed self-weighting scheme in (11) had the best reconstruction quality, capturing both contrast and structural details. In contrast, no loss weighting resulted in visible blurriness in the image. Using Pipe's density

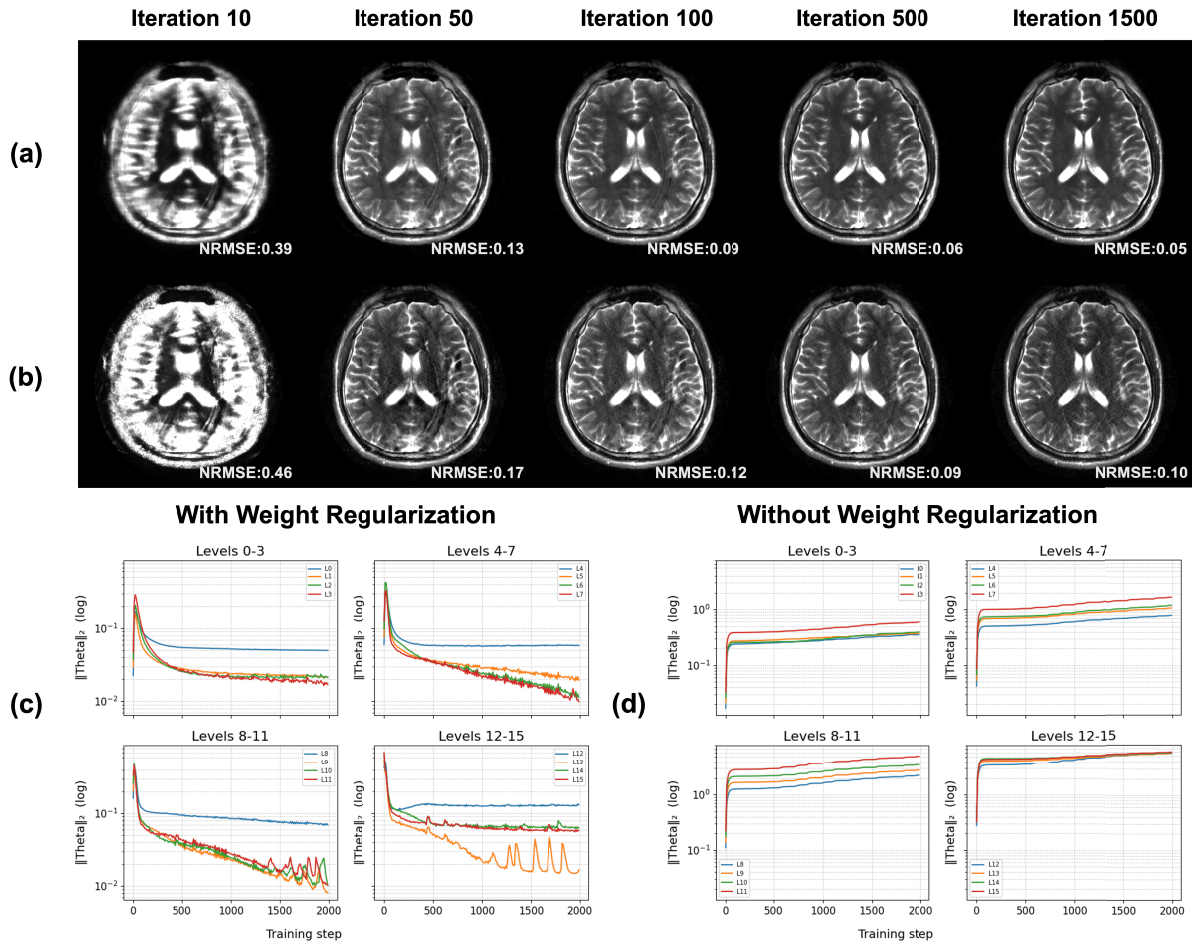


Fig. 8. Training history of the proposed INR on a bSSFP brain volunteer ($6\times$ Cartesian undersampled). (a) Reconstruction *with* weight regularization; (b) reconstruction *without* weight regularization; (c) per-level ℓ_2 -norm of hash encoder weights *with* regularization; (d) per-level ℓ_2 -norm *without* regularization.

compensation function [66] led to less aliased reconstruction, achieving 0.07 NRMSE and 0.97 SSIM compared to 0.03 and 0.99 with our approach. Substituting the acquired data for the predicted k-space in the proposed self-weighting led to slightly worsened performance, increasing NRMSE by 0.01.

Fig. 7(d) investigates the impact of the MLP activation function on reconstruction performance. ReLU activation outperformed both Sinusoidal [67] and LeakyReLU in both NRMSE and SSIM. Sinusoidal activation exhibited the worst performance with visible speckle noise across the whole brain and residual aliasing streaks in white-matter regions.

Fig. 8 evaluates the influence of weight domain regularization by comparing training histories with and without the Tikhonov terms applied to the INR weights.

Fig. 8(a)–(b) show the images reconstructed during the course of training. Without the penalty, the INR gradually reduced aliasing in the white-matter regions but speckle noise emerged after ~ 1500 iterations, and the NRMSE rose after ~ 500 iterations. With the proposed regularization, aliasing was removed while noise remained controlled, and the NRMSE dropped to 0.05 versus 0.10 without regularization.

The weight norm history in Fig. 8(c)–(d) supports this behavior. Without regularization, the per-level ℓ_2 -norms

increased steeply during the first 50 iterations and continued to grow, indicating overfitting. In contrast, with the penalty applied, the norms stabilized quickly, demonstrating that the model capacity was effectively constrained. The regularization strengths λ_{Enc} and λ_{MLP} are automatically selected by the bilevel optimization, ensuring scan-specific optimal performance.

C. Hyperparameter Optimization and Transferability Experiments

Fig. 9 illustrates the results of hyperparameter optimization. Fig. 9(a) and (b) demonstrate that Bayesian optimization achieved nearly optimal NRMSE and SSIM compared to the oracle optimization for both $6\times$ and $8\times$ undersampled reconstruction on volunteer 1. In contrast, empirical tuning led to residual artifacts in the images. The empirically tuned hyperparameters were selected based on previously successful reconstructions for T2-weighted TSE sequences and the same resolutions on brain. Table II lists the hyperparameter values for each case. Except for the loss-weighting bias δ , all other hyperparameters differed between $6\times$ and $8\times$ undersampled reconstruction, indicating that each sampling pattern benefits from tailored optimization.

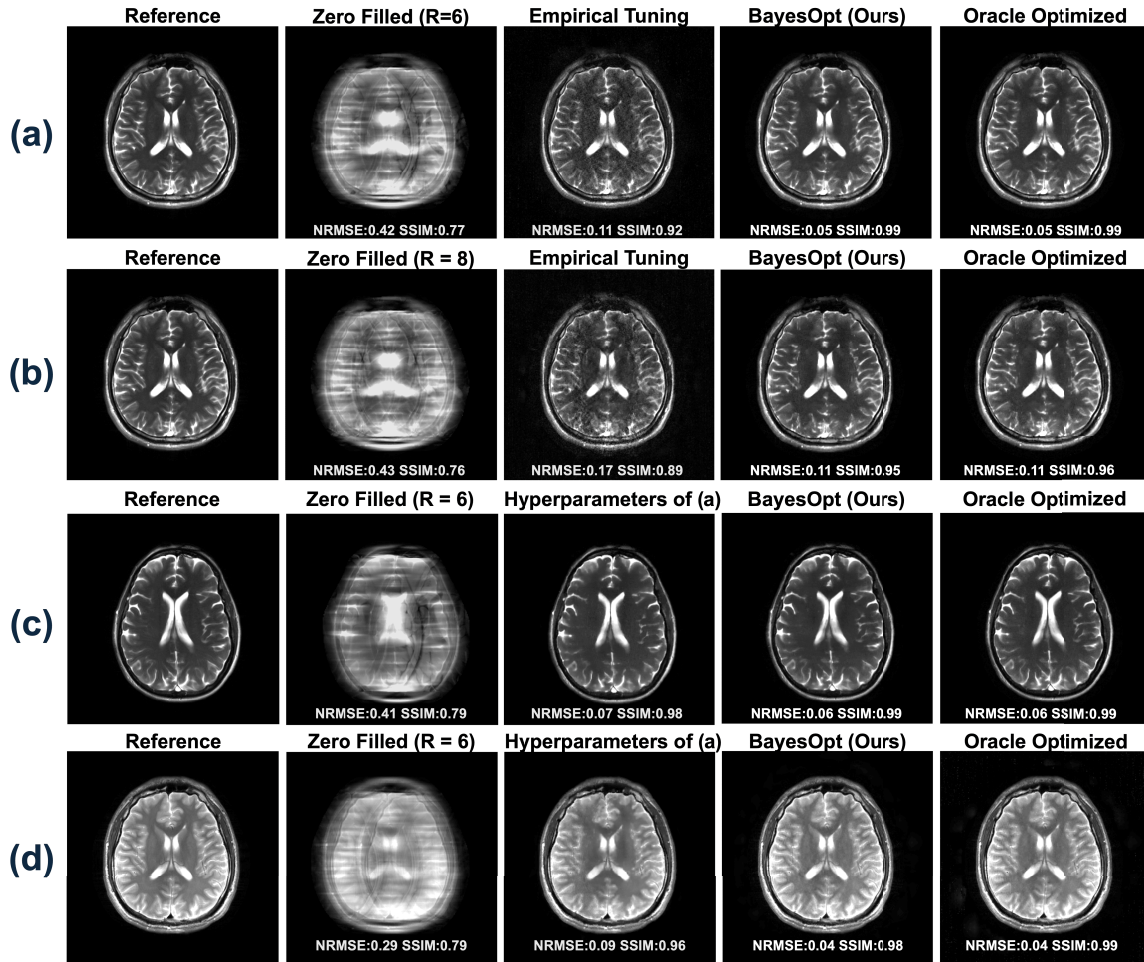


Fig. 9. Demonstration of hyperparameter optimization using Bilevel Optimization. (a) and (b) show comparisons of bSSFP acquisitions at acceleration factors $R = 6$ and $R = 8$ for volunteer 1. (c) shows comparison of bSSFP acquisition for volunteer 2 at $R = 6$, illustrating that hyperparameters are transferable to a different subject for the same imaging sequence. (d) shows a comparison of a T2-weighted TSE acquisition from volunteer 1 in (a) at $R = 6$, illustrating that different sequences require tailored hyperparameter optimization. Acquisition parameters are provided in Table I.

TABLE II
COMPARISON OF HYPERPARAMETER VALUES FOUND BY EMPIRICAL TUNING VS. BAYESIAN OPTIMIZATION. (FIG. 9 (A),(B))

	λ_{ENC}	λ_{MLP}	lr	δ	b
Empirical Tuning	1.0×10^{-5}	1.0×10^{-10}	1.0×10^{-3}	1×10^{-4}	1.50
BayesOpt (R=6)	2.6×10^{-4}	1.0×10^{-7}	9.3×10^{-3}	2×10^{-4}	1.57
BayesOpt (R=8)	3.6×10^{-3}	7.2×10^{-6}	4.4×10^{-4}	2×10^{-4}	1.60

Fig. 9(c) displays the reconstructed images acquired with the same bSSFP sequence with similar slice position for volunteer 2. Bayesian optimization achieved performance comparable to oracle optimization across all metrics. The hyperparameters optimized for volunteer 1 also adapted well to volunteer 2, resulting in only a slight drop of 0.01 in NRMSE and SSIM. This demonstrates that hyperparameters optimized for bilevel INR can be transferred between similar acquisitions with mild performance drop.

Fig. 9(d) uses a T2-weighted TSE scan acquired at the same slice position from volunteer 1 to demonstrate the sequence dependence of INR-based reconstruction. Using hyperparameters optimized for the bSSFP scan of same volunteer led

to an NRMSE that was 0.05 higher and an SSIM that was 0.02 lower than those of tailored Bayesian optimization. The resulting image appeared more blurry because the optimized weight regularization strength was larger for bSSFP sequence (i.e., $\lambda_{\text{Enc}} = 2.6e-4$ in (a)) compared to the optimized one for T2-weighted TSE scan (i.e., $\lambda_{\text{Enc}} = 1.2e-5$ in (d)).

Fig. 10 further demonstrates the transferability of hyperparameters across anatomically similar regions. Reconstructed images of two slices from a $6 \times$ Cartesian undersampled pelvic scan were compared using both tailored hyperparameters (optimized specifically for each slice) and hyperparameters transferred from the other slice. The transferred hyperparameters yielded good reconstructions, with only small decreases in PSNR (0.41 dB and 0.58 dB, respectively) compared to the tailored results.

Table III summarizes the comparison between population-based and scan-specific bilevel hyperparameter optimization. Scan-specific bilevel INR outperformed the population-based variant for every subject and on average. For the population-based reconstructions, we also observed that after an initial improvement, the quantitative metrics degraded as the lower-level optimization proceeded,

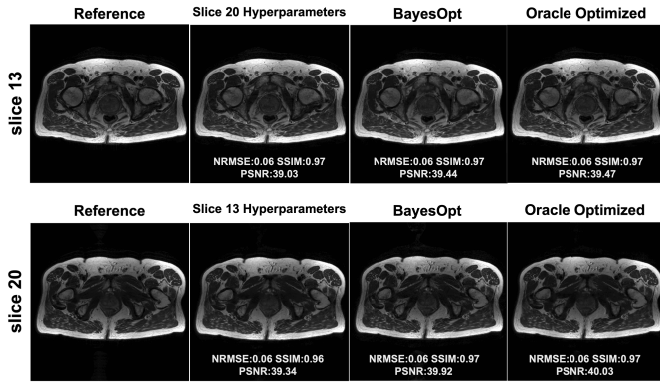


Fig. 10. Demonstration of hyperparameter transferability. Reconstructions of two slices from a multislice prostate acquisition are shown ($6\times$ Cartesian undersampled). Each slice is reconstructed with both its own optimized hyperparameters and those from the other slice, indicating that the optimized hyperparameters are transferable across similar anatomy under the same acquisition.

TABLE III

NRMSE AND SSIM OF POPULATION-BASED VS. SCAN-SPECIFIC BILEVEL-OPTIMIZED INR RECONSTRUCTIONS

	Population-based (N=20)		Scan-specific	
	NRMSE	SSIM	NRMSE	SSIM
Subject 1	0.072	0.952	0.066	0.959
Subject 2	0.054	0.961	0.049	0.968
Subject 3	0.073	0.961	0.060	0.965
Subject 4	0.079	0.952	0.073	0.955
Subject 5	0.076	0.958	0.071	0.960
Mean $\pm$ Std (N=10)	0.070 $\pm$ 0.008	0.956 $\pm$ 0.004	0.065 $\pm$ 0.008	0.962 $\pm$ 0.005

indicating overfitting under a fixed population hyperparameter vector.

These results confirmed the advantage of per-scan hyperparameter optimization, and is expected as INR-based reconstruction parameterizes the target image through both the network weights θ and the hyperparameters β ; a population-tuned β is not tailored to the specific acquisition and therefore cannot achieve the same performance as scan-specific optimization. Nonetheless, the improvements were modest (0.005 lower in NRMSE and 0.006 higher in SSIM on average) on this dataset, implying that the population-based reconstruction could still be used directly or as a warm start for the scan-specific optimization.

D. Examples of Tuned Hyperparameters

Fig. 11 demonstrates the impact of optimized hyperparameters in the upper level of (6). The encoder's ℓ_2 regularization strength λ_{Enc} adjusts the degree of regularization on the positionally encoded feature representation; sub-optimal choices lead to over- or under-regularization. Similarly, the loss-weighting bias δ balances the trade-off between image smoothness and fine details. According to (11), a larger δ diminished higher-frequency components relative to the lower-frequency terms, resulting in blurrier reconstructions. Conversely, smaller δ emphasized high-frequency components, producing clearer but potentially more aliased images.

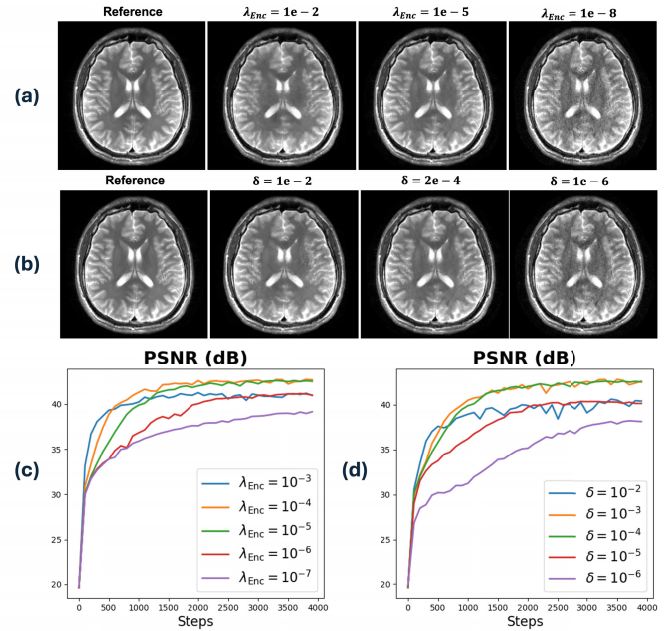


Fig. 11. Illustrating choosing hyperparameter for Bayesian Optimization; (a),(c) ℓ_2 regularization strength for Hash Encoder parameters. (b),(d) Self-weighting loss stability value δ that controls the emphasis on higher frequency k-space components.

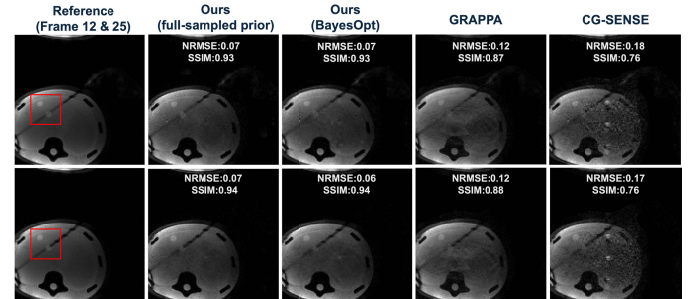


Fig. 12. Comparisons of real-time interventional MRI reconstructions. The INR was pre-trained on the fully-sampled first frame. Each subsequent frame is $6\times$ undersampled with 6 fully-sampled center k-space lines and reconstructed using only 50 residual-learning iterations. Reconstructions of two selected frames obtained using our method with the fully-sampled prior and frame-specific tailored reconstructions without prior are compared with those from GRAPPA and CG-SENSE. Table I provides acquisition parameters.

E. Real-Time Interventional MRI

Fig. 12 compares image quality reconstructed by various methods in real-time interventional MRI during a simulated biopsy procedure using an abdominal biopsy phantom. The first frame was fully sampled to train the INR, while subsequent frames were retrospectively undersampled by $6\times$. Our INR-based residual learning reconstruction outperformed GRAPPA and CG-SENSE by over 0.05 in NRMSE, and showed fewer visible artifacts and less noise, especially near the lesion mimics highlighted in red box. The needle and lesion mimics remained clearly distinguishable. Because the INR was pre-trained on the first frame, reconstructing each subsequent frame took only 180 milliseconds (< 50 iterations), which was near the 175 millisecond temporal resolution required for a $6\times$ undersampled real-time acquisition.

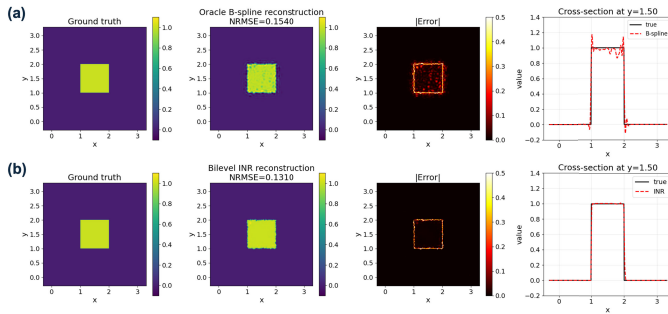


Fig. 13. Preliminary sparse function fitting comparison between a cubic B-spline model and the proposed bilevel-optimized INR on a 2D rect function. Both methods use only Tikhonov regularization without additional handcrafted priors and are trained from the same set of 10000 randomly sampled points. NRMSE is computed on a dense 501×501 evaluation grid. (a) Oracle-tuned cubic B-spline fit: ground truth, reconstruction with hyperparameters selected by grid search over knot density and regularization strength (NRMSE = 0.154), absolute error, and cross-section at $y = 1.50$, showing residual oscillations and edge blurring. (b) Bilevel-optimized INR fit: ground truth, reconstruction (NRMSE = 0.131), absolute error, and cross-section at $y = 1.50$, showing sharper edge transitions and reduced oscillations compared to the B-spline representation.

We also compared this residual learning scheme with frame-specific tailored reconstruction, in which the complete bilevel optimization was performed for each frame. The residual learning reconstruction achieved similar quantitative metrics. The supplementary material provides a movie showing the full time-series reconstruction, including comparisons at different acceleration factors ($4\times$ and $6\times$).

F. Function Fitting Comparison With B-Spline

Fig. 13 summarizes a sparse function fitting comparison between a fixed-basis cubic B-spline and the proposed INR. Because both models use only Tikhonov weight regularization, the comparison more directly reflects the representation without additional handcrafted priors. The oracle-tuned cubic B-spline achieved a NRMSE of 0.154 at $M^* = 75$ knots per dimension and $\lambda^* = 2.51 \times 10^{-2}$, but still showed visible edge blurring and oscillation near the discontinuities. The bilevel-optimized INR achieved a lower NRMSE of 0.131 with hyperparameters $\beta^* = (\lambda_{\text{Enc}}^*, \lambda_{\text{MLP}}^*, \tau^*, b^*) = (7.55 \times 10^{-3}, 9.87 \times 10^{-7}, 5.5 \times 10^{-5}, 1.37)$, showed sharper edge transitions and reduced oscillation in the 1D profile. Because both model classes are continuous and differentiable, this result suggests the benefit of INR is not continuity alone, but its adaptive, multiresolution, and continuous piecewise-linear function class, which may better capture discontinuous structure than a fixed cubic B-spline basis. Meanwhile, we emphasize that this result should be interpreted as empirical evidence for this sparse discontinuous task, which is linked to anatomical boundaries in MRI reconstruction, but not a general theoretical proof that INR is preferable to finite-series models. A more detailed comparison [52] found that all variants of weight-regularized INRs achieved lower NRMSE than the oracle-tuned B-spline (0.129–0.137 versus 0.154), whereas unregularized INR overfit (NRMSE = 0.382), indicating that any practical INR advantage depends on appropriate weight-domain regularization rather than parameter count alone.

VI. DISCUSSION

This paper proposes a bilevel-optimized INR framework for hyperparameter-optimized, scan-specific image reconstruction for accelerated MRI acquisitions. Existing self-supervised deep learning methods rely on task- and subject-dependent hyperparameter tuning, which typically involves prior population data [29] or a time-consuming grid search [24], [26], [28], [31]—both impractical in clinical settings. In contrast, we formulate the MRI reconstruction as a bilevel optimization problem, where the upper level optimizes hyperparameters and the lower level performs INR-based reconstruction given sampled hyperparameters, solely relying on the single under-sampled acquisition. This approach automatically determines scan-specific hyperparameters, providing improved and better-tailored reconstruction.

We validated the proposed method on *in vivo* scans of healthy volunteers across various anatomies, contrasts, sampling patterns, and field strengths. Across the experiments, our bilevel-optimized INR achieved comparable or better quantitative metrics (NRMSE and SSIM) than both model-based and existing self-supervised INR methods. The proposed method also achieved similar quantitative metrics to self-supervised unrolled network based methods with the advantage of much faster training speed and more efficient memory footprint.

Furthermore, we showed that the optimized hyperparameters can be transferred to other subjects, or to similar anatomies acquired with the same protocol, with only mild performance loss. We also showed that direct scan-specific hyperparameter optimization outperforms population-based bilevel optimization with a modest but consistent improvement. A practical next step is to use population-based hyperparameters as a warm start for the bilevel framework and then apply a brief scan-specific fine-tuning—an approach we view as a promising future direction.

A. Analysis of Self-Supervised Bilevel Optimization

The primary challenge of supervised deep learning methods is the need for large, application-specific training datasets and their limited generalization to unseen acquisitions.

This study addresses these limitations by integrating automatic hyperparameter optimization and scan-specific self-supervised reconstruction. Instead of using population-based hyperparameter optimization, we split the undersampled data into further undersampled training and validation sets, inspired by SSDU [23], to achieve self-supervision. We then employ Bayesian optimization for the upper-level search because of its efficiency when evaluating multiple hyperparameters in an expensive lower-level INR-based image reconstruction (i.e., training an INR). Bayesian Optimization also handles discrete inputs and only requires a fixed number of iterations (e.g., 20 for initialization and 40 for Gaussian regression).

Scan-specific hyperparameter tuning is particularly important for INR reconstruction, because the image is represented by a continuous function governed jointly by the network weights and by the hyperparameters. An inappropriate choice of these parameters can cause the network to overfit the undersampled k-space data, degrading image quality. The same bilevel framework can be applied to any reconstruction

model with tunable parameters—linear or nonlinear—to yield per-scan settings that are better matched to the noise level, sampling pattern, and anatomy than population-based defaults.

We focus on five hyperparameters in the proposed INR-based reconstruction that are highly application-dependent: the weight decay for encoder and decoder parameters, the learning rate, the loss-weighting bias, and the multiresolution encoding scale factor. Both weight decay parameters balance the trade-off between noise robustness and structural detail, where λ_{Enc} regularizes positionally encoded latent features and λ_{MLP} directly regularizes the learned function manifolds. λ_{Enc} dominates due to the hash encoder’s higher trainable weight count compared to the compact MLP and its learned multiresolution feature representation. As shown in Fig. 11, the loss-weighting bias further optimizes the balance between blurriness and resolution in the reconstruction. We incorporate the learning rate into hyperparameter optimization because it jointly influences reconstruction results with the other four parameters, and its optimal value is also interdependent on the others. The multiresolution scale factor adapts the encoder’s feature maps on a per-scan basis, selecting the level of detail—from local to global—that best captures the specific image content.

We have also considered additional hyperparameters such as the MLP’s width and depth, but observed that their influence remains relatively stable across different datasets. Fig. 7 shows that even a linear-layer decoder can produce high-quality reconstructions, demonstrating the positional encoder’s dominant contribution for the reconstruction. For the hash encoder, we also tested a table size of $\geq 2^{16}$, which yielded comparable performance across various scans. Overall, overparameterization with optimized self-regularization provided by weight decay is another reason why the hash-encoded INR-based approach can reconstruct highly accelerated MRI scans.

B. Analysis of Components of Bilevel Optimized INR

This study also provides experiments to analyze the effect and selection of key components in an INR for scan-specific MRI reconstruction. The main components of an INR are the positional encoder, the decoder MLP, and the loss function. Fig. 7(b) compares no encoding, frequency encoding [27], dense grid encoding [50], and hash grid encoding [51]. Unlike view synthesis, which leverages multiple view images for training, no encoding and frequency encoding are both slow to train and even unable to capture fine MR image details. By contrast, dense grid and hash grid encodings can reconstruct images effectively because of the learned multiresolution feature representation. However, the dense grid demands more memory and training time since it stores weights for every grid corner. Fig. 7(a) compares MLPs of various depths against a linear layer. All decoders achieve similar performance upon convergence, consistent with prior findings [51] that the weight-regularized encoder primarily solves the inverse problem. However, our activation function differs from SIREN [67], which applies sinusoidal activation for view synthesis. In this study, ReLU activation yielded more stable training and was less prone to overfitting, whereas SIREN tended to introduce residual artifacts. One potential reason is that this task is heavily overparametrized as only a 2D MR image

is reconstructed instead of full 3D view like in SIREN. In addition, the “dying ReLU” effect [70] provides an implicit regularization, which further helps limit overfitting. Additionally, we choose the INR to output the image in real/imaginary representation instead of magnitude/phase because phase periodicity, magnitude-scaled phase gradients, and the $(m, \phi) \sim (-m, \phi + \pi)$ redundancy make magnitude/phase optimization slow and unstable. Real/imaginary outputs avoid these issues and yielded faster reconstructions with less noise and residual artifacts in our tests.

The diagonal weighting matrix $W(\theta, \delta)$ in (7) can be viewed as a learned density compensation function (DCF). For non-Cartesian trajectories, a diagonal DCF is traditionally introduced to balance the non-uniform sampling density and to pre-condition the normal equations [57], [66]. Our weighting plays the same role but is updated each iteration from the current prediction $|Af_\theta(\rightarrow rr)|$, eliminating trajectory-specific tuning. The adaptive weighting is also generalized to Cartesian undersampling, where the magnitude of acquired k-space samples still varies widely; a plain ℓ_2 loss therefore drives the network to fit low-frequency terms first, reflecting the spectral bias of neural networks [71]. By flattening this spectrum, the adaptive weighting accelerates learning of high-frequency details and improves final image quality. Figure 7(c) shows that it consistently outperforms fixed schemes such as identity weighting, Pipe’s analytic DCF, or data weighting.

In principle, one could apply the continuous Fourier transform to the learned INR function and sample it directly at the measured k-space locations, matching continuous nature of the MR acquisition and INR. In practice, this is computationally inefficient, requiring one INR evaluation per k-space point, and still relies on numerical approximation. Hence, we still adopt a conventional discrete forward model.

Although INR-based reconstruction constrains the final image within its learnable function class (yielding implicit regularization), the network remains an overparameterized model and still raises identifiability issues, as shown in Fig. 8(b) and (d). Prior studies mitigated this problem by adding handcrafted image-domain terms such as total variation or low-rank penalties [29], [30], [31]. In contrast, we regularize directly in weight space, penalizing the ℓ_2 norm of both encoder and decoder weights; the associated strengths are chosen scan-specifically by the bilevel loop. This strategy controls overfitting at its source—the learnable function class—while avoiding extra image-domain regularizers. It is worth noting that the original InstantNGP [51] implementation omits weight decay on the hash encoder, because large-scale novel-view synthesis benefits from a maximally expressive multiresolution basis. However, in the single-scan, highly undersampled MRI, this excess capacity leads to rapid overfitting, so explicit weight-space Tikhonov penalties become essential, removing the need for the additional image-domain priors required by earlier hash-encoded INR reconstruction methods.

C. Representation of Bilevel Optimized INR

The multiresolution hash encoder provides the INR with a scan-adaptive, multi-resolution basis. Trainable feature vectors are stored on a cascade of Cartesian grids, with the finest

spacing and cross-level scale factor b automatically selected by the bilevel optimization. Coarse grids form an injective, low-resolution scaffold, while finer grids add local detail. Any hash collisions occur only at these fine levels. Since the gradients of colliding vertices are accumulated into the same table entry, those vertices are effectively coupled and must share a single high-resolution feature vector. High-error voxels (typically high-contrast or high-signal regions) dominate these shared updates, while small, noise-like deviations are averaged out. This coupling shrinks fine-scale parameters toward a common value across the collided locations and suppresses spiky high-frequency variations, i.e., collisions act as a data-driven implicit regularizer that discourages overly complex fine-grid structure.

With ReLU activation, the network represents a continuous, piecewise-linear function [72]. The break planes of this function coincide with both the encoder grid boundaries and the MLP activations, limiting the representable spatial frequency to $f_{\max} = 1/(2\Delta_{\min})$ where $\Delta_{\min} = ab^{K-1}$ is the finest grid spacing. The bilevel optimization automatically chooses the finest grid spacing in the positional encoding and therefore the effective bandwidth for each scan, so the network effectively learns a scan-adaptive multiresolution representation. This property provides a plausible explanation for the performance difference observed in Fig. 13, where the oracle-tuned cubic B-spline model is less effective than the bilevel optimized INR at fitting a discontinuous function, exhibiting higher NRMSE and more pronounced edge blurring and oscillations. Cubic B-spline bases are particularly suited for approximating smooth functions, whereas an INR with ReLU activations represents a continuous piecewise-linear function class that can more faithfully approximate sharp edges. Because *in vivo* MR images contain many anatomical boundaries, this property potentially gives the proposed INR representation an advantage over cubic B-splines for accelerated MRI reconstruction. We emphasize that a general theoretical analysis of when hash-encoded INR should outperform a fixed finite-series basis remains open. Accordingly, our discussion above should be interpreted as a concept-based intuition supported by empirical evidence in Fig. 13 and [52], rather than as a proof of universal superiority. For sufficiently smooth targets, or when interpretability and parameter efficiency are primary considerations, classical finite-series models may remain preferable.

One potential concern for a hash-encoded INR is the identifiability issue raised by the overparameterized representation. However, the effective model capacity is far smaller in practice. Even though the network stores $\sim 2.2\text{M}$ weights for a 256×256 image, only $K \times 2^d = 64$ vertex features are updated per coordinate, and many fine-grid vertices are hashed to the same table entry. Combined with scan-specific ℓ_2 penalties $\lambda_{\text{Enc}}, \lambda_{\text{MLP}}$, the effective degrees of freedom remain modest, avoiding overfitting and the need for explicit image-domain priors, while being less parameter-efficient than a finite-series model.

We emphasize that INR as a representation model alone does not provide strong regularization for highly accelerated MRI, as shown in Fig. 7(a). The good performance of the proposed method relies on weight-domain ℓ_2 regularization applied to both encoder and MLP parameters, with strengths

selected by bilevel optimization. In our experiments, this weight-regularized INR achieved better overall performance than an INR combined with a total variation image-domain regularizer as in [29].

Finally, because the encoder operates entirely in the image domain, the same network architecture is used for Cartesian, spiral, and Poisson-disc sampling, with only the forward model A being changed. This trajectory independence is shared with other continuous image-domain models such as finite-series expansions. In our framework, this property, together with the adaptive bandwidth and weight-domain regularization, makes the hash-encoded INR a convenient choice for scan-specific MRI reconstruction across diverse protocols.

D. Limitations and Future Work

The proposed bilevel optimized INR still requires coil sensitivity maps estimated from a low-resolution GRE pre-scan or the center of k-space. Although the method is faster than most self-supervised reconstruction approaches, it remains slower than SENSE, GRAPPA, and similar to other model-based compressed sensing algorithms if they are also implemented in PyTorch with GPU support and INR is trained from scratch. The time difference is only a few seconds for 2D reconstructions but can become significant for higher-dimensional datasets. However, we observed in Table III that population-based bilevel optimized INR achieved only a modest performance drop compared to per-scan optimization. Thus, a practical solution is to warm-start both levels of the proposed method: initialize the upper level with population-based hyperparameters derived offline, and initialize the lower level with weights from a prior reconstruction of the same protocol or by fitting conventional computationally-efficient reconstruction (e.g., SENSE and GRAPPA) directly in the image domain.

Extending bilevel optimized INR to 3D MRI is a promising direction, especially for free-breathing abdominal and cardiac imaging where undersampling is mandatory. A direct approach is to feed all coordinates (x, y, z) of the entire volume into the encoder so that the INR learns a single high-dimensional function, but this raises two practical challenges. First, both the upper-level hyperparameter search and the lower-level INR training become time-consuming in higher dimensions. A useful strategy is to warm-start both levels as mentioned above. Second, memory use grows rapidly with dimensionality. Unlike image-domain INR applications—such as gigapixel fitting [51]—MRI involves a forward model with an FFT that is most efficient when the complete image is processed at each iteration; the effective batch size is therefore one. Although a region-of-interest mask can reduce the coordinate set, this may still be insufficient for large 3D volumes.

If one phase encoding direction is fully sampled, then the 3D problem can be split into independent 2D slices, greatly reducing time and memory burdens and enabling parallel processing. Such sampling schemes (e.g., Poisson-disc with a fully-sampled partition dimension) are also less ill-posed and support higher accelerations, as suggested by the 2D results in Fig. 3. Thus, jointly designing the sampling pattern and the

INR reconstruction is another attractive direction for future 3D work.

Additionally, although this work focuses on INR as the lower-level representation, the proposed bilevel optimization framework is not restricted to INR and could, in principle, be applied to traditional finite-series models for MRI reconstruction. Investigating such extensions, along with comprehensive comparisons of representation capacity, parameter efficiency, and interpretability between INR and traditional models, remains an interesting direction for future work.

While the proposed method showed only modest differences in NRMSE and SSIM compared to ZS-SSL, the resolution analysis using the Fourier radial error plot (Fig. 5) indicates higher relative error in the high spatial-frequency range. However, the relationship between this type of resolution analysis and the diagnostic quality of the reconstructed images remains to be fully understood. The LPR study shows that scan-specific methods (e.g., INR and ZS-SSL) did not accurately reconstruct the small perturbation (0.5%) because the perturbed and unperturbed runs require independently trained network weights. This suggests the perturbation energy lies below the implicit regularization floor set by weight decay, network architectural bias, and loss weighting. With a larger 5% perturbation, the responses show blurring and signal leakage, but this violates the small-perturbation assumption underlying LPR analysis. Furthermore, LPR was originally designed for model-based and supervised deep learning methods with fixed weights between perturbed runs, which may not align well with scan-specific approaches that retrain per dataset. Extending perturbation-based analysis to scan-specific reconstruction methods, and improving their sensitivity to small perturbations, remain important directions for future work. Finally, although scan-specific bilevel optimization with weight domain regularization enhances noise robustness, INR as an expressive function approximator can still yield noisy reconstructions when the undersampled k -space data have low SNR or a high acceleration factor. Improving the noise robustness of INR-based MRI reconstruction with or without additional regularizers or denoisers remains an open research question, especially critical for applications like diffusion-weighted MRI where SNR inherently limits the reconstruction quality.

REFERENCES

- [1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, Nov. 1999.
- [2] M. A. Griswold et al., "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, Jun. 2002.
- [3] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Feb. 2008.
- [4] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint," *Magn. Reson. Med.*, vol. 57, no. 6, pp. 1086–1098, Jun. 2007.
- [5] J. P. Haldar, "Low-rank modeling of local k -space neighborhoods (LORAKS) for constrained MRI," *IEEE Trans. Med. Imag.*, vol. 33, no. 3, pp. 668–681, Mar. 2014.
- [6] 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, pp. 1499–1514, Apr. 2016.
- [7] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 480–495, Dec. 2016.
- [8] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "K-T FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, Jan. 2009.
- [9] L. Feng et al., "Golden-angle radial sparse parallel MRI: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI," *Magn. Reson. Med.*, vol. 72, no. 3, pp. 707–717, Sep. 2014.
- [10] L. Feng, L. Axel, H. Chandarana, K. T. Block, D. K. Sodickson, and R. Otazo, "XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing," *Magn. Reson. Med.*, vol. 75, no. 2, pp. 775–788, Feb. 2016.
- [11] G. L. da Cruz, A. Bustin, O. Jaubert, T. Schneider, R. M. Botnar, and C. Prieto, "Sparsity and locally low rank regularization for MR fingerprinting," *Magn. Reson. Med.*, vol. 81, no. 6, pp. 3530–3543, Jun. 2019.
- [12] J. Zbontar et al., "FastMRI: An open dataset and benchmarks for accelerated MRI," 2018, *arXiv:1811.08839*.
- [13] K. Kwon, D. Kim, and H. Park, "A parallel MR imaging method using multilayer perceptron," *Med. Phys.*, vol. 44, no. 12, pp. 6209–6224, Dec. 2017.
- [14] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Phys. Med. Biol.*, vol. 63, no. 13, Jun. 2018, Art. no. 135007.
- [15] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019.
- [16] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. CVPR*, Jun. 2018, pp. 1828–1837.
- [17] F. Knoll et al., "Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 128–140, Jan. 2020.
- [18] G. Mårtensson et al., "The reliability of a deep learning model in clinical out-of-distribution MRI data: A multicohort study," *Med. Image Anal.*, vol. 66, Dec. 2020, Art. no. 101714.
- [19] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Ugurbil, "Scan-specific robust artificial-neural-networks for k -space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magn. Reson. Med.*, vol. 81, no. 1, pp. 439–453, Jan. 2019.
- [20] T. H. Kim, P. Garg, and J. P. Haldar, "LORAKI: Autocalibrated recurrent neural networks for autoregressive MRI reconstruction in k -space," 2019, *arXiv:1904.09390*.
- [21] Y. Arefeen, O. Beker, J. Cho, H. Yu, E. Adalsteinsson, and B. Bilgic, "Scan-specific artifact reduction in k -space (SPARK) neural networks synergize with physics-based reconstruction to accelerate MRI," *Magn. Reson. Med.*, vol. 87, no. 2, pp. 764–780, Feb. 2022.
- [22] A. D. Desai et al., "Noise2Recon: Enabling SNR-robust MRI reconstruction with semi-supervised and self-supervised learning," *Magn. Reson. Med.*, vol. 90, no. 5, pp. 2052–2070, Nov. 2023.
- [23] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Ugurbil, and M. Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magn. Reson. Med.*, vol. 84, no. 6, pp. 3172–3191, Dec. 2020.
- [24] B. Yaman, S. A. H. Hosseini, and M. Akçakaya, "Zero-shot self-supervised learning for MRI reconstruction," in *Proc. Int. Conf. Learn. Represent.*, Apr. 2022.
- [25] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," 2017, *arXiv:1711.10925*.
- [26] A. P. Leynes, N. Deveshwar, S. S. Nagarajan, and P. E. Z. Larson, "Scan-specific self-supervised Bayesian deep non-linear inversion for undersampled MRI reconstruction," *IEEE Trans. Med. Imag.*, vol. 43, no. 6, pp. 2358–2369, Jun. 2024.
- [27] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in *Proc. ECCV*, 2020, pp. 405–421.
- [28] L. Shen, J. Pauly, and L. Xing, "NeRP: Implicit neural representation learning with prior embedding for sparsely sampled image reconstruction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 1, pp. 770–782, Jan. 2024.

- [29] R. Feng et al., "IMJENSE: Scan-specific implicit representation for joint coil sensitivity and image estimation in parallel MRI," *IEEE Trans. Med. Imag.*, vol. 43, no. 4, pp. 1539–1553, Apr. 2024.
- [30] B. Liu, H. She, and Y. P. Du, "Scan-specific unsupervised highly accelerated non-Cartesian CEST imaging using implicit neural representation and explicit sparse prior," *IEEE Trans. Biomed. Eng.*, vol. 71, no. 10, pp. 3032–3045, Oct. 2024.
- [31] J. Feng et al., "Spatiotemporal implicit neural representation for unsupervised dynamic MRI reconstruction," *IEEE Trans. Med. Imag.*, vol. 44, no. 5, pp. 2143–2156, May 2025.
- [32] W. C. Karl, "Regularization in image restoration and reconstruction," in *Handbook of Image and Video Processing*, A. C. Bovik, Ed., Amsterdam, The Netherlands: Elsevier, 2005, pp. 183–202.
- [33] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 322–336, Jul. 1992.
- [34] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Rev.*, vol. 34, no. 4, pp. 561–580, 1992.
- [35] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Scientific Comput.*, vol. 14, no. 6, pp. 1487–1503, Nov. 1993.
- [36] F.-H. Lin, K. K. Kwong, J. W. Belliveau, and L. L. Wald, "Parallel imaging reconstruction using automatic regularization," *Magn. Reson. Med.*, vol. 51, no. 3, pp. 559–567, Mar. 2004.
- [37] S. J. Reeves and R. M. Mersereau, "Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration," *Opt. Eng.*, vol. 29, no. 5, pp. 446–454, 1990.
- [38] D. A. Girard, "The fast Monte-Carlo cross-validation and C_L procedures: Comments, new results and application to image recovery problems," *Comput. Stat.*, vol. 10, pp. 205–231, Jun. 1995.
- [39] J. D. Carew, G. Wahba, X. Xie, E. V. Nordheim, and M. E. Meyerand, "Optimal spline smoothing of fMRI time series by generalized cross-validation," *NeuroImage*, vol. 18, no. 4, pp. 950–961, Apr. 2003.
- [40] S. Ramani, Z. Liu, J. Rosen, J.-F. Nielsen, and J. A. Fessler, "Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3659–3672, Aug. 2012.
- [41] F. Luisier, T. Blu, and P. J. Wolfe, "A CURE for noisy magnetic resonance images: Chi-square unbiased risk estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3454–3466, Aug. 2012.
- [42] A. Marin, C. Chaux, J.-C. Pesquet, and P. Ciuciu, "Image reconstruction from multiple sensors using Stein's principle: Application to parallel MRI," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Mar. 2011, pp. 465–468.
- [43] S. Ramani, D. S. Weller, J.-F. Nielsen, and J. A. Fessler, "Non-Cartesian MRI reconstruction with automatic regularization via Monte-Carlo SURE," *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1411–1422, Aug. 2013.
- [44] C. Crockett and J. A. Fessler, "Bilevel methods for image reconstruction," *Found. Trends Signal Process.*, vol. 15, nos. 2–3, pp. 121–289, May 2022.
- [45] P. I. Frazier, "A tutorial on Bayesian optimization," 2018, *arXiv:1807.02811*.
- [46] J. A. Fessler, "Optimization methods for magnetic resonance image reconstruction: Key models and optimization algorithms," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 33–40, Jan. 2020.
- [47] B. Zhao, J. P. Haldar, C. Brinegar, and Z.-P. Liang, "Low rank matrix recovery for real-time cardiac MRI," in *Proc. IEEE Int. Symp. Biomed. Imaging: From Nano to Macro*, Apr. 2010, pp. 996–999.
- [48] J. A. Fessler, "Model-based image reconstruction for MRI," *IEEE Signal Process. Mag.*, vol. 27, no. 4, pp. 81–89, Jul. 2010.
- [49] M. Tancik et al., "Fourier features let networks learn high frequency functions in low dimensional domains," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 1–11.
- [50] J. N. P. Martel, D. B. Lindell, C. Z. Lin, E. R. Chan, M. Monteiro, and G. Wetzstein, "Acorn: Adaptive coordinate networks for neural scene representation," *ACM Trans. Graph.*, vol. 40, no. 4, pp. 1–13, Jul. 2021.
- [51] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM Trans. Graph.*, vol. 41, no. 4, pp. 1–15, Jul. 2022.
- [52] H. Yu, Y. Jiang, and J. A. Fessler, "Comparing implicit neural representations and B-splines for continuous function fitting from sparse samples," 2026, *arXiv:2602.20535*.
- [53] B. Mildenhall, P. Hedman, R. Martin-Brualla, P. P. Srinivasan, and J. T. Barron, "NeRF in the dark: High dynamic range view synthesis from noisy raw images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 16169–16178.
- [54] M. Uecker et al., "ESPIRiT—An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn. Reson. Med.*, vol. 71, no. 3, pp. 990–1001, Mar. 2014.
- [55] F. Huang, S. Vijayakumar, Y. Li, S. Hertel, and G. R. Duensing, "A software channel compression technique for faster reconstruction with many channels," *Magn. Reson. Imag.*, vol. 26, no. 1, pp. 133–141, Jan. 2008.
- [56] M. S. Hansen and P. Kellman, "Image reconstruction: An overview for clinicians," *J. Magn. Reson. Imag.*, vol. 41, no. 3, pp. 573–585, Mar. 2015.
- [57] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003.
- [58] M. J. Muckley, R. Stern, T. Murrell, and F. Knoll, "TorchKbNufft: A high-level, hardware-agnostic non-uniform fast Fourier transform," in *Proc. ISMRM Workshop Data Sampling Image Reconstruction*, 2020.
- [59] Y.-H. Shih, G. Wright, J. Andén, J. Blaschke, and A. H. Barnett, "CuFINUFFT: A load-balanced GPU library for general-purpose nonuniform FFTs," 2021, *arXiv:2102.08463*.
- [60] A. B. Taylor, J. M. Hendrickx, and F. Glineur, "Exact worst-case performance of first-order methods for composite convex optimization," *SIAM J. Optim.*, vol. 27, no. 3, pp. 1283–1313, Jan. 2017.
- [61] L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, Aug. 2013.
- [62] G. Wang, N. Shah, K. Zhu, D. Noll, and J. Fessler, "MIRTorch: A PyTorch-powered differentiable toolbox for fast image reconstruction and scan protocol optimization," in *Proc. ISMRM Annu. Meeting*, 4982, p. 4982.
- [63] O. Maier et al., "CG-SENSE revisited: Results from the first ISMRM reproducibility challenge," *Magn. Reson. Med.*, vol. 85, no. 4, pp. 1821–1839, Apr. 2021.
- [64] T. H. Kim, K. Setsompop, and J. P. Haldar, "LORAKS makes better SENSE: Phase-constrained partial Fourier SENSE reconstruction without phase calibration," *Magn. Reson. Med.*, vol. 77, no. 3, pp. 1021–1035, Mar. 2017.
- [65] B. A. Hargreaves, D. G. Nishimura, and S. M. Conolly, "Time-optimal multidimensional gradient waveform design for rapid imaging," *Magn. Reson. Med.*, vol. 51, no. 1, pp. 81–92, Jan. 2004.
- [66] J. G. Pipe and P. K. Menon, "Sampling density compensation in MRI: Rationale and an iterative numerical solution," *Magn. Reson. Med.*, vol. 41, no. 1, pp. 179–186, Jan. 1999.
- [67] V. Sitzmann, J. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," in *Proc. NIPS*, vol. 33, 2020, pp. 7462–7473.
- [68] T. H. Kim and J. P. Haldar, "The Fourier radial error spectrum plot: A more nuanced quantitative evaluation of image reconstruction quality," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Washington, DC, USA: IEEE, Apr. 2018, pp. 61–64.
- [69] C.-C. Chan and J. P. Haldar, "Local perturbation responses and checkerboard tests: Characterization tools for nonlinear MRI methods," *Magn. Reson. Med.*, vol. 86, no. 4, pp. 1873–1887, Oct. 2021, doi: [10.1002/mrm.28828](https://doi.org/10.1002/mrm.28828).
- [70] L. Lu, Y. Shin, Y. Su, and G. E. Karniadakis, "Dying ReLU and initialization: Theory and numerical examples," *Commun. Comput. Phys.*, vol. 28, no. 5, pp. 1671–1706, Jun. 2020, doi: [10.4208/cicp.oa-2020-0165](https://doi.org/10.4208/cicp.oa-2020-0165).
- [71] N. Rahaman et al., "On the spectral bias of neural networks," in *Proc. 36th Int. Conf. Mach. Learn. (ICML)*, vol. 97, Jun. 2019, pp. 5301–5310.
- [72] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: <http://www.deeplearningbook.org>