A Bilevel Sensitivity-Corrected Reconstruction Framework with Deep Priors for Parallel MRI

Weipeng Chen[1], Yan-Ran Li[1*], Raymond H Chan[2,3], Lixin Shen[4], Xiaosheng Zhuang[5]

[1*]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, 518060, China .
[2*]School of Data Science and Department of Operations and Risk Management, Lingnan University, Tuen Mun, Hong Kong, China .
[3*]Hong Kong Centre for Cerebro-Cardiovascular Health Engineering, Sha Tin, Hong Kong, China .
[4*]Department of Mathematics, Syracuse University, Syracuse, NY, 13244, USA .
[5*]Department of Mathematics, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong, China .

*Corresponding author(s). E-mail(s): lyran@szu.edu.cn;
Contributing authors: chenweipeng2019@email.szu.edu.cn; raymond.chan@ln.edu.hk; lshen03@syr.edu; xzhuang7@cityu.edu.hk;

Abstract

Parallel magnetic resonance imaging (pMRI) accelerates data acquisition by undersampling multi-coil k-space. Its reconstruction quality, however, deteriorates when coil sensitivity maps (CSMs) or calibration kernels are inaccurate, especially when only limited auto-calibration signal (ACS) data are available. We propose SRSC+, a model-driven bilevel optimization framework that couples SENSE-based image reconstruction with SPIRiT-based k-space calibration through shared CSMs. The bilevel formulation explicitly decouples sensitivity estimation from kernel calibration, thereby enabling iterative correction of both components and reducing error accumulation that often arises in dual-domain methods. In addition, SRSC+ introduces a deep-prior-guided regularization strategy that preserves the structure of classical linear regularizers while adaptively learning spatially varying regularization weights from denoised intermediate reconstructions. Experiments on out-of-distribution datasets under

diverse sampling patterns show that SRSC+ achieves state-of-the-art performance across multiple fidelity and perceptual metrics, while remaining robust to scarce ACS data and imperfect CSM initialization. Visual comparisons further demonstrate effective artifact suppression without pseudo-structural distortions, together with strong generalization across scanners and acquisition protocols. The implementation code is available at https://github.com/chenGTMES/SRSC/.

Keywords: Bilevel optimization, Deep priors, Model-driven reconstruction, Regularization parameter estimation, Coil sensitivity correction

1 Introduction

Magnetic resonance imaging (MRI) is a central imaging modality because it provides high spatial resolution, strong soft-tissue contrast, and flexible acquisition protocols without ionizing radiation. In conventional spin-warp MRI [1], spatial information is encoded through gradient-modulated radiofrequency pulses, and the measured k-space data are mapped to the image domain by an inverse Fourier transform. A key limitation of this acquisition strategy is the need for repeated phase encoding, which leads to long scan times, greater sensitivity to motion, and reduced patient comfort.

Parallel MRI (pMRI) addresses this limitation by using multiple receiver coils to acquire undersampled k-space data simultaneously [2]. Because each coil has a spatially varying sensitivity profile, the multi-coil measurements contain additional spatial information that can be used to disentangle aliasing caused by undersampling. In practice, however, aggressive undersampling and imperfect sensitivity modeling can substantially degrade reconstruction quality, particularly when ACS data are limited and the estimated CSMs or calibration kernels are unreliable.

In this work, we propose SRSC+ (SENSE Reconstruction with SPIRiT Calibration and Deep Priors), a model-driven bilevel framework for pMRI reconstruction. SRSC+ integrates three components: a SENSE-based image-domain reconstruction model, a SPIRiT-based k-space calibration model for refining coil sensitivity information, and a score-based generative prior that adaptively determines regularization weights while preserving the structure of a classical linear-operator-based regularizer. This design enables controlled correction of sensitivity and calibration errors while leveraging learned priors without sacrificing interpretability or cross-domain robustness.

2 Related Work and Motivation

This section reviews model-driven and deep learning–driven MRI reconstruction methods, highlighting limitations in CSM estimation, kernel calibration, dual-domain error accumulation, and generalization under distribution shift. These challenges motivate a bilevel framework that adaptively corrects sensitivities, integrates image- and k-space models, and leverages deep priors in a controlled way.

2.1 Model-Driven Reconstruction

Sensitivity encoding (SENSE) is a widely used approach that reconstructs images by explicitly leveraging CSMs [3, 4]. Its performance depends critically on the accuracy of the CSMs; however, accurate sensitivity estimation is often challenging in clinical settings because of hardware geometry, noise, and limited ACS coverage.

In contrast, GRAPPA [5] avoids explicit sensitivity estimation by calibrating interpolation kernels from ACS data to recover missing k-space samples. SPIRiT [6] extends GRAPPA by enforcing self-consistency through a single convolutional kernel per coil and supports arbitrary sampling patterns via an inverse-problem formulation. However, when ACS lines are scarce, kernel estimation can be unstable, leading to high-frequency errors and loss of fine detail.

To mitigate the ill-posedness induced by undersampling, classical regularization techniques are often integrated into SENSE or SPIRiT, including total variation (TV) [7], wavelet-based [8], and frame-based priors [9]. TV effectively suppresses noise and artifacts but can introduce staircase artifacts; higher-order variants such as TGV [10] alleviate this at the cost of more complex optimization. Joint sparse regularization of multi-coil wavelet coefficients has been explored for SPIRiT reconstruction [11]. Tight-frame constructions, such as 3D directional Haar framelets [12], have shown strong performance in preserving structures while reducing artifacts.

In short, model-driven methods offer interpretability, stability, and good generalization across various acquisition settings; however, their performance is limited by the fidelity of CSMs or kernel calibration, especially when ACS lines are scarce.

2.2 Deep Learning-Driven Reconstruction

Deep learning has significantly advanced MRI reconstruction performance by learning rich data-driven priors from large image collections. Existing approaches can be broadly grouped into three categories:

1. **Deep Unrolled Networks.** These methods unroll iterative optimization algorithms into trainable networks that alternate data-consistency steps with learned proximal or denoising modules [13–16]. Hybrid designs, such as GrappaNet [17], integrate GRAPPA-style reconstruction with deep refinement, while GAN-based methods [18] synthesize missing k-space data. Although unrolled networks achieve strong empirical performance, they often entangle learned components with algorithmic hyperparameters, hindering theoretical analysis and cross-domain generalization.
2. **Generative Networks.** Generative models, including score-based and diffusion approaches, have recently been explored for MRI reconstruction [19–22]. By learning the distribution of fully sampled images or high-frequency k-space components, these methods can recover missing data under physical constraints. Representative examples include HFS-SDE [19], SPIRiT-diffusion [20], TRPA [21], and DKTM [22]. While capable of capturing rich image priors, generative pipelines are often computationally intensive and sensitive to mismatches between training and testing acquisition protocols.

3. **Plug-and-Play (PnP) Methods.** PnP methods [23, 24] incorporate pre-trained denoisers into iterative solvers, implicitly replacing explicit regularization. In MRI, ResiDe [25] and FRSGM [26] combine SENSE-type data consistency with CNN or ensemble denoisers. PnP and Regularization by Denoising (RED) frameworks offer flexibility in leveraging learned priors, but their implicit regularization is difficult to interpret and may overfit the training distribution, potentially introducing artifacts on out-of-distribution (OOD) data.

In addition to the above shortcomings, deep learning–based methods can be sensitive to changes in the field of view, slice thickness, and other acquisition parameters [27], which are commonplace in clinical practice.

2.3 Analysis of Existing Methods

Existing MRI reconstruction methods face three recurring challenges.

- **Errors in CSMs and interpolation kernels.** Model-driven approaches faithfully encode acquisition physics and often provide strong interpretability, but their performance is highly sensitive to the accuracy of the CSMs and the k-space interpolation kernel. In undersampled and noisy settings, both quantities can be difficult to estimate reliably.
- **Error accumulation in dual-domain models.** Recent dual-domain approaches [28, 29] jointly impose constraints in the image and k-space domains. Although this coupling can improve reconstruction quality, errors in sensitivity estimation and kernel calibration may reinforce one another and accumulate over iterations, ultimately degrading the final reconstruction.
- **Limited generalization of deep learning–driven methods.** Deep learning–driven methods exploit powerful data-driven priors, but their generalization can be limited by shifts in scanner type, field of view, slice thickness, or acquisition protocol [27]. Under such distribution shifts, tightly coupled learned components may become less reliable and can introduce visually plausible but incorrect structures.

2.4 Motivation

These observations suggest that an effective reconstruction framework should satisfy three requirements:

- It should update and correct the CSMs or k-space interpolation kernels adaptively.
- It should integrate image-domain and k-space-domain models in a principled manner without inducing uncontrolled error accumulation.
- It should benefit from deep learning while maintaining robustness under distribution shift.

Motivated by these requirements, we develop a bilevel optimization framework that decouples sensitivity estimation from k-space kernel calibration. SENSE-based reconstruction and SPIRiT-based calibration are formulated as separate but interacting subproblems during iteration [30]. We further adopt a hybrid strategy that preserves a classical linear regularizer while learning its parameters from deep priors guided by

empirical observations (Section 3.3). In this way, the proposed framework combines the flexibility of learned priors with the interpretability and stability of model-driven reconstruction.

2.5 Contributions

We summarize the main contributions of this work as follows.

1. **Bilevel framework.** We propose a bilevel reconstruction framework that combines SENSE-based image reconstruction with SPIRiT-based k-space calibration. By explicitly separating sensitivity-estimation errors from kernel-calibration errors, the framework reduces error accumulation and improves robustness relative to stand-alone SENSE or SPIRiT formulations.
2. **Deep-prior-guided regularization.** We introduce a deep-prior-guided regularization strategy based on a score-based generative model. A pretrained denoiser is applied to intermediate reconstructions, and the resulting multi-coil framelet coefficients are used to construct spatially adaptive regularization weights. Importantly, the linear regularization operator remains unchanged; only its parameters are adapted.
3. **Algorithms.** We develop efficient algorithms for the resulting SRSC and SRSC+ models. The upper-level reconstruction problem is solved by PD3O, while the lower-level calibration problem is solved by conjugate gradients. The updates of the CSMs and the deep-prior-based regularization weights are performed on different time scales to balance accuracy and efficiency.
4. **Ablation studies.** We conduct ablation studies to quantify the contributions of SENSE reconstruction, SPIRiT calibration, and deep-prior-guided regularization. We also compare the proposed parameter-adaptation strategy with a PnP-RED alternative that uses the same denoiser.
5. **Performance and robustness.** We demonstrate strong performance and robustness across multiple datasets and sampling patterns. In particular, SRSC+ achieves competitive or state-of-the-art results under limited ACS data and remains stable on OOD data, where several deep learning baselines degrade more noticeably.

2.6 Outline of the Paper

The remainder of this paper is organized as follows. Section 3 introduces the SENSE-based reconstruction model and the SPIRiT-based calibration model, along with empirically observed priors on regularization parameters, which are later transformed into deep priors to improve reconstruction. Section 4 analyzes existing methods and motivates our approach, then presents the proposed bilevel optimization framework SRSC and its extension with deep-prior-guided regularization, termed SRSC+. Ablation studies are conducted to validate the contribution of each component. Section 5 reports extensive experimental results demonstrating the superiority of SRSC+. Finally, Section 6 concludes the paper.

3 Preliminary Work

In this section, we review the SENSE model and the definition of CSMs, together with the SENSE3d algorithm based on 3D directional Haar framelet regularization. We then recall the SPIRiT k-space calibration model, which forms the lower-level component of our bilevel framework. Finally, we summarize empirical observations on regularization parameters that will later be converted into deep regularization priors.

3.1 SENSE Reconstruction and Coil Sensitivity Maps

Let c denote the number of receive coils, and let $u \in \mathbb{R}^n$ ($n = h \times w$) be the target image. In SENSE, the acquisition model for the ℓ-th coil is

$$g_\ell = PFS_\ell u + \epsilon_\ell, \quad \ell = 1, 2, \ldots, c, \tag{1}$$

where $g_\ell \in \mathbb{C}^n$ are the undersampled k-space measurements, $P \in \{0,1\}^{n\times n}$ is a diagonal sampling matrix, $F \in \mathbb{C}^{n\times n}$ is the discrete Fourier transform (DFT), $S_\ell \in \mathbb{C}^{n\times n}$ is the CSM diagonal matrix for the ℓ-th coil, and $\epsilon_\ell \in \mathbb{C}^n$ is complex Gaussian noise. Stacking all coils yields the compact pMRI acquisition model

$$g = \mathcal{P}\mathcal{F}\mathcal{S}u + \epsilon, \tag{2}$$

where

$$g = \begin{bmatrix} g_1 \\ \vdots \\ g_c \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} P & & \\ & \ddots & \\ & & P \end{bmatrix}, \quad \mathcal{F} = \begin{bmatrix} F & & \\ & \ddots & \\ & & F \end{bmatrix}, \quad \mathcal{S} = \begin{bmatrix} S_1 \\ \vdots \\ S_c \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_c \end{bmatrix}.$$

The quality of SENSE reconstruction depends critically on the accuracy of the CSMs $\{S_\ell\}_{\ell=1}^c$. In practice, however, k-space data are undersampled, and fully sampled coil images u_ℓ are not directly available. Instead, CSMs are typically pre-estimated using only the ACS region. Let g_ℓ^{ACS} denote the ACS data for the ℓ-th coil and $u_\ell^{\mathrm{ACS}} = F^\top g_\ell^{\mathrm{ACS}}$. The CSMs are then approximated as

$$S_\ell[i,i] = \frac{u_\ell^{\mathrm{ACS}}[i]}{\sqrt{\sum_{j=1}^c (u_j^{\mathrm{ACS}}[i])^2}}. \tag{3}$$

where $S_\ell[i,i]$ is the i-th diagonal entry of S_ℓ, and $u_\ell^{\mathrm{ACS}}[i]$ is the i-th pixel of u_ℓ^{ACS}. Due to the limited spatial support of the ACS region, the estimated CSMs may suffer from noise and modeling inaccuracies, which can significantly degrade the final reconstruction.

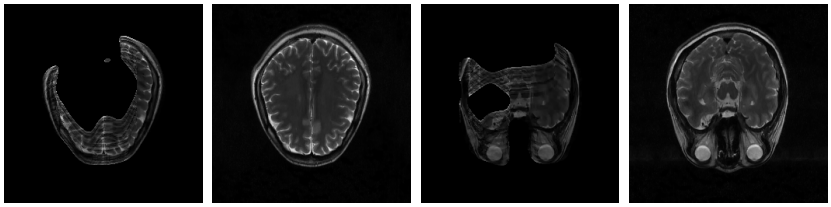
To mitigate these effects, the SENSE3d algorithm [9] incorporates a 3D directional Haar tight framelet (3DHTF) regularization into the SENSE framework. Denoting by $\mathcal{Q} = \mathrm{diag}(Q, \ldots, Q)$ the diagonal mask that selects unsampled k-space locations (with

$Q = I - P$), by $W \in \mathbb{R}^{Lcn \times cn}$ the 3DHTF transform, and by $\Gamma \in \mathbb{R}^{Lcn \times Lcn}$ a diagonal matrix of regularization parameters, SENSE3d solves

$$\min_u \left\{ \frac{1}{2} \|\mathcal{P}\mathcal{F}\mathcal{S}u - g\|_2^2 + \|\Gamma W \mathcal{F}^\top (\mathcal{Q}\mathcal{F}\mathcal{S}u + g)\|_1 \right\}, \tag{4}$$

where L is the number of framelet filters.

Since the CSMs in (3) are estimated solely from ACS data, they are prone to errors, particularly when the ACS lines are insufficient, which propagate through (4) and degrade reconstruction quality. ESPIRiT [31] mitigates this by exploiting multi-coil k-space redundancy and estimating CSMs via eigen-decomposition of a calibration matrix, but this process discards absolute phase information. Phase-corrected variants such as VCC-ESPIRiT [32] recover absolute phase using virtual conjugate coils; however, they still rely on ACS-based estimation and thus remain sensitive to high-frequency information loss and instability.



(a1) ESPIRiT (b1) SRSC+ (a2) ESPIRiT (b2) SRSC+

Fig. 1 Comparison of (a) ESPIRiT-based CSMs estimation and (b) SPIRiT-based periodic calibration in k-space for updating CSMs. The proposed SPIRiT-based approach effectively corrects CSMs estimation errors that cause severe reconstruction errors in ESPIRiT.

However, as shown in Figure 1, the ESPIRiT-based CSMs pre-estimation method is inherently an ACS-based approach. When ACS lines are insufficient, severe estimation errors may occur, leading to notable reconstruction information loss. Therefore, one of the core objectives of this work is to address the potential limitations of ESPIRiT-based CSMs pre-estimation. As illustrated in the figure, the proposed SRSC+ achieves accurate reconstruction even under limited ACS conditions, without sacrificing image information.

3.2 SPIRiT Calibration in k-space

In contrast to SENSE, the SPIRiT framework performs self-consistent k-space interpolation without explicitly estimating CSMs. Let $z \in \mathbb{C}^{cn}$ denote the multi-coil k-space data to be reconstructed. The acquisition model simplifies to $g = \mathcal{P}z + \epsilon$. SPIRiT assumes local redundancy in multi-coil k-space: missing samples can be expressed as linear combinations of their neighbors. Let $G \in \mathbb{R}^{cn \times cn}$ be the interpolation kernel estimated from ACS data. The self-consistency condition is $z = Gz$, which, when combined with undersampled data, can be written as

$$(\mathcal{Q}z + g) = G(\mathcal{Q}z + g).$$

Hence, SPIRiT reconstructs z by solving the least-squares problem

$$\min_z \left\{ \frac{1}{2} \|(G - I)(\mathcal{Q}z + g)\|_2^2 \right\}, \tag{5}$$

where I is the identity matrix.

As in the SENSE case, accurate calibration of G is critical. Limited ACS lines or noise can cause kernel estimation errors, leading to inaccurate high-frequency k-space interpolation and loss of fine image details.

3.3 Empirical Priors for Regularization Parameters

The SENSE3d algorithm in (4) iteratively updates the regularization parameters Γ based on intermediate multi-coil reconstructions. Let $U = \mathcal{F}^\top(\mathcal{Q}\mathcal{F}\mathcal{S}u + g)$ denote the reconstructed multi-coil image at the current iterate. In prior work, Γ is adaptively updated as a function of U. To analyze this behavior, we conduct the experiment shown in Figure 2: the top row presents reconstructions at different iterations alongside the truth target image; the middle row shows the corresponding regularization parameters Γ_t and Γ_{Target}; and the bottom row reports reconstructions obtained using each Γ_t and Γ_{Target}. Qualitatively, artifacts are progressively reduced as t increases, and image obtained with Γ_{Target} are nearly artifact-free in the zoomed-in regions.

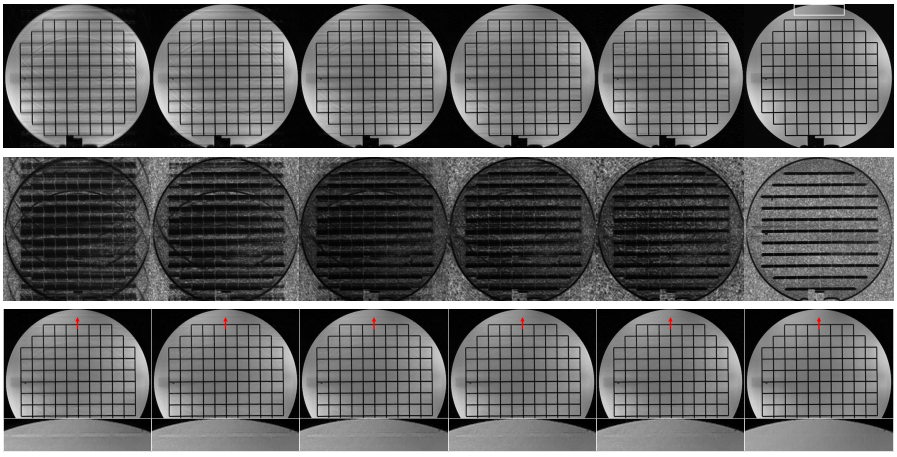


Fig. 2 Top row: Reconstructed images at iteration t and target image. Middle row: Corresponding regularization parameters Γ computed at each iteration. Bottom row: Reconstruction results using different Γ settings, along with zoomed-in regions. The zoomed-in area corresponds to the white box in the top row of the target Image. The red arrow indicates a noticeable artifact that is gradually removed across iterations.

Table 1 reports PSNR, SSIM, DISTS [33], and HaarPSI [34] for different Γ_t. All metrics improve monotonically with t, with the best performance achieved using

Γ_{Target}. This indicates that regularization parameters derived from artifact-free, detail-preserving images can significantly enhance reconstruction quality, motivating the learning of such parameters from improved reconstructions.

Table 1 Quantitative comparison of reconstruction performance under different regularization parameters Γ, measured by PSNR, SSIM, DISTS, and HaarPSI. The best results across different numbers of iterations t are highlighted in bold.

Reg. Para.	PSNR↑	SSIM↑	DISTS↓	HaarPSI↑
Γ_0	30.9271	0.9065	0.1168	0.8038
Γ_5	31.2161	0.9086	0.1045	0.8347
Γ_{10}	31.3928	0.9087	0.0994	0.8449
Γ_{15}	31.5852	0.9091	0.0945	0.8573
Γ_{20}	**31.6441**	**0.9109**	**0.0929**	**0.8638**
Γ_{Target}	**34.4253**	**0.9332**	**0.0902**	**0.9224**

4 Bilevel Optimization with Deep Priors

We now present the proposed bilevel optimization framework and the integration of deep priors for regularization parameters, with the overall algorithmic architecture shown in Figure 3.

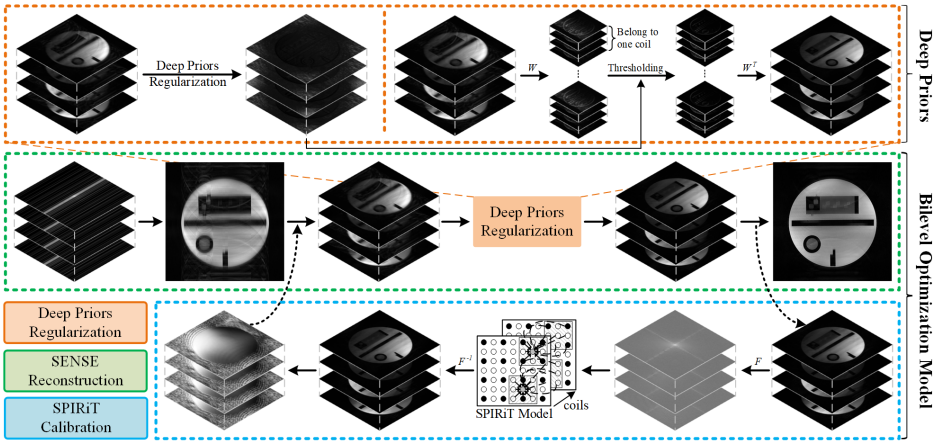


Fig. 3 The proposed bilevel optimization model with deep-prior-guided regularization follows a two-step strategy. We use the SENSE-based model for image reconstruction, while the SPIRiT-based model is employed to calibrate the coil sensitivity maps in k-space. Additionally, deep learning-driven techniques are used to provide deep priors for regularization parameter setting, which are then applied to the SENSE-based reconstruction process.

4.1 Bilevel Optimization for Reconstruction

The upper-level component reconstructs the target slice image u in the SENSE model, whereas the lower-level component operates on the multi-coil k-space variable z in the SPIRiT model. These two variables are linked through

$$z = \mathcal{F}\mathcal{S}u,$$

which allows the two subproblems to exchange information through the current estimate of the coil sensitivities.

To formalize this coupling, we introduce the multi-coil image $x = \mathcal{F}^\top z = \mathcal{S}u$, whose blocks correspond to the individual coil images. From x, we define a mapping $x \mapsto \mathcal{S}(x)$ that constructs the corresponding diagonal CSM matrices by normalizing the coil images pixelwise. This mapping provides a natural bridge between the image-domain reconstruction variable and the k-space calibration variable. Concretely, we write $\mathcal{S}(x) = [\mathcal{S}_1(x)^\top, \cdots, \mathcal{S}_c(x)^\top]^\top \in \mathbb{C}^{cn\times n}$, where each $\mathcal{S}_\ell(x) \in \mathbb{C}^{n\times n}$ is a diagonal matrix whose entries are given by

$$\mathcal{S}_\ell(x)[i,i] = \frac{x_\ell[i]}{\sqrt{\sum_{j=1}^{c} |x_j[i]|^2}}, \qquad i = 1, 2, \ldots, n. \tag{6}$$

In other words, $\mathcal{S}(x)$ encodes the discrete CSMs as diagonal blocks whose i-th diagonal entry is the normalized coil image value at pixel i. This mapping converts a multi-coil image x into the corresponding $\mathcal{S}(x)$. Conversely, given $\mathcal{S}$ and a slice image u, the multi-coil image is $x = \mathcal{S}u$.

Using the above notation, we now couple the SENSE-based image-domain reconstruction with the SPIRiT-based k-space calibration in a bilevel way. The key idea of SRSC is to use the current multi-coil image estimate $\tilde{x}$ to define $\mathcal{S}(\tilde{x})$, reconstruct a slice image $\tilde{u}$ through the SENSE3d objective, and then refine the multi-coil image by enforcing SPIRiT self-consistency while keeping it close to the SENSE-predicted coil images. In practice, the algorithm alternates between these two subproblems so that the image estimate, the multi-coil image, and the induced CSMs are progressively refined. Formally, we propose the following bilevel model:

$$\begin{cases} \tilde{u} \in \arg\min\limits_{u\in\mathbb{R}^n} \left\{ \dfrac{1}{2}\|\mathcal{P}\mathcal{F}\mathcal{S}(\tilde{x})u - g\|_2^2 + \|\Gamma W \mathcal{F}^\top(\mathcal{Q}\mathcal{F}\mathcal{S}(\tilde{x})u + g)\|_1 \right\}, \\ \hat{x} = \mathcal{S}(\tilde{x})\,\tilde{u}, \\ \tilde{x} \in \arg\min\limits_{x\in\mathbb{C}^{cn}} \left\{ \dfrac{1}{2}\|(G-I)(\mathcal{Q}\mathcal{F}x + g)\|_2^2 + \dfrac{\eta}{2}\|x - \hat{x}\|_2^2 \right\}. \end{cases} \tag{7}$$

Here, the upper-level problem reconstructs the slice image $\tilde{u}$ using SENSE3d with CSMs given by $\mathcal{S}(\tilde{x})$, where $\tilde{x}$ is the current multi-coil image estimate. The intermediate quantity $\hat{x} = \mathcal{S}(\tilde{x})\,\tilde{u}$ is the multi-coil image predicted by the SENSE model. The

lower-level problem then updates the multi-coil image $\tilde{x}$ by combining SPIRiT self-consistency in k-space with a quadratic penalty that pulls x toward $\hat{x}$. We refer to (7) as the SENSE Reconstruction with SPIRiT Calibration (SRSC) model.

The SENSE3d subproblem [9] can be solved by various existing algorithms. In this work, we adopt the Proximal Dual Three-Operator Splitting (PD3O) algorithm [35]. Choosing step sizes $\rho = 1.999$ and $\delta = 0.499$ ensures the convergence [9]. Defining $A = W\mathcal{F}^\top \mathcal{Q}\mathcal{F}\mathcal{S}$ and $w = W\mathcal{F}^\top g$, the PD3O iterations read

$$u^{k+1} = \mathrm{Re}(v^k), \tag{8a}$$
$$y^{k+1} = (I - \rho\delta AA^\top)s^k + \delta A\left(\bar{v}^k - \rho\mathcal{S}^\top\mathcal{F}^\top\left(\mathcal{P}\mathcal{F}\mathcal{S}u^k - g\right)\right), \tag{8b}$$
$$s^{k+1} = \left(y^{k+1} + \delta w\right) - \mathrm{soft}\left(y^{k+1} + \delta w, \Gamma\right), \tag{8c}$$
$$v^{k+1} = u^k - \rho\mathcal{S}^\top\mathcal{F}^\top\left(\mathcal{P}\mathcal{F}\mathcal{S}u^k - g\right) - \rho A^\top s^{k+1}. \tag{8d}$$

Here, $\mathrm{Re}(\cdot)$ denotes the real-part operation, $\bar{v}^k$ denotes the complex conjugate of v^k, and $\mathrm{soft}(\cdot, \Gamma)$ is the element-wise soft-thresholding operator with thresholds given by the diagonal of Γ. The initialization is given by $v^0 = u^0 = \mathrm{SoS}(\mathcal{F}^{-1}g)$, $s^0 = Au^0$, and $\mathcal{S}$ is initialized using the mapping defined in (3). A single iteration is denoted by

$$(u^{k+1}, s^{k+1}, v^{k+1}) = \mathrm{PD3O}(u^k, s^k, v^k).$$

We adopt the parameter choices and convergence conditions from SENSE3d [9]. In implementation, we set the maximum number of PD3O iterations to $K_{\max}$ (discussed in Section 4.2.3) and terminate the outer iteration when the infinity-norm difference between two successive image iterates, namely $\Delta = \|u^{k+1} - u^k\|_\infty$, falls below 10^{-4}.

The SPIRiT-based calibration subproblem is smooth and strictly convex, and can therefore be solved efficiently by the conjugate gradient (CG) method. In practice, we apply CG to the linear system

$$\left(\mathcal{F}^\top\mathcal{Q}(G-I)^\top(G-I)\mathcal{Q}\mathcal{F} + \eta I\right)x = \eta\hat{x} - \mathcal{F}^\top\mathcal{Q}(G-I)^\top(G-I)g. \tag{9}$$

To avoid hyperparameter tuning, we simply fix $\eta = 1$, and we denote the corresponding CG solver by $x = \mathrm{CG}(\hat{x})$. Given the fast convergence rate of CG, we set its maximum number of inner iterations to 5 in subsequent experiments.

The resulting SRSC algorithm is summarized in Algorithm 1. The maximum number of iterations $K_{\max}$ and the SPIRiT-based calibration is applied periodically (every K_{csm} iterations) to update CSMs via $\mathrm{CG}(\hat{x})$ followed by the mapping $\mathcal{S}(\cdot)$. The regularization parameters Γ follow the design in earlier SENSE3d work, and convergence is discussed in Appendix A.

4.2 SRSC with Deep Priors

We next augment SRSC with deep-prior-guided regularization derived from score-based generative models (SGM). The empirical findings in Section 3.3 indicate that regularization parameters estimated from high-quality images can substantially

Algorithm 1 SRSC algorithm for solving model (7)

Inputs: Step sizes $\rho = 1.999$ and $\delta = 0.499$; $K_{\max}$, K_{csm}.
Initialization: Set $\mathcal{S}$ using the mapping in (3); initialize $u^0 = v^0 = \text{SoS}(\mathcal{F}^{-1}g)$, and $s^0 = Au^0$; initialize $k = 0$ and $\Delta = +\infty$.
while $k < K_{\max}$ and $\Delta > 10^{-4}$ **do**
    if $k \bmod K_{\text{csm}} = 0$ **and** $k > 0$ **then** *// SPIRiT-based calibration of CSMs*
        $\hat{x} = \mathcal{S}u^k$
        $\tilde{x} = \text{CG}(\hat{x})$
        $\mathcal{S} = \mathcal{S}(\tilde{x})$
    end if
    $(u^{k+1}, s^{k+1}, v^{k+1}) = \text{PD3O}(u^k, s^k, v^k)$ *// PD3O update*
    $\Delta = \|u^{k+1} - u^k\|_\infty$
    $k = k + 1$
end while
Output: Reconstructed MRI slice $\hat{u} = u^{k+1}$.

improve reconstruction. Our goal is to automate this process using a learned denoiser and a principled feature-based mapping to Γ.

4.2.1 Score-Based Generative Priors

Let $p(u)$ denote the distribution of clean MRI images. A score-based generative model aims to learn the score function $\nabla_u \log p(u)$, which points toward regions of higher data density [26]. Because $p(u)$ is not directly available, we instead consider noisy samples $\tilde{u} = u + \sigma\epsilon$, where $\epsilon \sim \mathcal{N}(0, I)$, and learn the score of the corresponding conditional density through denoising score matching:

$$\nabla_{\tilde{u}} \log p_\sigma(\tilde{u}|u) = \frac{u - \tilde{u}}{\sigma^2}.$$

We then train a neural network $s_\Theta(\tilde{u})$, parameterized by Θ, to approximate this score by minimizing the denoising score-matching loss

$$\mathcal{L}_\Theta = \mathbb{E}_{u\sim p(u),\tilde{u}\sim p_\sigma(\tilde{u}|u)} \left[\frac{1}{2} \left\| s_\Theta(\tilde{u}) - \frac{u - \tilde{u}}{\sigma^2} \right\|_2^2 \right]. \tag{10}$$

Because the injected noise is relatively small, the marginal distribution of $\tilde{u}$ remains concentrated near the true data manifold, so the learned score s_Θ provides a good approximation to the underlying data distribution.

Once trained, the SGM can be used as a powerful denoiser. In our framework, the current reconstruction is treated as a noisy intermediate image, and the pretrained network $\mathcal{N}_\Theta$ is applied to obtain a refined image

$$\hat{u} = \mathcal{N}_\Theta(u).$$

This denoised image serves as a higher-quality surrogate for estimating the regularization weights used in the subsequent reconstruction step.

4.2.2 From Denoised Images to Regularization Parameters

Given the denoised image $\hat{u}$, we construct spatially adaptive regularization parameters that remain fully compatible with the classical 3DHTF-based SENSE3d regularizer. Specifically, $\hat{u}$ is used to extract multiscale, multi-coil framelet-domain statistics, which are then mapped to local weights that adaptively modulate the ℓ_1 regularization strength. The procedure has four steps: denoising, 3DHTF decomposition, local variation estimation, and energy normalization.

1. **Denoising.** We apply an SGM-based denoiser to the current reconstruction, $\hat{u} = \mathcal{N}_\Theta(u)$, where u may still contain aliasing and noise due to imperfect CSMs and kernel calibration. Acting as a data-driven denoiser, $\mathcal{N}_\Theta$ suppresses artifacts and restores fine structures using learned MRI priors while preserving global anatomy. The enhanced image $\hat{u}$ therefore provides a more reliable surrogate of the ground truth for estimating regularization parameters.
2. **3DHTF decomposition.** Next, we extract multiscale and multi-directional features from $\hat{u}$ in the same framelet domain used by the SENSE3d regularizer. To ensure consistency with the data-fidelity term and the actual sampling pattern, we form a multi-coil image by inserting $\hat{u}$ into the forward model and combining it with the acquired data:
$$\iota = W\mathcal{F}^\top(\mathcal{Q}\mathcal{F}\mathcal{S}\hat{u} + g).$$
Here, $\mathcal{Q}\mathcal{F}\mathcal{S}\hat{u} + g$ denotes the hybrid k-space data that incorporates the acquired measurements g. Applying $\mathcal{F}^\top$ maps this hybrid k-space data back to the multi-coil image domain. The operator W represents the 3DHTF transform applied across both spatial dimensions and coil channels. Consequently, the coefficient vector $\iota \in \mathbb{C}^{Lcn}$ captures local edges, textures, and directional features in a representation that is fully consistent with the regularization term in (4).
3. **Local variation estimation.** To turn the 3DHTF coefficients into spatially varying weights, we estimate local variation in the framelet domain. We first reshape $\iota \in \mathbb{C}^{Lcn} \mapsto \zeta \in \mathbb{C}^{L\times c\times h\times w}$, so that $\zeta_o^\ell[i,j]$ denotes the coefficient at filter index o, coil ℓ, and spatial location (i,j). For each filter, coil, and location, we compute the average magnitude over a 3×3 neighborhood:
$$\varsigma_o^\ell[i,j] = \frac{1}{\#\{(\tau_1,\tau_2)\}} \sum_{\tau_1,\tau_2\in\{-1,0,1\}} \left|\zeta_o^\ell[i+\tau_1,\, j+\tau_2]\right|.$$
This local average magnitude serves as a simple, translation-invariant measure of local framelet energy: large values indicate regions with strong edges or textures in the corresponding subband, while small values correspond to smoother, more homogeneous areas. By vectorizing $\varsigma \in \mathbb{R}^{L\times c\times h\times w}$, we obtain a diagonal matrix ϱ. For convenience, we define
$$\varrho = H(\iota) \in \mathbb{C}^{Lcn\times Lcn},$$

where $H(\cdot)$ is a mapping that assigns the local variation estimates in ς to the diagonal entries of ϱ, while all off-diagonal elements remain zero.

4. **Energy normalization.** Finally, for the ℓ-th filter of 3DHTF, we normalize the raw weights in $\varrho_\ell \in \mathbb{C}^{cn\times cn}$ and match their overall scale to the energy level of the current coefficients. Let $\|\iota_\ell\|_\infty$ and $\|\varrho_\ell\|_\infty$ denote the maximum magnitude among all entries of ι_ℓ and ϱ_ℓ, respectively. We define

$$\Gamma_\ell = \begin{cases} 0, & \text{if } \ell \text{ is low-pass filter or } \|\varrho_\ell\|_\infty = 0, \\ \|\iota_\ell\|_\infty \cdot \dfrac{\varrho_\ell}{\|\varrho_\ell\|_\infty}, & \text{otherwise.} \end{cases}$$

Here, scalar–matrix multiplication and division are implemented via element-wise operations. By stacking the regularization parameter matrices Γ_ℓ calculated from different filters, the final unified Γ can be obtained. In this way, Γ inherits the spatial adaptivity and multi-scale structure of the framelet coefficients but remains on a comparable scale across iterations and slices.

Because Γ enters the SENSE3d regularizer through $\|\Gamma W F^\top (QFSu + g)\|_1$, it directly controls how strongly each local framelet coefficient is penalized. Regions in which the denoised image $\hat{u}$ exhibits prominent structure lead to distinct patterns in ϱ (and hence in Γ), allowing the algorithm to modulate shrinkage in a feature-aware manner. Conversely, in smooth or low-contrast areas, the weights can promote stronger regularization to suppress noise and residual artifacts. For the sake of notational simplicity, we define the aforementioned SGM-Guided Adaptive Regularization Parameter Estimate (SGarpe) as

$$\Gamma = \mathrm{SGarpe}(u).$$

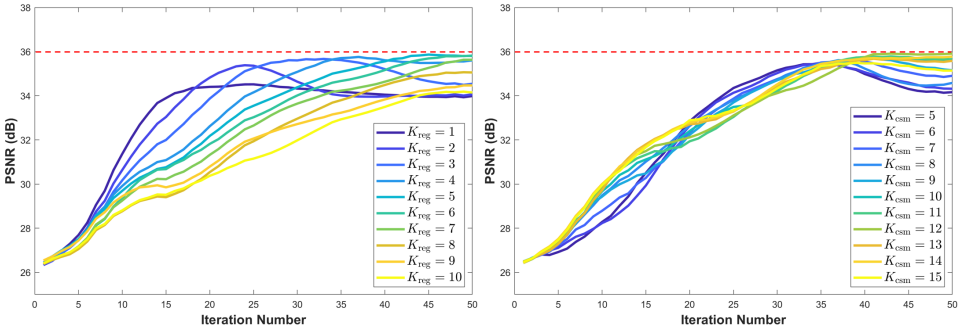
This procedure yields an SGM-guided, reconstruction-oriented regularization-parameter network: the classical linear operator W is kept fixed, while its associated weights Γ are adapted from denoised reconstructions via local framelet statistics. This preserves the structure and convergence properties of the SENSE3d regularizer while injecting powerful deep priors into the choice of regularization parameters.

4.2.3 Parameter Study and the SRSC+ Algorithm

Three hyperparameters are used in SRSC+: the maximum number of outer iterations $K_{\max}$, the update frequency K_{reg} for the regularization parameters, and the update frequency K_{csm} for the coil sensitivity maps. Empirically, PD3O converges rapidly in this setting, and $K_{\max} = 50$ is sufficient to achieve stable reconstruction quality across the tested scenarios.

To study K_{reg} and K_{csm}, we evaluate how the PSNR evolves over iterations under different update schedules. In Figure 4(a), K_{csm} is fixed at 12 while K_{reg} varies from 1 to 10. In Figure 4(b), K_{reg} is fixed at 5 while K_{csm} varies from 5 to 15. The experiments are conducted on a phantom dataset (image size 512×512 with 4 coils), and multiple independent trials are performed under different sampling patterns. Based on the

resulting performance–efficiency trade-off, we use $K_{\text{reg}} = 5$ and $K_{\text{csm}} = 12$ in all subsequent experiments.



(a) Effect of K_{reg} on PSNR (fixed $K_{\text{csm}} = 12$) (b) Effect of K_{csm} on PSNR (fixed $K_{\text{reg}} = 5$)

Fig. 4 Parameter sensitivity analysis of update frequencies. The PSNR evolution curves under different update frequencies are shown. (a) Influence of the regularization update frequency K_{reg} with K_{csm} fixed. (b) Influence of the coil sensitivity map update frequency K_{csm} with K_{reg} fixed. The results demonstrate that appropriate choices of update frequencies can balance reconstruction quality and computational efficiency.

Algorithm 2 SRSC+ algorithm for solving model (7)

Inputs: Step sizes $\rho = 1.999$ and $\delta = 0.499$; $K_{\max} = 50$, $K_{\text{reg}} = 5$, $K_{\text{csm}} = 12$.
Initialization: Set $\mathcal{S}$ using the mapping in (3); initialize $u^0 = v^0 = \text{SoS}(\mathcal{F}^{-1}g)$, and $s^0 = Au^0$. initialize $k = 0$ and $\Delta = +\infty$.
while $k < K_{\max}$ and $\Delta > 10^{-4}$ **do**
  if $k \bmod K_{\text{reg}} = 0$ **then** *// Deep prior regularization*
    $\Gamma = \text{SGarpe}(u^k)$
  end if
  if $k \bmod K_{\text{csm}} = 0$ **and** $k > 0$ **then** *// SPIRiT-based calibration of CSMs*
    $\hat{x} = \mathcal{S}u^k$
    $\tilde{x} = \text{CG}(\hat{x})$
    $\mathcal{S} = \mathcal{S}(\tilde{x})$
  end if
  $(u^{k+1}, s^{k+1}, v^{k+1}) = \text{PD3O}(u^k, s^k, v^k)$ *// PD3O update*
  $\Delta = \|u^{k+1} - u^k\|_\infty$
  $k = k + 1$
end while
Output: Reconstructed MRI slice $\hat{u} = u^{k+1}$.

Algorithm 2 summarizes the full SRSC+ procedure. Every K_{reg} iterations, the SGM-based module updates the regularization weights; every K_{csm} iterations, the SPIRiT-based calibration step updates the CSMs; and at every iteration, PD3O updates the image and auxiliary variables. This schedule decouples the time scales

of regularization-parameter adaptation and sensitivity-map refinement while allowing both to improve the reconstruction progressively. In principle, the convergence behavior of SRSC+ can be analyzed within a Kurdyka–Łojasiewicz-type framework, provided the inexact inner updates are appropriately controlled.

4.3 Ablation Study

We now examine the contributions of the three main components of our framework: SENSE-based Reconstruction (SR), SPIRiT-based Calibration (SC), and Deep Priors Regularization (DP). For comparison, we also consider a plug-and-play Regularization by Denoising (RED) scheme [23, 24], in which the explicit classical linear regularizer is replaced by a deep denoiser.

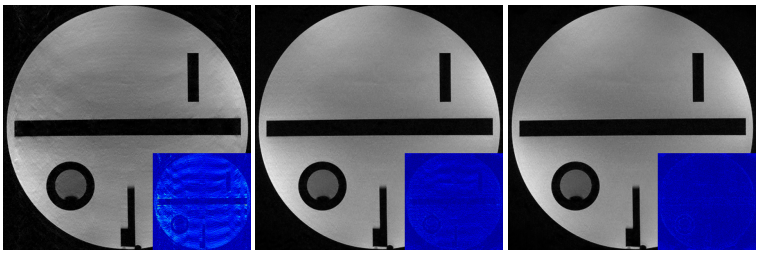
4.3.1 SR, SC, and DP

We perform 100 randomized experiments on non-training datasets under a Cartesian phase-encoding sampling pattern. Reconstruction quality is evaluated using PSNR, SSIM, DISTS, and HaarPSI. Table 2 reports mean metrics across all trials.

Adding SC to SR (i.e., moving from SR to SRSC) yields substantial gains in all metrics, confirming the benefit of bilevel CSMs calibration. Incorporating DP on top of SRSC (i.e., SRSC+) further improves performance and achieves the best overall results, with particularly strong gains in perceptual metrics (DISTS and HaarPSI).

Table 2 Ablation study comparing SR, SC, and DP methods, evaluated on PSNR, SSIM, DISTS, and HaarPSI. Bold values indicate the best performance for each metric.

	Module			**Metrics**			
Model	SR	SC	DP	PSNR↑	SSIM↑	DISTS↓	HaarPSI↑
SR	✓			31.5860	0.8566	0.1211	0.8347
SRSC	✓	✓		35.2092	0.9150	0.0796	0.9080
SRSC+	✓	✓	✓	**35.4965**	**0.9203**	**0.0737**	**0.9207**



(a) SR (b) SRSC (c) SRSC+

Fig. 5 Reconstruction results of a Phantom slice under a random phase-encoding sampling pattern with a sampling rate of 22% and 27 ACS lines. Panels (a)–(c) show the reconstructed images and corresponding error maps with respect to the target MRI for SR, SRSC, and SRSC+, respectively.

Figure 5 illustrates phantom reconstructions, which serve as fully OOD data relative to the SGM's fastMRI training set. Pure SENSE-based reconstruction (SR) exhibits pronounced residual aliasing and noise. Adding SPIRiT-based calibration (SRSC) significantly reduces artifacts, while SRSC+ with deep-prior-guided regularization produces the lowest error maps and best preserves phantom structures, demonstrating robust generalization to OOD cases.

4.3.2 Deep Priors vs PnP-RED

To isolate the effect of deep priors on the regularization parameters, we compare our DP strategy with a PnP-RED scheme. For fairness, both approaches use the same SGM-based denoiser. The RED-based formulation and its implementation within the SRSC framework follow standard constructions.

We consider two baselines, SR and SRSC, and report results over 100 randomized OOD experiments in Table 3. In all scenarios, DP consistently outperforms RED across all metrics. These results indicate that adapting regularization parameters via deep priors is more accurate than replacing the regularizer with a PnP denoiser.

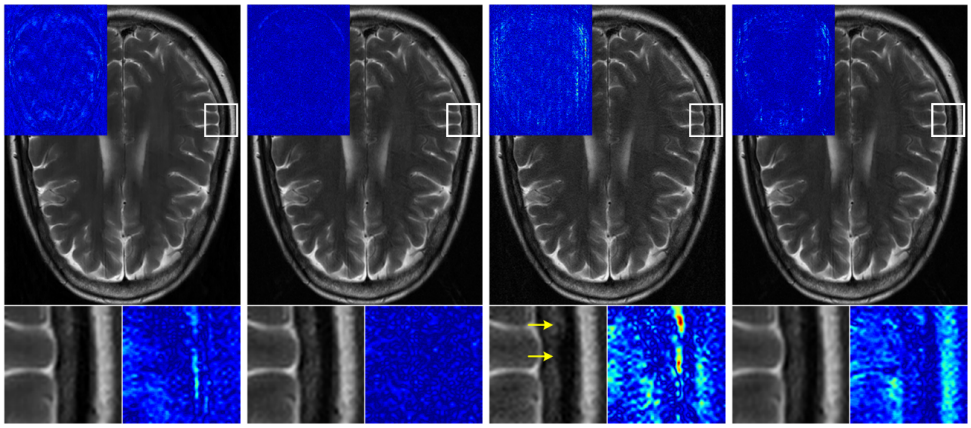
Table 3 Ablation Study comparing the DP and RED methods is conducted under both baseline models. The evaluation is performed using PSNR, SSIM, DISTS, and HaarPSI. Bold values highlight the best performance for each metric.

	Module		**Metrics**			
Baseline	DP	RED	PSNR↑	SSIM↑	DISTS↓	HaarPSI↑
SR	✓		**34.6731**	**0.9157**	**0.1007**	**0.8929**
SR		✓	31.2532	0.8358	0.1226	0.8461
SRSC	✓		**36.1452**	**0.9261**	**0.0684**	**0.9318**
SRSC		✓	32.1402	0.9124	0.0986	0.8555

Figure 6 compares DP and RED on OOD data. Under identical baselines, DP consistently yields lower reconstruction errors than RED. For SR-based reconstructions, (a) and (c), RED introduces visible pseudo-structures, whereas DP produces cleaner results with fewer hallucinations. For SRSC-based reconstructions, (b) and (d), both methods improve; but DP shows more pronounced gains. In particular, RED tends to oversmooth images and suppress fine textures, while DP better preserves anatomical details. These results demonstrate the advantage of retaining a classical linear regularizer and using deep learning only for parameter modulation, and further validate the robustness of the proposed CSMs update strategy.

5 Experiments and Results

We now present quantitative and qualitative experiments that compare the proposed SRSC+ algorithm with several classical and state-of-the-art methods. We first describe the experimental setup, then report computational efficiency, numerical performance



(a) SR-DP (b) SRSC-DP (c) SR-RED (d) SRSC-RED

Fig. 6 Reconstruction results of a brain slice under a random phase-encoding sampling pattern with a sampling rate of 28% and 27 ACS lines. Panels (a) SR with DP, (b) SRSC with DP, (c) SR with RED, and (d) SRSC with RED. Each panel includes the full reconstruction error map, along with a locally zoomed-in region (indicated by the white box) and its corresponding error map. Yellow arrows in panel (c) highlight pseudo-structures introduced by the RED method.

across multiple OOD datasets and sampling patterns, and detailed visual comparisons. Throughout Section 5, the target image refers to the SoS reconstruction from fully sampled k-space data.

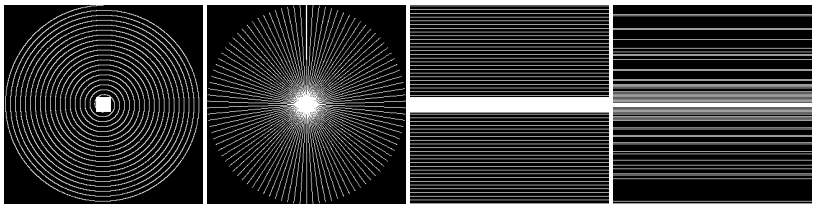
5.1 Experiment Setup

- **Network training:** The SGM-based denoiser is trained on the fastMRI dataset, using SoS-combined slices as training images. We use 128×128 patches, train for 500 epochs with Adam, and decay the initial learning rate of 0.005 by a factor of 0.1. Because the U-Net is fully convolutional, it can be applied to images of different sizes at test time through adaptive patch-based inference.
- **Evaluation metrics:** Reconstruction quality is evaluated using four complementary metrics: PSNR and SSIM as standard reference metrics, together with DISTS and HaarPSI as perceptual metrics. All reported results are averaged over repeated experiments.
- **Sampling masks:** We consider four undersampling masks: spiral, radial, uniform, and random, as shown in Figure 7. Spiral and radial masks use sampling rates of 15%–20% together with a fully sampled 24×24 k-space center for CSM estimation, whereas the Cartesian 1D masks (uniform and random) use sampling rates of 20%–30% with 10–70 ACS lines.
- **Datasets:** Experiments are conducted on multiple datasets, including fastMRI[1], the NYU dataset, the Stanford dataset[2], and additional 3T MRI data reported in [9]. The fastMRI training set is used to train the deep learning components, whereas the

[1]The fastMRI dataset is available at: http://fastmri.med.nyu.edu/.
[2]The NYU Machine Learning and Stanford MRI datasets are available at: http://mridata.org/.

fastMRI test set and all non-fastMRI datasets are used exclusively for evaluation. The non-fastMRI datasets are treated as fully OOD test sets.

- **Compared methods:** We compare SRSC+ with ESPIRiT [31], SENSE3d [9], MoDL [36][3], TRPA [21][4], and FRSGM [26][5]. For the deep learning baselines, we use the authors' official implementations and pretrained models with their default hyperparameters. Both SENSE3d and SRSC+ initialize CSMs using the ACS-based mapping (3), while the other methods use CSMs estimated by ESPIRiT. All methods follow a unified training and evaluation protocol, and no additional re-tuning is performed for the OOD experiments. The undersampling masks and preprocessing pipelines are kept consistent across methods to ensure fair comparison.



(a) Spiral (b) Radial (c) Uniform (d) Random

Fig. 7 Sampling masks used in the simulations. From left to right: (a) spiral, (b) radial, (c) uniform, and (d) random, respectively.

5.2 Efficiency Comparison

Table 4 reports the runtime and memory consumption of different methods evaluated on a Tesla P100 GPU with 16 GB RAM. The experiments are conducted on three datasets with representative dimensions: fastMRI-Brain ($768 \times 396 \times 20$), NYU-Knee ($768 \times 616 \times 15$), and Stanford-Foot ($320 \times 512 \times 8$), where $h \times w \times c$ denotes image height, width, and the number of coils.

While SRSC+ incurs moderately higher memory usage due to the 3DHTF-based multi-coil feature extraction, it achieves competitive or superior runtime compared with other deep learning–based methods, especially on larger 2D slices. This demonstrates that the proposed bilevel framework, with periodic calibration and deep-prior-guided updates, remains computationally practical.

5.3 Numerical Evaluation

We benchmark all methods on the OOD datasets under four sampling modes. For spiral and radial masks, the sampling rates range from 15% to 20%; for uniform and random Cartesian masks, the rates range from 20% to 30% with varying ACS lines. Table 5 summarizes average PSNR, SSIM, DISTS, and HaarPSI across more than 50 randomized trials for each configuration. Key observations include:

[3] Available at: https://github.com/hkaggarwal/modl.
[4] Available at: https://github.com/Houruizhi/TRPA.
[5] Available at: https://github.com/Houruizhi/FRSGM.

Table 4 Running time (in seconds) / memory consumption (in MB) of the compared algorithms on different datasets, evaluated on a Tesla P100 GPU.

	Model-Driven		**Model-Based Deep Learning-Driven**			
Dataset	ESPIRiT	SENSE3d	MoDL	TRPA	FRSGM	SRSC+
fastMRI-Brain	58/1274	28/5117	62/14658	85/1275	158/3405	22/3851
NYU-Knee	77/1492	33/5973	80/15846	118/1835	209/3928	24/4491
Stanford-Foot	52/280	6/1103	30/13142	68/594	91/1144	5/960
# Iterations	50	50	10	100	100	50

- Across sampling patterns, SRSC+ achieves the strongest overall quantitative performance.
- The gains are especially clear in the perceptual metrics DISTS and HaarPSI, indicating that the proposed framework improves visual fidelity in addition to standard reconstruction accuracy.
- Under Cartesian 1D phase encoding, where limited ACS data make CSM initialization more difficult, SRSC+ remains particularly robust because the bilevel calibration mechanism progressively corrects sensitivity errors during reconstruction.

Table 5 PSNR, SSIM, DISTS, and HaarPSI metrics of reconstructed images on the OOD datasets introduced in Section 5.1, using ESPIRiT, SENSE3d, MoDL, TRPA, FRSGM, and the proposed SRSC+ model under spiral, radial, uniform, and random sampling modes. Bold indicates the best result for each metric.

		Model-Driven		**Model-Based Deep Learning-Driven**			
Mode	**Metrics**	**ESPIRiT**	**SENSE3d**	**MoDL**	**TRPA**	**FRSGM**	**SRSC+**
Spiral	PSNR	31.3687	32.9280	31.0189	33.4784	33.4253	**34.0530**
	SSIM	0.7432	0.8611	0.7362	0.8083	0.8081	**0.8965**
	DISTS	0.1574	0.1331	0.1624	0.1237	0.1269	**0.0977**
	HaarPSI	0.8287	0.8505	0.8298	0.8744	0.8709	**0.8860**
Radial	PSNR	32.4346	34.3875	31.6120	33.5822	33.5439	**36.5859**
	SSIM	0.7090	0.8445	0.6894	0.7526	0.7526	**0.9160**
	DISTS	0.1683	0.1373	0.1885	0.1369	0.1400	**0.0956**
	HaarPSI	0.9154	0.9095	0.9064	0.9104	0.9073	**0.9594**
Uniform	PSNR	32.1303	33.2594	31.9817	32.8024	32.7518	**35.4008**
	SSIM	0.7144	0.8821	0.7132	0.7404	0.7398	**0.9187**
	DISTS	0.1469	0.1209	0.1504	0.1420	0.1445	**0.0910**
	HaarPSI	0.8315	0.8529	0.8314	0.8314	0.8281	**0.9017**
Random	PSNR	30.8020	35.1189	29.6815	32.8815	32.8230	**36.3579**
	SSIM	0.7980	0.9216	0.7648	0.8634	0.8636	**0.9385**
	DISTS	0.1475	0.0933	0.1685	0.1112	0.1155	**0.0583**
	HaarPSI	0.8507	0.9185	0.8416	0.8572	0.8531	**0.9496**

Figure 8 further organizes the results by sampling rate. Here, '•' denotes the 2D sampling patterns (spiral and radial) with sampling rates ranging from 15% to 20%, while '⋆' denotes the Cartesian 1D sampling patterns (uniform and random) with sampling rates ranging from 20% to 30%. To reduce sensitivity to very small rate differences, we group the sampling rates into intervals: "15–16", "17–18", and "19–20" for the 2D patterns, and "21–23", "24–26", and "27–29" for the Cartesian 1D patterns. This analysis shows that SRSC+ not only performs well at a fixed sampling rate but also improves consistently as the sampling rate increases, suggesting that the method is not brittle with respect to sampling conditions.

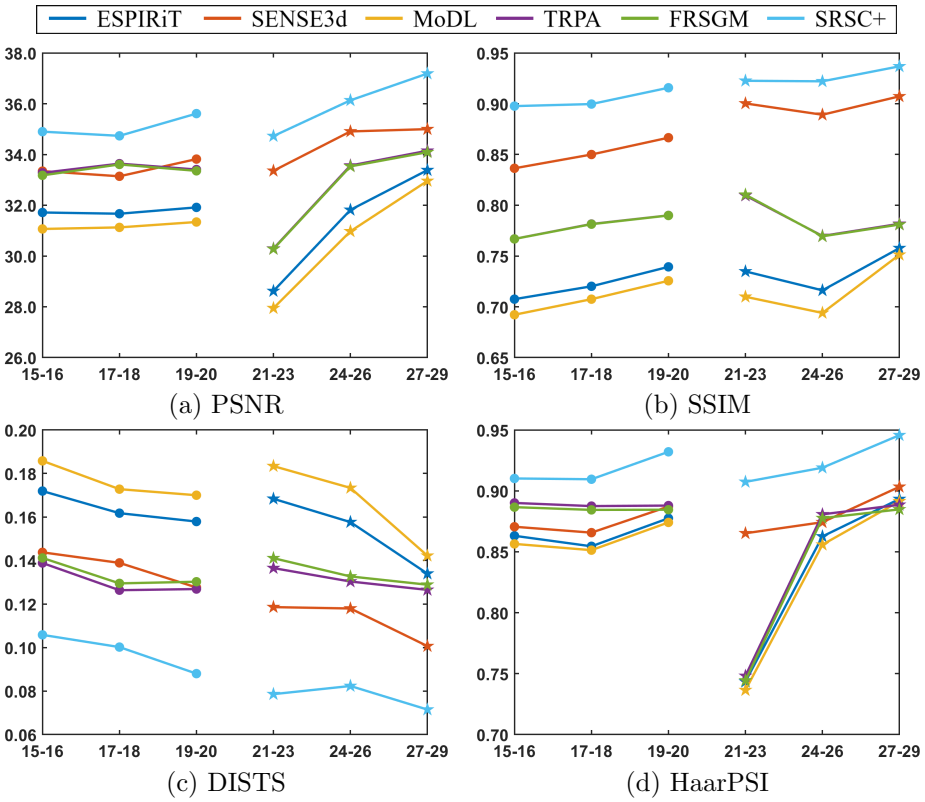


Fig. 8 Performance of different algorithms under '•' Cartesian 2D (spiral and radial) and '⋆' Cartesian 1D (uniform and random) sampling patterns: (a) PSNR, (b) SSIM, (c) DISTS, and (d) HaarPSI as functions of the sampling rate (x-axis). The results are derived from the quantitative experiments reported in Table 5.

5.4 Visualization Comparison

We also provide visual comparisons under Cartesian phase encoding, which closely resembles clinical acquisition, focusing on NYU knee and 3T MRI brain datasets.

NYU knee slice. Figure 9 presents reconstructions of a knee slice under random phase encoding with 25% sampling and 30 ACS lines. The OOD nature of this dataset stresses model generalization:

- The model-driven algorithms ESPIRiT and SENSE3d provide faithful reconstructions overall; however, both still exhibit some loss of fine textural detail in the zoomed-in regions.
- ESPIRiT and MoDL suffer from noticeable noise amplification and reconstruction artifacts around the knee joint, as indicated by the blue arrow.
- The RED-based algorithms TRPA and FRSGM demonstrate significant texture degradation in muscle and soft-tissue regions, as highlighted by the green, yellow, and red arrows.
- SRSC+ better preserves the textures of muscles and vessels while avoiding obvious pseudo-structures, highlighting the advantages of learned regularization parameters within a stable linear regularization framework.

3T MRI brain slice. Figure 10 shows a representative brain slice reconstructed under uniform phase encoding with a 40% sampling rate and 70 ACS lines. All methods produce reasonable global reconstructions, but the error maps reveal clear differences:

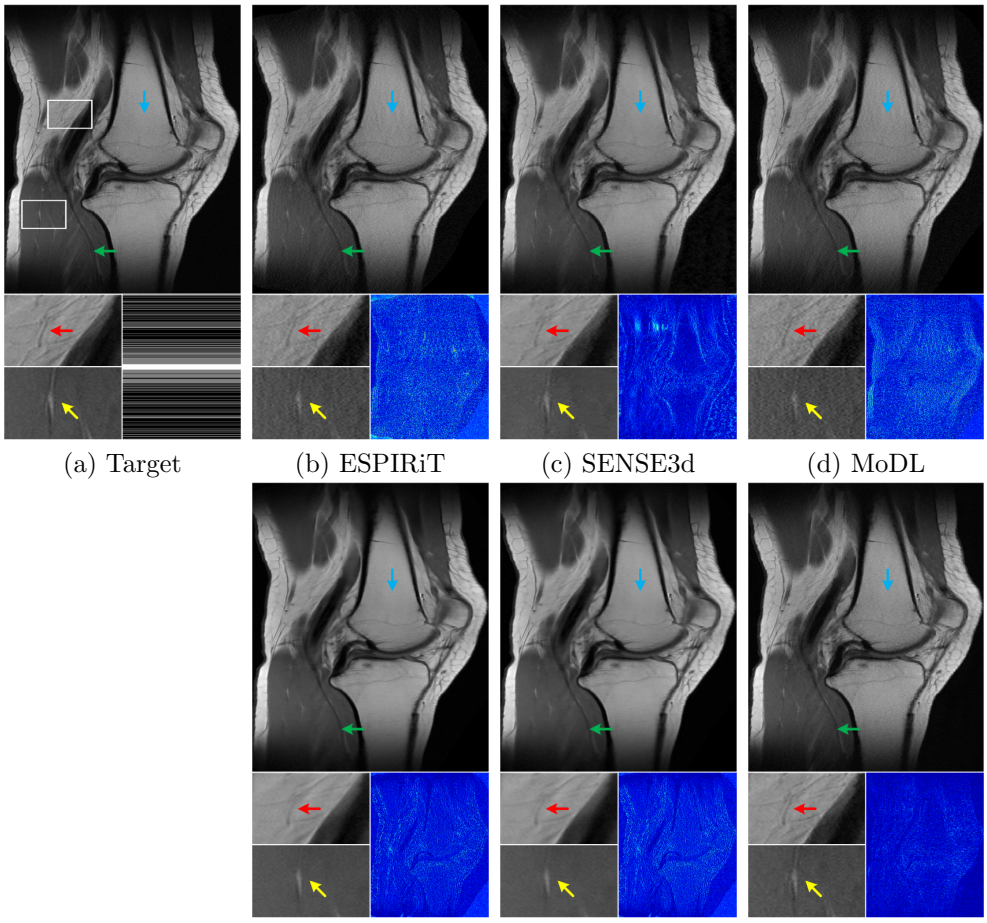
- ESPIRiT, SENSE3d, and MoDL exhibit clearly visible residual artifacts, as indicated by the green arrows, which lead to reconstruction errors in the zoomed-in regions highlighted by the red arrows.
- Although TRPA and FRSGM remove artifacts using advanced deep denoisers, excessive denoising results in over-smoothing and loss of structural information, as shown by the yellow arrow in the zoomed-in region.
- The proposed SRSC+ matches the target image more closely than the other methods, demonstrating superior reconstruction accuracy.

Overall, the visual comparisons support the quantitative results: SRSC+ generalizes well to OOD data and preserves anatomical structure without introducing obvious pseudo-structures. The quantitative metrics from the two visualization experiments are summarized in Table 6, showing that the proposed SRSC+ achieves the best performance.

6 Conclusion

We have proposed SRSC+, a model-driven bilevel framework for parallel MRI reconstruction that integrates SENSE-based image reconstruction, SPIRiT-based k-space calibration, and deep-prior-guided regularization. By decoupling sensitivity estimation from k-space calibration, the framework reduces error accumulation across domains and enables iterative refinement of the CSMs under limited or noisy ACS data.

Building on empirical observations from SENSE3d, SRSC+ further uses an SGM-based denoiser to derive spatially adaptive regularization weights while preserving the structure of a classical linear regularizer. This hybrid design combines interpretability and stability with the expressive power of learned priors.



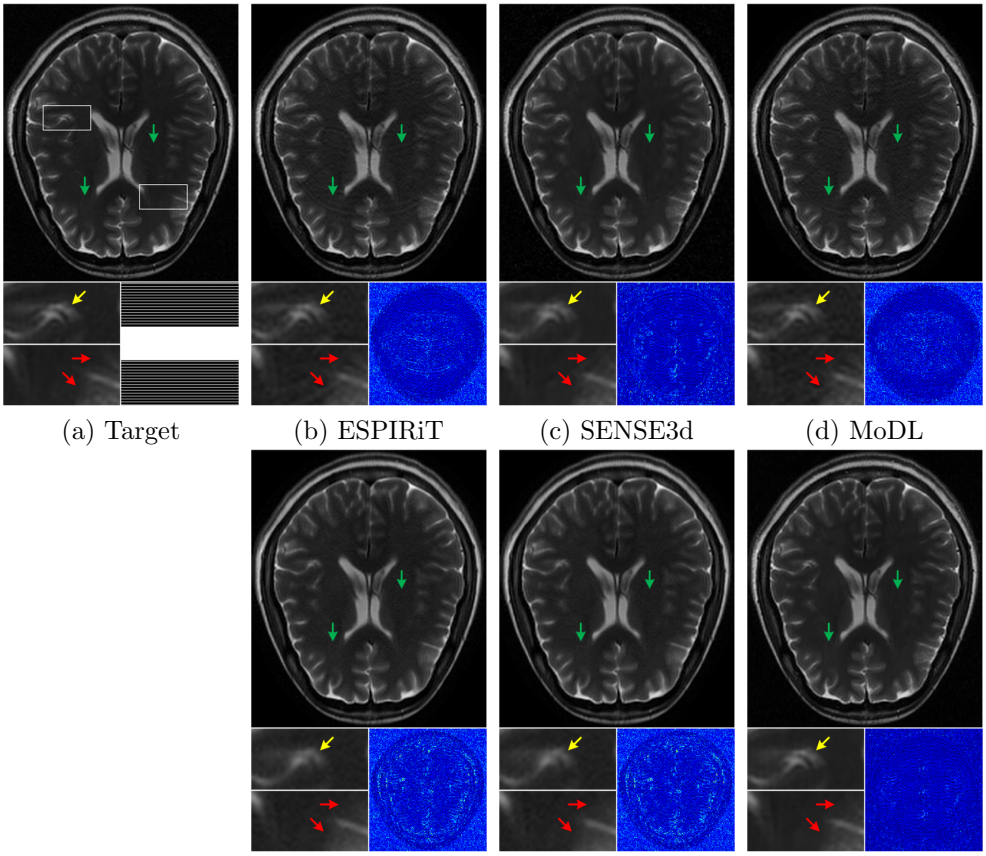
(a) Target (b) ESPIRiT (c) SENSE3d (d) MoDL

(e) TRPA (f) FRSGM (g) SRSC+

Fig. 9 Visual comparison of a NYU knee slice under random phase-encoding sampling pattern with a 25% sampling rate and 30 ACS lines. Panels (a)–(g) correspond to the target image, ESPIRiT, SENSE3d, MoDL, TRPA, FRSGM, and SRSC+, respectively. Each panel includes an error map and two zoomed-in regions indicated by white boxes for detailed comparison.

Extensive experiments and ablation studies show that SRSC+ achieves strong quantitative and perceptual performance across multiple datasets and sampling patterns, while remaining robust on OOD data. These results suggest that bilevel sensitivity correction combined with deep-prior-guided parameter adaptation is a promising direction for reliable model-driven MRI reconstruction.

Future work includes extending SRSC+ to multi-contrast and dynamic MRI, incorporating uncertainty quantification for CSMs and kernel estimates, and exploring task-aware training strategies in which the deep priors are optimized not only for reconstruction metrics but also for downstream clinical tasks.



(a) Target (b) ESPIRiT (c) SENSE3d (d) MoDL

(e) TRPA (f) FRSGM (g) SRSC+

Fig. 10 Visual comparison of a 3T MRI brain slice under uniform phase-encoding sampling pattern with a 40% sampling rate and 70 ACS lines. Panels (a)–(g) correspond to the target image, ESPIRiT, SENSE3d, MoDL, TRPA, FRSGM, and SRSC+, respectively. Each panel includes an error map and two zoomed-in regions indicated by white boxes for detailed comparison.

Table 6 PSNR, SSIM, DISTS, and HaarPSI metrics of reconstructed images on the NYU knee and fastMRI brain experiments shown in Figure 9 and Figure 10, using ESPIRiT, SENSE3d, MoDL, TRPA, FRSGM, and the proposed SRSC+ model. Bold indicates the best result for each metric.

		Model-Driven		**Model-Based Deep Learning-Driven**			
Figure	**Metrics**	**ESPIRiT**	**SENSE3d**	**MoDL**	**TRPA**	**FRSGM**	**SRSC+**
Figure 9	PSNR	31.7549	35.7590	29.7170	35.2066	34.9770	**36.6436**
	SSIM	0.8594	0.9299	0.8255	0.9082	0.9082	**0.9360**
	DISTS	0.1107	0.0695	0.1325	0.0837	0.0861	**0.0548**
	HaarPSI	0.9035	0.9277	0.8931	0.9135	0.9085	**0.9459**
Figure 10	PSNR	35.7757	38.6495	35.9708	35.1550	35.1656	**41.4010**
	SSIM	0.5647	0.9110	0.5848	0.5645	0.5642	**0.9475**
	DISTS	0.1724	**0.0695**	0.1723	0.1651	0.1561	0.0780
	HaarPSI	0.9437	0.9659	0.9470	0.9414	0.9427	**0.9849**

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (NSFC 12471400) and the Shenzhen Science and Technology Program (JCYJ20230808105610021); HKRGC Grants No. LU13300125 and LU BGR 105824; the National Science Foundation under grant DMS-2208385; The work of L. Shen was supported in part by the National Science Foundation under grant DMS-2208385; The work of X. Zhuang was supported by the Research Grants Council of Hong Kong (Project No. CityU 11302023, CityU 11301224, and CityU 11300825) and a grant from the Innovation and Technology Commission of Hong Kong (Project No. MHP/054/22).

Appendix A Convergence Proof for SRSC

In this section, we provide a rigorous convergence proof for the proposed SRSC model. We analyze the iterative process used to solve the bilevel optimization problem and establish the conditions under which the algorithm converges. Specifically, we show that the alternating scheme employed in model (7) converges to a stationary point under appropriate assumptions.

We define the following joint energy (or surrogate objective) in (u, x):

$$\begin{aligned}\mathcal{J}(u,x) :=& \frac{1}{2}\left\|\mathcal{P}\mathcal{F}\mathcal{S}(x)u - g\right\|_2^2 + \left\|\Gamma W \mathcal{F}^\top\left(\mathcal{Q}\mathcal{F}\mathcal{S}(x)u + g\right)\right\|_1 + \\ & \frac{1}{2}\left\|(G-I)(\mathcal{Q}\mathcal{F}x + g)\right\|_2^2 + \frac{\eta}{2}\left\|x - \mathcal{S}(x)u\right\|_2^2.\end{aligned} \tag{A1}$$

For fixed x, the first two terms in (A1) reduce to a SENSE3d-type objective in u. For fixed u, the last two terms form a SPIRiT-based calibration problem in x with a quadratic coupling to $\mathcal{S}(x)u$. In this sense, $\mathcal{J}$ is consistent with the bilevel formulation discussed in model (7).

We collect the assumptions used in the analysis.

Assumption A1 (Bounded linear operators). All linear operators $\mathcal{P}, \mathcal{Q}, \mathcal{F}, W, G$ are bounded, and $\mathcal{F}$ is unitary (up to normalization). The SPIRiT operator G is such that $G - I$ is bounded.

Assumption A2 (Regularity of the CSMs mapping). The CSMs mapping $\mathcal{S}(\cdot)$ is continuously differentiable and locally Lipschitz on the region of interest; in particular, whenever $\sum_{\ell=1}^{c} |x_\ell[i]|^2$ is bounded away from zero, the entries of $\mathcal{S}(x)$ and their first derivatives are bounded.

Assumption A3 (Coercivity and KL property). The function $\mathcal{J}$ is coercive, i.e.,

$$\|(u,x)\|_2 \to \infty \quad \Longrightarrow \quad \mathcal{J}(u,x) \to +\infty,$$

and $\mathcal{J}$ satisfies the Kurdyka–Łojasiewicz (KL) property at every point of its domain.

Assumption A3 is standard in KL-based convergence analysis. In our setting, $\mathcal{J}$ is a finite sum of compositions of polynomials, rational, and ℓ_1-type functions with linear

and smooth mappings. This makes $\mathcal{J}$ semi-algebraic (or more generally definable in an o-minimal structure), hence it has the KL property; see, e.g., [37].

We now record some basic consequences of these assumptions.

Lemma 1 (Basic properties of $\mathcal{J}$) Under Assumptions A1–A3, the function $\mathcal{J} : \mathbb{C}^n \times \mathbb{C}^{cn} \to \mathbb{R} \cup \{+\infty\}$ defined in (A1) satisfies:

1. $\mathcal{J}$ is proper and lower semicontinuous;
2. Every sublevel set $\{(u, x) : \mathcal{J}(u, x) \le a\}$ is bounded for any $a \in \mathbb{R}$;
3. $\mathcal{J}$ has the KL property at every point of its domain.

Proof Properness and lower semicontinuity follow from continuity of the quadratic terms and lower semicontinuity of the ℓ_1-norm, plus Assumption A2 on $\mathcal{S}(\cdot)$. Coercivity in Assumption A3 implies boundedness of sublevel sets. The KL property follows from the semi-algebraic (or definable) structure of $\mathcal{J}$; see, for example, [37]. □

For the convergence analysis, we consider an *idealized* SRSC scheme in which each subproblem is solved exactly. Starting from an initial pair (u^0, x^0), the algorithm generates a sequence $\{(u^k, x^k)\}_{k\in\mathbb{N}}$ as follows:

Algorithm A.1 (Idealized SRSC). For $k = 0, 1, 2, \ldots$

1. **(u-step)** Given x^k, set
$$u^{k+1} \in \arg\min_{u\in\mathbb{C}^n} \mathcal{J}(u, x^k).$$
This is a convex but nonsmooth problem in u (SENSE3d-type objective).
2. **(x-step)** Given u^{k+1}, set
$$x^{k+1} \in \arg\min_{x\in\mathbb{C}^{cn}} \mathcal{J}(u^{k+1}, x).$$
This is a smooth, strongly convex problem in x (SPIRiT calibration plus quadratic coupling) with a unique minimizer.

In the implemented Algorithm 1, the u-step is carried out by PD3O iterations and the x-step by a few conjugate gradient iterations on the linear system associated with (A1). For the theoretical analysis, we work with Algorithm A.1 and then briefly comment on the practical scheme at the end of this appendix.

Lemma 2 (Well-posedness of the block subproblems) Under Assumptions A1–A3, the following hold for every fixed k:

1. For any fixed x^k, the u-subproblem $u \mapsto \mathcal{J}(u, x^k)$ is convex and has at least one minimizer. If in addition the operator $\mathcal{PFS}(x^k)$ has trivial null space (up to constants), the minimizer is unique.
2. For any fixed u^{k+1}, the x-subproblem $x \mapsto \mathcal{J}(u^{k+1}, x)$ is smooth and strongly convex, hence has a unique minimizer.

Proof For fixed x^k, the terms in $\mathcal{J}(u, x^k)$ involving u consist of a quadratic data-fidelity term and an ℓ_1-norm composed with a linear operator, which is convex. Coercivity implies the existence of minimizers, and uniqueness follows from the strict convexity of the quadratic part under mild conditions on $\mathcal{PFS}(x^k)$.

For fixed u^{k+1}, the terms in $\mathcal{J}(u^{k+1}, x)$ involving x are smooth quadratics: $\frac{1}{2}\|(G - I)(\mathcal{QF}x + g)\|_2^2$ and $\frac{\eta}{2}\|x - \mathcal{S}(x^k)u^{k+1}\|_2^2$. The first term defines a positive semidefinite quadratic form and the second adds strict convexity because $\eta > 0$. Consequently, $x \mapsto \mathcal{J}(u^{k+1}, x)$ is strongly convex, whence it has a unique minimizer. □

The next lemma records the basic descent property of Algorithm A.1.

Lemma 3 (Descent and finite-length property) Let $\{(u^k, x^k)\}$ be the sequence generated by Algorithm A.1. Under Assumptions A1–A3, we have:

1. $\mathcal{J}(u^{k+1}, x^{k+1}) \le \mathcal{J}(u^k, x^k)$ for all k, and the sequence $\{\mathcal{J}(u^k, x^k)\}$ is convergent;
2. The sequence $\{(u^k, x^k)\}$ has bounded sublevel sets and satisfies

$$\sum_{k=0}^{\infty}\Big(\|u^{k+1} - u^k\|_2^2 + \|x^{k+1} - x^k\|_2^2\Big) < +\infty.$$

Proof By definition of u^{k+1},

$$\mathcal{J}(u^{k+1}, x^k) \le \mathcal{J}(u^k, x^k),$$

since u^{k+1} minimizes $\mathcal{J}(\cdot, x^k)$. Similarly, x^{k+1} minimizes $\mathcal{J}(u^{k+1}, \cdot)$, so

$$\mathcal{J}(u^{k+1}, x^{k+1}) \le \mathcal{J}(u^{k+1}, x^k).$$

Combining the two inequalities yields

$$\mathcal{J}(u^{k+1}, x^{k+1}) \le \mathcal{J}(u^k, x^k), \qquad \forall k.$$

By Lemma 1 and Assumption A3, $\mathcal{J}$ is bounded from below, so $\{\mathcal{J}(u^k, x^k)\}$ converges to some limit $\mathcal{J}^\star$.

The finite-length property follows from standard arguments in the analysis of block coordinate descent for KL functions; see, e.g., [38, Theorem 3.2] and [39, Theorem 3.1]. Intuitively, the strict convexity of each subproblem yields a sufficient decrease condition $\mathcal{J}(u^k, x^k) - \mathcal{J}(u^{k+1}, x^{k+1}) \ge c\|(u^{k+1}, x^{k+1}) - (u^k, x^k)\|^2$ for some $c > 0$, from which the summability of the squared increments follows. □

We are now ready to state the main convergence result for the idealized SRSC algorithm.

Theorem 1 (Convergence of Algorithm A.1) *Let Assumptions A1–A3 hold, and let $\{(u^k, x^k)\}$ be the sequence generated by Algorithm A.1. Then:*

1. *The sequence $\{(u^k, x^k)\}$ is bounded and has finite length, i.e., $\sum_{k=0}^{\infty}\|(u^{k+1}, x^{k+1}) - (u^k, x^k)\|_2 < +\infty$.*
2. *The sequence $\{(u^k, x^k)\}$ converges to a limit $(u^\star, x^\star)$.*

3. *The limit* $(u^\star, x^\star)$ *is a critical point of* $\mathcal{J}$*, in the sense that*

$$0 \in \partial \mathcal{J}(u^\star, x^\star),$$

where ∂ *denotes the limiting subdifferential. Equivalently,* $(u^\star, x^\star)$ *satisfies the first-order optimality conditions of the SENSE3d subproblem in* u *and the SPIRiT-calibration subproblem in* x*.*

Proof Boundedness of the sequence and the finite-length property follow from Lemma 1 and Lemma 3. The KL property of $\mathcal{J}$ and the sufficient decrease and relative error conditions implied by the exact minimization of each block (Lemma 2) place Algorithm A.1 within the framework of block coordinate descent on KL functions. The conclusion then follows from standard results on the global convergence of such methods, see, for example, [38, Theorem 3.2] and [39, Theorem 3.1]. □

The implemented SRSC algorithm differs from Algorithm A.1 in two respects:

- the u-subproblem is solved approximately by a fixed number of PD3O iterations with fixed $\mathcal{S}$ and Γ;
- the x-subproblem is solved approximately by a fixed number of conjugate gradient iterations on the normal equations.

In practice, these inexact solves still produce a decreasing sequence of energies and empirically stable behavior. A fully rigorous treatment of the inexact case can be obtained by invoking KL-based convergence results for inexact block coordinate or PALM-type methods, provided that the inner errors are suitably controlled (e.g., summable). For clarity of presentation, we restrict the formal analysis here to the idealized exact-solve setting.

Appendix B SRSC with RED

When the PnP-based Regularization by Denoising (RED) is adopted and the explicit classical linear regularizer is replaced by a deep denoiser, the bilevel SRSC model (7) degenerates into:

$$\begin{cases} \tilde{u} \in \arg\min\limits_{u \in \mathbb{R}^n} \left\{ \dfrac{1}{2} \|\mathcal{P}\mathcal{F}\mathcal{S}(\tilde{x})u - g\|_2^2 + \lambda \mathcal{R}(u) \right\}, \\ \hat{x} = \mathcal{S}(\tilde{x})\, \tilde{u}, \\ \tilde{x} \in \arg\min\limits_{x \in \mathbb{C}^{cn}} \left\{ \dfrac{1}{2} \|(G - I)(\mathcal{Q}\mathcal{F}x + g)\|_2^2 + \dfrac{\eta}{2} \|x - \hat{x}\|_2^2 \right\}. \end{cases} \tag{B2}$$

Here, λ denotes the regularization parameter and $\mathcal{R}(\cdot)$ represents an implicitly defined regularization function. The upper-level u-subproblem can be solved using the ADMM-PnP method [23, 24, 40]. By constructing the augmented Lagrangian and performing variable splitting, the iterative update process can be formulated as follows:

$$v^{k+1} = \arg\min_{v} \left\{ \lambda \mathcal{R}(v) + \frac{\rho_k}{2} \|v - \hat{v}^k\|_2^2 \right\}, \tag{B3a}$$

$$u^{k+1} = \arg\min_u \left\{ \frac{1}{2}\|\mathcal{P}\mathcal{F}\mathcal{S}(\tilde{x})u - g\|_2^2 + \frac{\rho_k}{2}\|u - \hat{u}^k\|_2^2 \right\}, \tag{B3b}$$

$$\alpha^{k+1} = \alpha^k + u^{k+1} - v^{k+1}. \tag{B3c}$$

Here, $\hat{v}^k = u^k + \alpha^k$ and $\hat{u}^k = v^{k+1} - \alpha^k$. Based on the RED framework, the v-subproblem (B3a) can be solved using a deep denoiser. To ensure fairness in the ablation study, we employ the same SGM-based denoiser for reconstruction, i.e.,

$$v^{k+1} = \mathcal{N}_\Theta(\hat{v}^k).$$

For the u-subproblem (B3b), its smooth and differentiable form allows for an efficient solution using the Conjugate Gradient (CG) method, which corresponds to solving the following linear system:

$$(\mathcal{S}^\top \mathcal{F}^\top \mathcal{P}\mathcal{F}\mathcal{S} + \rho_k I)u = \rho_k \hat{u}^k + \mathcal{S}^\top \mathcal{F}^\top g. \tag{B4}$$

Based on the above preparation, an approximate solution to the upper-level subproblem can be obtained. Following the overall procedure described in Algorithm 1, the SRSC model with RED regularization is summarized in Algorithm 3, and we set $\rho_k \equiv 1$ to eliminate the need for tuning this hyperparameter.

Algorithm 3 The SRSC with RED algorithm for solving model (B2)

Ensure: $\hat{v}^k = u^k + \alpha^k$ and $\hat{u}^k = v^{k+1} - \alpha^k$.
for $k = 0, 1, \ldots$ **do**
  // SPIRiT-based Calibration for CSMs:
  if k mod $10 = 0$ and $k > 0$ **then**
    $\hat{x} = \mathcal{S}u^k$
    $\tilde{x} = \mathrm{CG}(\hat{x})$
    $\mathcal{S} = \mathcal{S}(\tilde{x})$
  end if
  $v^{k+1} = \mathcal{N}_\Theta(\hat{v}^k)$
  $u^{k+1} =$ Use CG to solve linear system (B4)
  $\alpha^{k+1} = \alpha^k + u^{k+1} - v^{k+1}$
end for

References

[1] Edelstein, W.A., Hutchison, J.M., Johnson, G., Redpath, T.: Spin warp NMR imaging and applications to human whole-body imaging. Physics in Medicine & Biology **25**(4), 751 (1980)

[2] Doneva, M.: Mathematical models for magnetic resonance imaging reconstruction: An overview of the approaches, problems, and future research areas. IEEE Signal Processing Pagazine **37**(1), 24–32 (2020)

[3] Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P.: SENSE: sensitivity encoding for fast MRI. Magnetic Resonance in Medicine **42**(5), 952–962 (1999)

[4] Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing MRI. IEEE Signal Processing Magazine **25**(2), 72–82 (2008)

[5] Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A.: Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magnetic Resonance in Medicine **47**(6), 1202–1210 (2002)

[6] Lustig, M., Pauly, J.M.: SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space. Magnetic Resonance in Medicine **64**(2), 457–471 (2010)

[7] Ye, X., Chen, Y., Huang, F.: Computational acceleration for MR image reconstruction in partially parallel imaging. IEEE Transactions on Medical Imaging **30**(5), 1055–1063 (2010)

[8] Chaâri, L., Pesquet, J.-C., Benazza-Benyahia, A., Ciuciu, P.: A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging. Medical Image Analysis **15**(2), 185–201 (2011)

[9] Li, Y.-R., Chan, R.H., Shen, L., Zhuang, X., Wu, R., Huang, Y., Liu, J.: Exploring structural sparsity of coil images from 3-dimensional directional tight framelets for SENSE reconstruction. SIAM Journal on Imaging Sciences **17**(2), 888–916 (2024)

[10] Bredies, K., Kunisch, K., Pock, T.: Total generalized variation. SIAM Journal on Imaging Sciences **3**(3), 492–526 (2010)

[11] Murphy, M., Alley, M., Demmel, J., Keutzer, K., Vasanawala, S., Lustig, M.: Fast ℓ_1-SPIRiT compressed sensing parallel imaging MRI: Scalable parallel implementation and clinically feasible runtime. IEEE Transactions on Medical Imaging **31**(6), 1250–1262 (2012)

[12] Li, Y.-R., Shen, L., Zhuang, X.: A tailor-made 3-dimensional directional Haar semi-tight framelet for pMRI reconstruction. Applied and Computational Harmonic Analysis **60**, 446–470 (2022)

[13] Aggarwal, H.K., Mani, M.P., Jacob, M.: Modl: Model-based deep learning architecture for inverse problems. IEEE Transactions on Medical Imaging **38**(2), 394–405 (2018)

[14] Li, Z., Xiao, S., Wang, C., Li, H., Zhao, X., Duan, C., Zhou, Q., Rao, Q., Fang, Y., Xie, J., *et al.*: Encoding enhanced complex CNN for accurate and highly accelerated MRI. IEEE Transactions on Medical Imaging **43**(5), 1828–1840

(2024)

[15] Duan, J., Ren, X.: Improved complex convolutional neural network based on SPIRiT and dense connection for parallel MRI reconstruction. IET Signal Processing **2024**(1), 7006156 (2024)

[16] Wang, B., Lian, Y., Xiong, X., Zhou, H., Liu, Z., Zhou, X.: DCT-net: dual-domain cross-fusion transformer network for MRI reconstruction. Magnetic Resonance Imaging **107**, 69–79 (2024)

[17] Sriram, A., Zbontar, J., Murrell, T., Zitnick, C.L., Defazio, A., Sodickson, D.K.: GrappaNet: Combining parallel imaging with deep learning for multi-coil MRI reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14315–14322 (2020)

[18] Wei, H., Li, Z., Wang, S., Li, R.: Undersampled multi-contrast MRI reconstruction based on double-domain generative adversarial network. IEEE Journal of Biomedical and Health Informatics **26**(9), 4371–4377 (2022)

[19] Cao, C., Cui, Z.-X., Wang, Y., Liu, S., Chen, T., Zheng, H., Liang, D., Zhu, Y.: High-frequency space diffusion model for accelerated MRI. IEEE Transactions on Medical Imaging **43**(5), 1853–1865 (2024)

[20] Cui, Z.-X., Cao, C., Wang, Y., Jia, S., Cheng, J., Liu, X., Zheng, H., Liang, D., Zhu, Y.: SPIRiT-diffusion: Self-consistency driven diffusion model for accelerated MRI. IEEE Transactions on Medical Imaging (2024)

[21] Hou, R., Li, F., Zhang, G.: Truncated residual based plug-and-play ADMM algorithm for MRI reconstruction. IEEE Transactions on Computational Imaging **8**, 96–108 (2022)

[22] Hou, R., Li, F.: Denoising knowledge transfer model for zero-shot MRI reconstruction. IEEE Transactions on Computational Imaging (2025)

[23] Venkatakrishnan, S.V., Bouman, C.A., Wohlberg, B.: Plug-and-play priors for model based reconstruction. In: 2013 IEEE Global Conference on Signal and Information Processing, pp. 945–948 (2013). IEEE

[24] Dong, W., Wang, P., Yin, W., Shi, G., Wu, F., Lu, X.: Denoising prior driven deep neural network for image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence **41**(10), 2305–2318 (2018)

[25] Shafique, M., Liu, S., Schniter, P., Ahmad, R.: MRI recovery with self-calibrated denoisers without fully-sampled data. Magnetic Resonance Materials in Physics, Biology and Medicine **38**(1), 53–66 (2025)

[26] Hou, R., Li, F., Zeng, T.: Fast and reliable score-based generative model for parallel MRI. IEEE Transactions on Neural Networks and Learning Systems **36**(1), 953–966 (2025)

[27] Muckley, M.J., Riemenschneider, B., Radmanesh, A., Kim, S., Jeong, G., Ko, J., Jun, Y., Shin, H., Hwang, D., Mostapha, M., *et al.*: Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. IEEE Transactions on Medical Imaging **40**(9), 2306–2317 (2021)

[28] Liu, X., Pang, Y., Jin, R., Liu, Y., Wang, Z.: Dual-domain reconstruction network with V-Net and K-Net for fast MRI. Magnetic Resonance in Medicine **88**(6), 2694–2708 (2022)

[29] Arshad, M., Najeeb, F., Khawaja, R., Ammar, A., Amjad, K., Omer, H.: Cardiac MR image reconstruction using cascaded hybrid dual domain deep learning framework. PLoS One **20**(1), 0313226 (2025)

[30] Crockett, C., Fessler, J.A., *et al.*: Bilevel methods for image reconstruction. Foundations and Trends® in Signal Processing **15**(2-3), 121–289 (2022)

[31] Uecker, M., Lai, P., Murphy, M.J., Virtue, P., Elad, M., Pauly, J.M., Vasanawala, S.S., Lustig, M.: ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA. Magnetic Resonance in Medicine **71**(3), 990–1001 (2014)

[32] Uecker, M., Lustig, M.: Estimating absolute-phase maps using ESPIRiT and virtual conjugate coils. Magnetic Resonance in Medicine **77**(3), 1201–1207 (2017)

[33] Ding, K., Ma, K., Wang, S., Simoncelli, E.P.: Image quality assessment: Unifying structure and texture similarity. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(5), 2567–2581 (2020)

[34] Reisenhofer, R., Bosse, S., Kutyniok, G., Wiegand, T.: A Haar wavelet-based perceptual similarity index for image quality assessment. Signal Processing: Image Communication **61**, 33–43 (2018)

[35] Yan, M.: A new primal–dual algorithm for minimizing the sum of three functions with a linear operator. Journal of Scientific Computing **76**, 1698–1717 (2018)

[36] Aggarwal, H.K., Mani, M.P., Jacob, M.: MoDL: model-based deep learning architecture for inverse problems. IEEE Transactions on Medical Imaging **38**(2), 394–405 (2018)

[37] Attouch, H., Bolte, J., Redont, P.: Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka–Łojasiewicz inequality. Mathematics of Operations Research **35**, 438–457 (2010)

[38] Attouch, H., Bolte, J.: On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. Mathematical Programming, Ser. A **116**, 5–16 (2009)

[39] Bolte, J., Sabach, S., Teboulle, M.: Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Mathematical Programming, Ser. A **146**, 459–494 (2014)

[40] Chan, S.H., Wang, X., Elgendy, O.A.: Plug-and-play ADMM for image restoration: Fixed-point convergence and applications. IEEE Transactions on Computational Imaging **3**(1), 84–98 (2016)