arXiv:2603.04438v3 [eess.IV] 16 Jun 2026

CogGen: Cognitive-Load-Inspired Fully Unsupervised Deep Generative Modeling for Compressively Sampled MRI Reconstruction

Qingyong Zhu, Yumin Tan, Xiang Gu and Dong Liang

***Abstract*—Fully unsupervised deep generative modeling (FU-DGM) offers significant potential for compressively sampled magnetic resonance imaging (CS-MRI) reconstruction. Representative FU-DGM formulations, such as deep image prior (DIP) and implicit neural representation (INR), employ architectural bias to induce a low-dimensional manifold in the image space that aligns with the forward observation. However, as the underlying inverse system is highly ill-posed, prolonged iterative fitting in FU-DGM typically leads to poor efficiency and noise amplification. In this paper, guided by the cognitive principle of easy-to-hard learning, we propose CogGen, an FU-DGM framework that reformulates CS-MRI reconstruction as a staged inversion problem. Specifically, CogGen implements an self-paced curriculum learning (SPCL)-driven progressive scheduling strategy through an MRI-aware dual-threshold weighting criterion, which adaptively regulates k-space measurement participation. The data-consistency residual thresholding evaluates the fitting reliability of the current generator, while the k-space radius thresholding controls stage-wise measurement exposure, thereby avoiding uniform fitting throughout optimization. Theoretically, our analysis shows that, when early stages favor easy-to-fit measurements, CogGen yields a reduced local sufficient-iteration bound and a smaller cumulative noise-amplification bound, explaining the improved convergence behavior and reconstruction fidelity of CogGen within a finite iteration budget. Numerical experiments demonstrate that both CogGen instantiations, CogGen-DIP and CogGen-INR, achieve superior performance over prevailing CS-MRI reconstruction techniques, including unsupervised and supervised pipelines.**

***Index Terms*—CS-MRI reconstruction, CogGen, cognitive load, progressive scheduling, dual-threshold weighting, SPCL**

I. Introduction

MAgnetic resonance imaging (MRI) is a leading medical-imaging modality that noninvasively visualizes internal tissues, providing critical information for assessing not only anatomical structures but also pathological conditions [1], [2]. However, protracted scan times due to hardware limitations impede scanner throughput and patient comfort, thereby hindering the deployment of advanced clinical applications such as multi-contrast [3] and real-time imaging [4]. Consequently, compressively sampled MRI (CS-MRI) [5]–[7], in which k-space data are deliberately acquired below the Nyquist rate, has long been a central topic of investigation; yet achieving accurate and reliable image reconstruction from such incomplete measurements remains a substantial challenge.

In recent years, deep generative modeling (DGM) [8], [9] has attracted increasing attention for CS-MRI reconstruction, frequently demonstrating notable advantages over conventional baselines based on variational or iterative regularization [10]–[13] through a generative parameterization that implicitly encodes a low-dimensional image manifold whose elements are consistent with the acquired k-space measurements. Existing DGM-based methods can be broadly categorized into two families. The first family [14]–[16] employs supervised DGM (S-DGM), wherein generators trained on extensive corpora pairing measurements with the corresponding ground-truths (GTs) learn a conditional generative mapping. Nevertheless, the dependence on task-matched, paired data imposes real-world constraints, particularly in specialized or resource-limited settings where GT labels are scarce, costly, or unattainable. The second family [17]–[21], namely unsupervised DGM (U-DGM), dispenses with paired datasets. In particular, fully unsupervised DGM (FU-DGM) represents a notable subcategory that requires no external training data and reconstructs each target image by exploiting the architectural bias of an untrained generator, thereby implicitly encoding image regularities such as spatial continuity, local correlations, and multiscale structures. Although classical FU-DGM formulations such as deep image prior (DIP) [22] and implicit neural representation (INR) [23] yield visually plausible reconstructions, they often suffer from suboptimal accuracy, especially in fine details, due to the amplification of noise corruption in ill-posed systems [24]. Additionally, the prolonged iterative procedures required by such systems lead to low inversion efficiency. Existing improvements typically stabilize FU-DGM reconstruction by refining the generator architecture [25], [26], introducing explicit regularization [27], [28] or applying early stopping [29]. Although these strategies can alleviate noise amplification or overfitting to some extent, they still provide limited control over the accuracy-efficiency tradeoff in highly ill-posed CS-MRI reconstruction.

To address the above issues, inspired by cognitive-load theory [30], [31], which formalizes the easy-to-hard learning principle in human cognition, we introduce a novel FU-DGM framework termed CogGen. CogGen reformulates CS-MRI reconstruction as a staged inversion problem, in

This work was supported in part by the National Natural Science Foundation of China under Grants 62125111 and 62331028; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515110476; and in part by the Guangdong Provincial Key Laboratory of Multimodality Non-Invasive Brain-Computer Interfaces under Grant 2024B1212010010. (Corresponding author: Dong Liang.)

Qingyong Zhu is with the Medical AI Research Centre, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China (qy.zhu@siat.ac.cn).

Yumin Tan is with the School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, Guangdong, China.

Xiang Gu is with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China.

Dong Liang is with the Paul C. Lauterbur Research Centre for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China (dong.liang@siat.ac.cn).

which the cognitive load imposed by heterogeneous k-space measurements, including sample-specific difficulty and noise-induced interference, is progressively scheduled to enhance the efficiency and robustness of scan-specific parameterization. Operationally, progressive scheduling is instantiated through a self-paced curriculum learning (SPCL) scheme [32]–[34] featuring an MRI-aware dual-threshold weighting criterion. Specifically, data-consistency residual thresholding evaluates the fitting reliability of the current generator, whereas k-space radius thresholding governs stage-wise measurement exposure. Explicitly regulating what to fit and when during reconstruction avoids uniform fitting throughout optimization, allowing early iterations to focus on structurally informative, low-complexity samples while gradually introducing more challenging high-frequency or noise-dominated measurements, thereby mitigating overfitting and accelerating practical inversion. Theoretically, we show that, with an early-stage preference for easy-to-fit measurements, CogGen admits a reduced local sufficient-iteration bound and a smaller cumulative noise-amplification bound within a finite iteration budget, offering a theoretical explanation for improved convergence behavior and image fidelity. Numerical experiments further demonstrate that both CogGen instantiations, CogGen-DIP and CogGen-INR, consistently outperform representative CS-MRI reconstruction techniques, including unsupervised and supervised pipelines.

The key contributions of our work are summarized as follows, covering the framework design, scheduling mechanism, theoretical analysis, and experimental validation.

1. We propose CogGen, a cognitive-load-inspired FU-DGM framework that reformulates CS-MRI reconstruction as a staged inversion problem and progressively schedules k-space measurement to improve the efficiency and accuracy of untrained generative reconstruction.
2. We design an MRI-aware SPCL-based k-space scheduling strategy equipped with a dual-threshold sample-weighting criterion, where data-consistency residual thresholding evaluates the fitting reliability of the current generator, whereas k-space radius thresholding defines a stage-wise feasible region to regulate measurement exposure.
3. We provide the local and conditional analyses showing that stage-wise measurement weighting can reduce the sufficient-iteration bound and suppress cumulative noise amplification during finite-iteration reconstruction.
4. We instantiate CogGen with DIP and INR, and extensive experiments demonstrate that the resulting CogGen-DIP and CogGen-INR outperform representative CS-MRI reconstruction techniques, including both unsupervised and supervised pipelines.

The remainder of this paper is organized as follows. Section II reviews the preliminaries of CS-MRI reconstruction and classical FU-DGM backbones. Section III presents the proposed CogGen framework, including the staged inversion formulation and the SPCL-driven k-space scheduling strategy. Section IV reports the experimental settings, comparative results, and ablation studies. Section V concludes the paper. Finally, the Appendix provides theoretical analyses of CogGen.

II. Preliminaries

A. CS-MRI Reconstruction

A discrete forward observation model for CS-MRI reconstruction can be mathematically expressed as

$$y = Ax + \epsilon, \tag{1}$$

where $x \in \mathbb{C}^N$ denotes the target MR image and $y \in \mathbb{C}^M$ represents the acquired undersampled k-space measurements. The observation operator $A : \mathbb{C}^N \to \mathbb{C}^M$ characterizes the MR acquisition process, typically including coil-sensitivity encoding, Fourier transformation, and k-space undersampling, with M denoting the total number of acquired measurements. Due to the inherent undersampling with $M \ll N$, the operator A is rank-deficient or severely ill-posed. The term ϵ denotes measurement noise, usually modeled as additive white Gaussian noise (AWGN). Therefore, CS-MRI reconstruction aims to recover a high-quality image x from the undersampled observation y.

B. Classical FU-DGM: DIP and INR

Definition 1 (Deep Network Prior [35], [36])**.** *A particular image x is deemed to follow an untrained neural network prior (i.e., a fully unsupervised generative prior) if it falls within a set S defined as $S := \{x|x = f_\theta(z)\}$, where z denotes the underlying latent representation driving the generation.*

We build upon two established FU-DGM frameworks, DIP and INR, both of which serve as deep network prior[1]-based image generators with architectural bias and employ the following objective function,

$$\hat{\theta} = \arg\min_{\theta} \parallel Af_\theta(z) - y \parallel_2^2,\ \hat{x} = f_{\hat{\theta}}(z) \tag{2}$$

where $f_\theta : \mathbb{R}^p \to \mathbb{C}^N$ is a parametric mapping with learnable parameters θ that generates a complex-valued image. Unlike supervised generative reconstruction which requires external training data, the parameters θ are optimized separately for each individual y. DIP injects randomness through z as a latent code, while INR treats z as a coordinate descriptor of the signal domain. As a result, classical FU-DGM formulations, despite yielding visually plausible results, often suffer from suboptimal fine-detail accuracy because prolonged iterative fitting in noisy and ill-posed systems can gradually amplify noise corruption in the reconstruction [37]–[39]. Specifically, we have

$$x^{(l+1)} = x^{(l)} - \eta A^{\mathrm{H}}\big(Ax^{(l)} - y\big), \quad \eta > 0 \tag{3}$$

Substituting $y = Ax^* + \epsilon$ yields

$$e^{(l+1)} = \big(I - \eta A^{\mathrm{H}}A\big)e^{(l)} + \eta A^{\mathrm{H}}\epsilon \tag{4}$$

where $e^{(l)} = x^{(l)} - x^*$. Let the singular value decomposition (SVD) be $A = U\Sigma V^{\mathrm{H}}$, with singular values $\{\sigma_i\}$. Projecting onto the i-th right singular vector v_i gives the scalar recursion

$$e_i^{(l+1)} = (1 - \eta\sigma_i^2)\, e_i^{(l)} + \eta\sigma_i\, \epsilon_i \tag{5}$$

where $e_i^{(l)} = v_i^{\mathrm{H}} e^{(l)}$, $\varepsilon_i = u_i^{\mathrm{H}}\epsilon$. Assuming $0 < \eta < 2/\sigma_{\max}^2$ so that $|1 - \eta\sigma_i^2| < 1$, the steady-state limit satisfies $e_i^{(\infty)} \approx \frac{\epsilon_i}{\sigma_i}$.

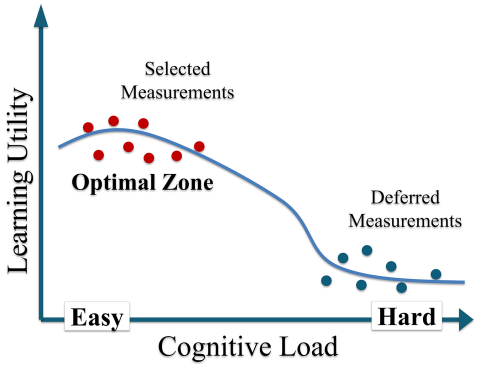

Fig. 1. Cognitive-load view of CogGen: measurements with lower load exhibit higher learning utility and are therefore selected earlier, whereas high-load measurements with reduced suitability are deferred to the further stages.

In CS-MRI, such small-singular-value directions are typically associated with weakly sampled or high-frequency components, where measurement noise and aliasing artifacts are more easily imprinted into the reconstruction. Furthermore, from an optimization perspective, the ill-conditioning of A, characterized by a large condition number $\kappa(A) = \sigma_{\max}/\sigma_{\min}$, also slows down iterative reconstruction, i.e., the components corresponding to extremely small singular values exhibit a convergence factor $|1 - \eta\sigma_i^2|$ very close to 1.

III. Proposed Framework: CogGen

In cognitive psychology, cognitive-load theory emphasizes that optimal learning utility is achieved by dynamically aligning task demands with the finite processing capacity of the learner to avoid information overload. Inspired by this principle, we propose CogGen, a cognitive-load-inspired FU-DGM framework that reformulates CS-MRI reconstruction as a staged inversion problem rather than a uniform iterative fitting task. In the context of CS-MRI, the cognitive load can be interpreted as the effective inversion burden imposed by heterogeneous k-space measurements, including sample-specific difficulty and noise-induced interference. Instead of estimating the target image by matching all undersampled k-space measurements uniformly throughout the optimization process, CogGen performs progressive scheduling over k-space measurements (As illustrated in Fig.1) and decomposes the original ill-posed inverse problem into a sequence of progressively refined subproblems, each associated with a controlled subset of measurements. This easy-to-hard inversion design [40], which mirrors human cognitive learning patterns, enables the model to first solve structurally simpler and more stable inversion objectives, and then gradually incorporate increasingly difficult components, thus matching the fitting burden to the evolving representation capacity of the generator while improving reconstruction quality within a finite iteration budget, and mitigating semi-convergence by delaying the fitting of noise-dominated measurements.

A. SPCL-based Scheduling Principle

Mathematically, progressive scheduling can be formulated as a sample-weighting problem [41]–[43] defined over the k-space measurements. Conventional uniform fitting assigns equal weights to all measurements, forcing the generator to match the acquired data with equal strength. However, this ignores the intrinsic heterogeneity of k-space: frequency-domain coefficients near the center encode dominant low-frequency anatomical structures with high SNR and stability, whereas peripheral high-frequency coefficients are highly vulnerable to undersampling-induced ill-posedness and noise contamination. Therefore, a scheduling scheme requires non-uniform, stage-dependent weighting. In this work, we adopt an MRI-aware dual-threshold weighting criterion-based SPCL strategy, which integrates curriculum learning [44] and self-paced learning [45] by combining a predefined easy-to-hard prior with adaptive sample selection according to the current learning state. Consequently, sample importance under soft-weighting or hard-selection regimes is assigned in a difficulty-aware and state-adaptive manner, enabling the generator to fit simpler and more reliable structures before gradually incorporating more challenging measurements.

B. Detailed Formulation

We first define a stage sequence over k-space,

$$C_1 \subset C_2 \subset \cdots \subset C_n \subset \cdots \subset C_T \tag{6}$$

where C_n denotes the set of acquired k-space measurements included in the n-th stage, and each measurement is treated as a schedulable sample. Thus, C_n specifies the measurements that are progressively exposed to the generator at stage n, rather than forcing all acquired measurements to participate with equal strength from the beginning. In general, the entire set of k-space measurements is covered only at the final stage. Accordingly, conditioned on the stage-wise weights $s_{C_n} \in [0,1]^M$, a regularized optimization problem is formulated over the network parameters θ_{C_n},

$$\begin{aligned}\hat{\theta}_{C_n} = \arg\min_{\theta_{C_n}} &\left\|s_{C_n} \circ \left(A f_{\theta_{C_n}}(z) - y\right)\right\|_2 / \|s_{C_n} \circ y\|_2 \\ &+ \gamma\left(\left\|\nabla_h f_{\theta_{C_n}}(z)\right\|_1 + \left\|\nabla_v f_{\theta_{C_n}}(z)\right\|_1\right)\end{aligned} \tag{7}$$

where the first term enforces normalized weighted data-fidelity and the second applies anisotropic total-variation (TV) constraint. $\circ$ denotes the Hadamard product. ∇_h and ∇_v denote the horizontal and vertical discrete gradient operators, respectively. The hyperparameter $\gamma > 0$ balances the data fidelity against the regularizer.

Consequently, we first define the stage-wise weighting vector s_{C_n} with the following closed-form rule. Here, the larger weight w emphasizes reliable measurements, whereas 1-w retains the remaining measurements with reduced influence; in particular, w = 1 corresponds to hard selection, while $w < 1$ yields soft weighting.

$$(s_{C_n})_i^* = w\mathbf{1}_{\mathcal{I}_{C_n}}(i) + (1-w)\mathbf{1}_{\mathcal{I}_{C_n}^c}(i),\ \frac{1}{2} < w \le 1 \tag{8}$$

with

$$\mathcal{I}_{C_n} = \left\{ i \in t_{C_n} \,\middle|\, \left|\left(A f_{\theta_{C_n}}(z)\right)_i - y_i\right| / |y_i| < \lambda_{C_n} \right\} \tag{9}$$

where $\mathbf{1}_{\mathcal{I}_{C_n}}$ is the indicator function of $\mathcal{I}_{C_n}$, yielding 1 if $i \in \mathcal{I}_{C_n}$ and 0 otherwise, and $\mathcal{I}_{C_n}^c$ denotes its set-theoretic

Algorithm 1 CogGen

Input: A, y, w, z, λ_{C_n}, r_{C_n}, $\Delta\lambda$, Δr, T, L

```
for n = 1 to T do
    Update s_{C_n} according to Eq. (8).
    for l = 1 to L(n) do
        Update θ_{C_n} according to Eq. (12).
    end for
    Update λ_{C_n}, r_{C_n} according to Eq. (11).
end for
```

Output: $\hat{x} = f_\theta(z)$

complement. The scalar λ_{C_n} serves as an adaptive threshold for data-consistency residual thresholding. t_{C_n} designates the feasible region, formally defined as the following index set,

$$t_{C_n} = \{i \mid \|\mathbf{k}_i - \mathbf{k}_{\text{center}}\|_2 < r_{C_n}\} \tag{10}$$

where $\mathbf{k}_i \in \mathbb{R}^2$ denotes the coordinate vector of the i-th k-space measurement, and $\mathbf{k}_{\text{center}} \in \mathbb{R}^2$ is the frequency-domain origin. r_{C_n} is a dynamic k-space radius that expands progressively across stages. Therefore, $\mathcal{I}_{C_n}$ contains only those measurements that simultaneously satisfy the feasible-region constraint $i \in t_{C_n}$ and the residual criterion, ensuring that a sample is strongly weighted only when it is both scheduled to be exposed and currently reliable to fit.

After each scheduling stage, the parameters λ_{C_n} and r_{C_n} are subjected to a deterministic update rule,

$$\begin{aligned} \lambda_{C_n} &= \lambda_{C_1} \left(\frac{\lambda_{C_T}}{\lambda_{C_1}}\right)^{\frac{n-1}{T-1}} \\ r_{C_n} &= r_{C_1} + \frac{n-1}{T-1}\left(r_{C_T} - r_{C_1}\right) \end{aligned} \tag{11}$$

where λ_{C_1} and r_{C_1} were initialized with small values and progressively enlarged across stages. Conditioned on the resulting weights, the network parameters are updated for $L(n)$ inner iterations as

$$\begin{aligned} \theta_{C_n}^{(l+1)} = \theta_{C_n}^{(l)} - \tau \nabla_{\theta_{C_n}^{(l)}} \big(&\|s_{C_n} \odot (A f_{\theta_{C_n}}(z) - y)\|_2 / \|s_{C_n} \odot y\|_2 \\ &+ \gamma\big(\|\nabla_h f_{\theta_{C_n}}(z)\|_1 + \|\nabla_v f_{\theta_{C_n}}(z)\|_1\big)\big) \end{aligned} \tag{12}$$

where $\tau > 0$ is the step size. Algorithm 1 provides a summary.

IV. Experiments and Results

A. Experimental Setup

To assess the performance of our CogGen framework, the retrospective CS-MRI experiments were conducted on the following four publicly available in-vivo human datasets, for which ethical approval and informed-consent procedures were handled according to the original data-acquisition protocols of the corresponding data sources:

Data#1 [46]: 3D brain data from healthy subjects were acquired using a 12-channel coil and a 3D T2 CUBE sequence (readout (RO) × phase-encoding (PE) × Slice: 256×232×208). A representative slice was retrospectively undersampled with a 2D variable-density (VD) random pattern at an acceleration factor (AF) of 8.

Data#2 [47]: 2D knee data were acquired on a 3T Siemens scanner using a TSE sequence and a 15-channel coil (RO × PE × Slice: 320 × 320 × 227). A representative slice was retrospectively undersampled with a 1D VD random pattern at an AF of 6.

Data#3 [48]: Brain data were acquired on a 3T Siemens scanner using a gradient-echo sequence and a 12-channel head coil array (RO × PE × Slice: 256×256×20). A 1D VD random undersampling pattern at an AF of 5 was retrospectively applied.

Data#4 [49]: An independent multicontrast MRI dataset was acquired on a Siemens Trio Tim 3.0T system using a TSE sequence (RO × PE × Contrast: 384 × 324 × 3). A 2D VD random undersampling pattern at an AF of 10 was used during acquisition.

We provide two instantiations of the CogGen framework, namely CogGen-DIP and CogGen-INR, benchmarking them against several SOTA CS-MRI reconstruction techniques including DIP-TV [27], BM3D-FISTA [50], DISCUS [51], SSDU [52], aSeq-DIP [28], Hash-INR-ℓ_2-ℓ_1 [26], PDAC [53] and MoDL [46]. Except for SSDU, PDAC, and MoDL, all remaining competing methods operate in a training-free manner. PDAC and MoDL were trained using approximately 300∼400 external training samples, whereas SSDU adopted a scan-specific self-supervised data-splitting strategy without requiring external paired training data. Notably, PDAC is included as a mechanistically relevant white-box baseline because it also performs progressive reconstruction. The DIP-based baselines were implemented using a deep U-Net variant with a 5-level encoder-decoder, 128 feature channels, and skip connections between corresponding encoder and decoder stages. Conversely, the INR counterpart adopted a hash-encoded multilayer perceptron with sinusoidal activations, consisting of 8 hidden layers with 256 neurons each, and further incorporated Fourier feature mappings to better represent high-frequency image details, optimized with a hybrid ℓ_2-ℓ_1 data-consistency loss. All deep learning-based models were implemented in PyTorch and executed on an NVIDIA RTX A6000 GPU, and were optimized using Adam with a learning rate of 1×10^{-4}. Moreover, for all methods, we systematically tuned the model and regularization hyperparameters to ensure a fair and controlled comparison. For CogGen, we set $T = 5$ and $L = [1000, 1000, 2000, 2000, 10000]$; In addition, λ_{C_1} and r_{C_1} are set within the ranges of $[1 \times 10^{-6}, 1 \times 10^{-5}]$ and $[60, 80]$, respectively. At the final stage, λ_{C_T} is set within $[10, 15]$, and r_{C_T} is chosen as $r_{C_T} = \max_{i \in \Omega} |\mathbf{k}_i - \mathbf{k}_{\text{center}}|_2$. Quantitative evaluation was conducted using two metrics, the relative ℓ_2-norm error (RLNE) and the peak signal-to-noise ratio (PSNR), both computed within a region of interest (ROI) that encompassed the brain and knee anatomical structures while excluding the background, and whose explicit definitions are given below:

$$\text{RLNE}_{\text{ROI}} = \frac{\|x_{\text{ROI}} - \hat{x}_{\text{ROI}}\|_2}{\|x_{\text{ROI}}\|_2} \times 100\% \tag{13}$$

$$\text{PSNR}_{\text{ROI}} = 20 \log_{10}\left(\frac{\max(x_{\text{ROI}})}{\sqrt{\frac{1}{N}\|x_{\text{ROI}} - \hat{x}_{\text{ROI}}\|_2^2}}\right) \tag{14}$$

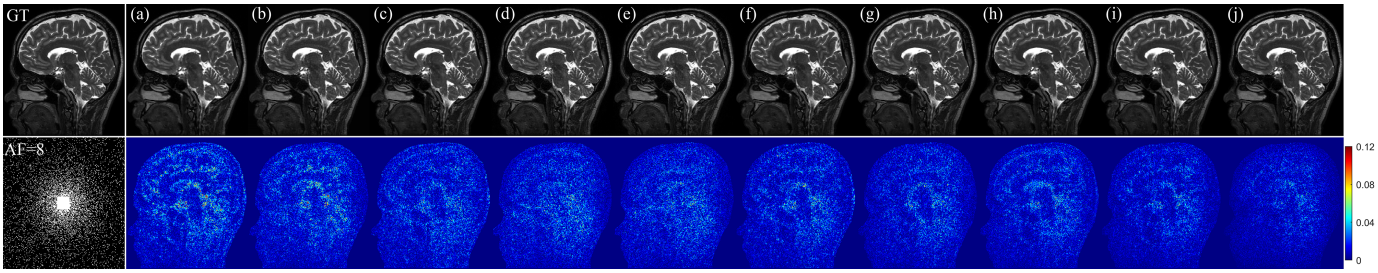

Fig. 2. Reconstruction results on Data#1 at AF = 8. (a)-(j): DIP-TV, BM3D-FISTA, DISCUS, SSDU, aSeq-DIP, Hash-INR-ℓ_2-ℓ_1, PDAC, MoDL, CogGen-DIP and CogGen-INR. The corresponding error maps are presented in the bottom row, alongside the GT and downsampling mask displayed at the far left.

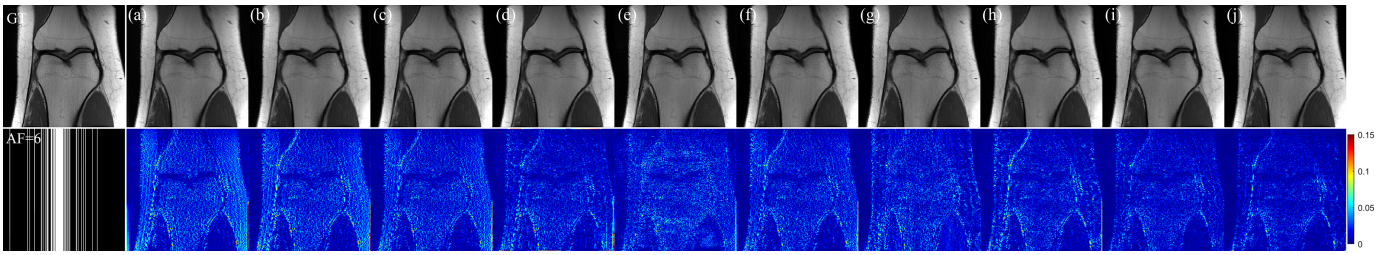

Fig. 3. Reconstruction results on Data#2 at AF = 6. (a)-(j): DIP-TV, BM3D-FISTA, DISCUS, SSDU, aSeq-DIP, Hash-INR-ℓ_2-ℓ_1, PDAC, MoDL, CogGen-DIP and CogGen-INR. The corresponding error maps are presented in the bottom row, alongside the GT and downsampling mask displayed at the far left.

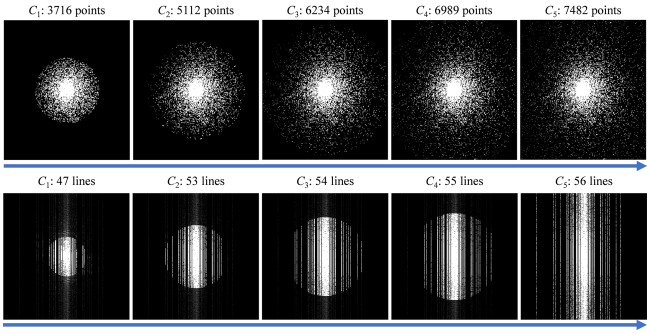

Fig. 4. Evolution of sample selection across five stages (C_1-C_5), showing the expansion from central k-space to the full acquired pattern. Top: Data#1 (2D, $w = 1$). Bottom: Data#2 (1D, $w = 0.8$). Here, C_5 denotes the final stage, which contains the original undersampled k-space measurements.

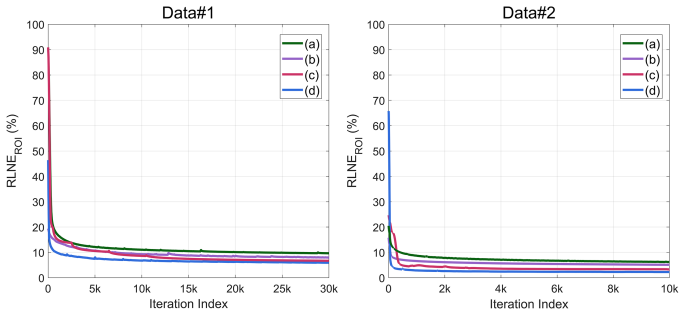

Fig. 5. Performance gains from CogGen on Data#1 (2D, AF = 8) and Data#2 (1D, AF = 6). Integrating CogGen into DIP- and INR-based baselines notably boosts accuracy and convergence under the same iteration budget. (a)-(d): DIP-TV, Hash-INR-ℓ_2-ℓ_1, CogGen-DIP, and CogGen-INR.

TABLE I
QUANTITATIVE RESULTS (RLNE$_{ROI}$ AND PSNR$_{ROI}$) ON DATA#1 (2D, AF = 8) AND DATA #2 (1D, AF = 6). THE TOP TWO PERFORMING ENTRIES ARE BOLDED.

Methods	Data#1 (2D, AF = 8)		Data#2 (1D, AF = 6)	
	RLNE$_{ROI}$(%)	PSNR$_{ROI}$(dB)	RLNE$_{ROI}$(%)	PSNR$_{ROI}$(dB)
DIP-TV	9.84	38.17	6.87	30.65
BM3D-FISTA	9.60	38.45	6.28	31.42
DISCUS	8.59	38.91	5.66	32.32
SSDU	8.52	38.91	4.83	33.79
aSeq-DIP	8.41	39.24	5.42	32.77
Hash-INR-ℓ_2-ℓ_1	8.17	39.43	5.30	32.90
PDAC	7.56	40.45	4.76	33.84
MoDL	7.89	39.97	4.65	34.03
CogGen-DIP	**6.70**	**41.50**	**3.89**	**35.58**
CogGen-INR	**6.06**	**42.42**	**3.45**	**36.63**

B. Reconstruction Performances

We present the reconstruction results in Fig. 2 (Data#1, AF = 8) and Fig. 3 (Data#2, AF = 6). It is evident that DIP-TV and BM3D-FISTA yield inferior reconstructions, with noticeably blurred structures and oversmoothed details, whereas DISCUS, SSDU, aSeq-DIP and Hash-INR-ℓ_2-ℓ_1 achieve broadly comparable visual quality by effectively suppressing noise-like artifacts while preserving delicate anatomical patterns. Notably, SSDU exhibits pronounced performance degradation under the more challenging AF = 8 setting. Meanwhile, PDAC and MoDL, functioning as the fully supervised methods, consistently outperform the remaining non-supervised competing baselines. Empowered by the CogGen scheduling procedure illustrated in Fig. 4, with $w = 1$ for Data#1 and $w = 0.8$ for Data#2, our proposed CogGen-DIP and CogGen-INR yield reconstructions that most closely match the GTs, with higher fidelity in recovering intricate textures and sharp structural boundaries. This improvement is also evident from the corresponding error maps, where CogGen produces weaker residual artifacts around anatomical edges and fine structures, indicating that progressive k-space

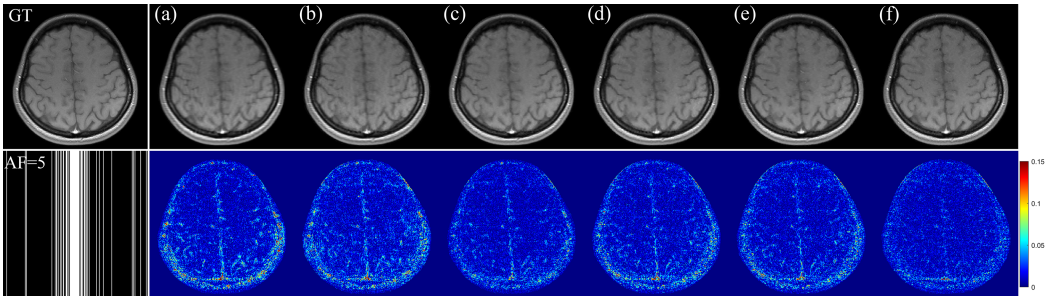

Fig. 6. Comparison between MRI-Aware and Random Scheduling on Data#3 at an AF of 5. (a)-(c): DIP-TV, SRS-DIP, CogGen-DIP; (d)-(f): Hash-INR-ℓ_2-ℓ_1, SRS-INR, CogGen-INR. The corresponding error maps are presented in the bottom row, alongside the GT and downsampling mask displayed at the far left.

TABLE II
QUANTITATIVE EVALUATION OF MRI-AWARE VS. RANDOM SCHEDULING ON DATA#3 AT AF = 5.

Methods	Data#3 (1D, AF = 5)	
	RLNE_{ROI}(%)	PSNR_{ROI}(dB)
DIP-TV	8.13	37.58
SRS-DIP	7.96	37.76
CogGen-DIP	**6.04**	**40.13**
Hash-INR-ℓ_2-ℓ_1	7.01	38.86
SRS-INR	6.95	38.95
CogGen-INR	**5.77**	**40.57**

scheduling helps suppress noise imprint while preserving diagnostically relevant details. These visual gains are reflected in Table.I, where both variants achieve the best performances across the primary evaluation metrics.

C. Discussion

1) CogGen superiority: In Fig. 5, we plot the reconstruction curves of RLNE_{ROI} and PSNR_{ROI} for DIP- and INR-based baselines (DIP-TV and Hash-INR-ℓ_2-ℓ_1) on Data#1 and Data#2, with and without the CogGen strategy. The results clearly reveal that CogGen consistently boosts both reconstruction accuracy and convergence efficiency across both backbones. Notably, to achieve comparable fidelity, CogGen requires substantially fewer iterations, highlighting a distinct low-latency advantage. Compared with DIP-TV, the Hash-INR-ℓ_2-ℓ_1-based instantiation enjoys a more pronounced gain, suggesting that the global functional parameterization of INR may benefit more directly from the easy-to-hard measurement pacing. Furthermore, during the early optimization stages, reconstructing from a dynamically constrained data subset yields noticeably higher precision than naively fitting the entire acquired measurement set from the start. This reveals that the capacity of FU-DGMs can be better exploited when measurement exposure is aligned with the current learning state. In other words, CogGen keeps the backbone generator unchanged and instead controls which measurements are emphasized during optimization. Structurally informative measurements are fitted first, while more difficult or noise-sensitive components are gradually introduced in later stages.

To further evaluate the superiority of CogGen, we conduct an additional ablation study of MRI-aware scheduling on Data#3 at AF = 5, as shown in Fig. 6. Specifically, we construct two staged random-scheduling variants based on DIP-TV and Hash-INR-ℓ_2-ℓ_1, namely SRS-DIP and SRS-INR. At each stage, the two variants activate exactly the same number of k-space measurements as their corresponding CogGen counterparts, but randomly select the participating samples instead of using the proposed dual-threshold criterion. The results indicate an interesting observation: even random progressive exposure slightly improves reconstruction compared with fitting the entire acquired measurement set from the beginning. This suggests that gradually enlarging the active measurement set may itself provide a mild regularizing effect for FU-DGM optimization. Nevertheless, CogGen consistently yields cleaner reconstructions and lower errors than its random-scheduling counterparts, demonstrating that MRI-aware scheduling plays a critical role in guiding a more accurate reconstruction trajectory. The quantitative results are reported in Table.II.

2) The influence of stage size: To isolate the effect of the stage size while avoiding excessive computational cost, we perform this controlled evaluation using CogGen-INR on Data#4 at AF = 10. As shown in Fig. 7, the quantitative results exhibit a clear dependence on the stage size, indicating the presence of an appropriate stage count at which the generative capacity of INR can be more effectively exploited. This behavior is consistent with the easy-to-hard scheduling rationale: a properly paced stage facilitates stable inversion, whereas an overly coarse schedule with too few stages, e.g., $T < 3$, may expose difficult measurements prematurely. In contrast, an overly fragmented schedule with too many stages, e.g., $T > 5$, may allocate too few iterations to each stage and slow down the incorporation of useful measurements. Both cases can hinder the staged inversion process and eventually degrade reconstruction quality.

3) Single-Thresholding vs. Dual-Thresholding: To validate the necessity of this dual-thresholding design, we conduct an component-wise experiment on Data#4 at AF = 10, using an identical INR architecture and the same total iteration budget. As illustrated in Fig. 8, removing the weighting design reduces CogGen to the standard Hash-INR-ℓ_2-ℓ_1 baseline and leads to visibly degraded reconstructions. While applying either weighting strategy alone yields clear accuracy gains, each single-thresholding variant has inherent limitations. When only the radius-based constraint is used, difficult samples within the current feasible region may still be strongly weighted before

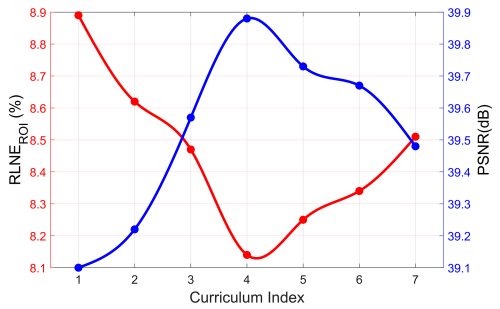

Fig. 7. Influence of stage count on Data#4 at AF = 10. Performance peaks at C_4 and decreases when additional stages are introduced, highlighting the importance of appropriate stage-count selection for reconstruction performance.

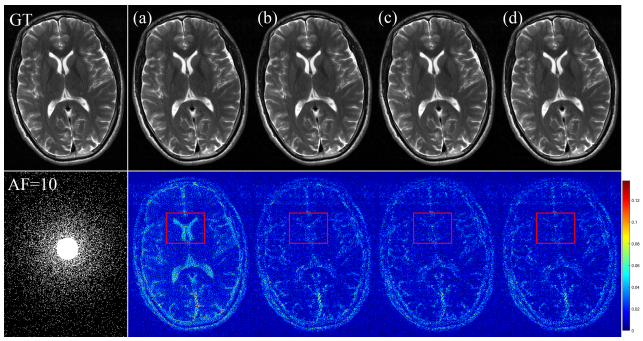

Fig. 8. Single-mode vs. dual-mode thresholding (Data#4, AF = 10). Dual-mode scheduling enhances detail recovery and reconstruction fidelity. (a-d): Hash-INR-ℓ_2-ℓ_1, CogGen$_{\mathrm{CL}}$-INR, CogGen$_{\mathrm{SPL}}$-INR and CogGen-INR.

they become reliably fitable, causing suboptimal updates before the optimization has sufficiently stabilized. In contrast, relying solely on the data-consistency residual threshold rule may penalize informative high-frequency components at later stages because it lacks frequency-structured guidance, resulting in blurred textures or localized artifacts. Combining the two modes achieves the best overall performance, improving detail recovery while preserving structural consistency, as highlighted by the red boxes. The quantitative results in Table.III further support the effectiveness of the dual-thresholding strategy across all metrics.

TABLE III
QUANTITATIVE COMPARISON FOR TWO THRESHOLDING SCHEMES ON DATA#4 (AF = 10).

Methods	Data#4 (2D, AF = 10)	
	RLNE$_{\mathrm{ROI}}$(%)	PSNR$_{\mathrm{ROI}}$(dB)
Hash-INR-ℓ_2-ℓ_1	8.89	39.10
CogGen$_{\mathrm{CL}}$-INR	8.41	39.59
CogGen$_{\mathrm{SPL}}$-INR	8.29	39.71
CogGen-INR	**8.14**	**39.88**

V. CONCLUSIONS

We introduce CogGen, a novel FU-DGM framework that employs a simple yet effective strategy to improve both reconstruction accuracy and convergence efficiency by reformulating conventional uniform k-space fitting as a staged inversion process. The framework is instantiated in two distinct forms, CogGen-DIP and CogGen-INR, whose effectiveness and generality are supported by numerical experiments and theoretical analyses. In future work, we will explore the potential submodular structure of k-space measurements with respect to image-quality assessment to enable principled, adaptive sample selection. Moreover, incorporating submodular optimization strategies [54], [55], such as vertex-cover-inspired selection [56], [57], may further tighten performance bounds while improving theoretical interpretability.

VI. APPENDIX

This appendix provides the local and conditional analyses: First, we show that, under an early-stage preference for reliably fitted measurements, the CogGen objective admits a reduced local sufficient-iteration bound compared with uniform fitting. Second, we show that early-stage weighting gives a smaller cumulative noise-amplification bound. These analyses explain the observed convergence and accuracy advantages.

A. Local Convergence Behavior of CogGen

To isolate the role of stage-wise measurement weighting, we analyze the local behavior of the non-squared weighted data-fidelity term used in CogGen. The TV regularizer and the normalization factor are omitted or absorbed into constants, since the present analysis focuses on how the stage-wise measurement operator affects the local optimization trajectory. For a fixed stage C_n, consider

$$\mathcal{L}_{C_n}(\theta) = \left\| S_{C_n}\left(Af_\theta(z) - y\right)\right\|_2, \tag{A.1}$$

where $S_{C_n} = \mathrm{diag}(s_{C_n})$ is the diagonal stage-wise weighting operator. Following a standard local linearization (NTK style) approximation [58], around a reference iterate $\theta^{(j)}$, the generator is locally linearized as

$$f_\theta(z) \approx f_{\theta^{(j)}}(z) + J^{(j)}\left(\theta - \theta^{(j)}\right), \tag{A.2}$$

where $J^{(j)} = \nabla_\theta f_{\theta^{(j)}}(z)$. Substituting (A.2) into (A.1) gives the weighted local surrogate

$$\mathcal{L}_{C_n}(\theta) \approx \left\| S_{C_n}\left(AJ^{(j)}(\theta - \theta^{(j)}) - d^{(j)}\right)\right\|_2, \tag{A.3}$$

where

$$d^{(j)} = y - Af_{\theta^{(j)}}(z). \tag{A.4}$$

For brevity, define

$$g_{C_n}(\theta) = S_{C_n}\left(AJ^{(j)}(\theta - \theta^{(j)}) - d^{(j)}\right). \tag{A.5}$$

Then $\mathcal{L}_{C_n}(\theta) \approx \|g_{C_n}(\theta)\|_2$. On a local neighborhood where $g_{C_n}(\theta) \neq 0$, the gradient is

$$\nabla\mathcal{L}_{C_n}(\theta) = \frac{(J^{(j)})^H A^H S_{C_n}^H g_{C_n}(\theta)}{\|g_{C_n}(\theta)\|_2}. \tag{A.6}$$

Assume that the local residual is bounded away from zero,

$$\|g_{C_n}(\theta)\|_2 \geq \delta_{C_n} > 0. \tag{A.7}$$

Then $\nabla\mathcal{L}_{C_n}$ is locally Lipschitz continuous. Let $\beta_{C_n} > 0$ denote a local Lipschitz constant. We further assume that $\mathcal{L}_{C_n}$ satisfies the PL inequality [59], [60]

$$\frac{1}{2}\left\|\nabla\mathcal{L}_{C_n}(\theta)\right\|_2^2 \geq \mu_{C_n}\big(\mathcal{L}_{C_n}(\theta) - \mathcal{L}_{C_n}^*\big), \tag{A.8}$$

where $\mathcal{L}_{C_n}^* = \inf_\theta \mathcal{L}_{C_n}(\theta)$, and $\mu_{C_n} > 0$ is a local effective PL curvature. The constants β_{C_n} and μ_{C_n} are local trajectory-dependent quantities, not global Hessian eigenvalues of a quadratic objective. In the early stages of CogGen, the scheduler tends to expose measurements that are structurally informative and readily fitable by the current generator, while deferring measurements that are harder to fit or more sensitive to noise. Under this early-stage fitability preference, the weighted local surrogate may admit a more favorable local geometry than uniform fitting along the identifiable trajectory.

Assumption A.1. Local early-stage fitting condition. Let n_0 be the last stage index of the early stage phase. Define

$$\mu_{\mathrm{early}} = \inf_{1\leq n\leq n_0} \mu_{C_n}. \tag{A.9}$$

Let μ_{uniform} denote the corresponding local effective PL curvature under uniform fitting, i.e., $S_{C_n} \equiv I$. We assume that, along the locally identifiable optimization trajectory,

$$\mu_{\mathrm{early}} > \mu_{\mathrm{uniform}}. \tag{A.10}$$

This assumption does not state that down-weighting measurements universally improves conditioning. It only describes the early-stage regime where selecting readily fitable measurements makes the local surrogate better aligned with the current generator representation.

Consider the stage-wise gradient descent update

$$\theta^{(\ell+1)} = \theta^{(\ell)} - \tau\nabla\mathcal{L}_{C_n}\big(\theta^{(\ell)}\big), \qquad 0 < \tau \leq \frac{1}{\beta_{C_n}}. \tag{A.11}$$

Lemma A.1. Stage-wise local linear convergence. Assume that the iterates remain in the local neighborhood where (A.7) and (A.8) hold. Under (A.8) and (A.11), with $0 < \tau\mu_{C_n} \leq 1$, the iterates satisfy

$$\begin{aligned}&\mathcal{L}_{C_n}\big(\theta^{(\ell)}\big) - \mathcal{L}_{C_n}^* \\ &\leq \big(1-\tau\mu_{C_n}\big)^{\ell}\Big[\mathcal{L}_{C_n}\big(\theta^{(0)}\big) - \mathcal{L}_{C_n}^*\Big].\end{aligned} \tag{A.12}$$

Proof. By local β_{C_n}-smoothness,

$$\begin{aligned}&\mathcal{L}_{C_n}\big(\theta^{(\ell+1)}\big) \\ &\leq \mathcal{L}_{C_n}\big(\theta^{(\ell)}\big) - \tau\left(1 - \frac{\tau\beta_{C_n}}{2}\right)\Big\|\nabla\mathcal{L}_{C_n}\big(\theta^{(\ell)}\big)\Big\|_2^2.\end{aligned} \tag{A.13}$$

Since $0 < \tau \leq 1/\beta_{C_n}$,

$$1 - \frac{\tau\beta_{C_n}}{2} \geq \frac{1}{2}. \tag{A.14}$$

Combining (A.13), (A.14), and the PL inequality (A.8) gives

$$\begin{aligned}&\mathcal{L}_{C_n}\big(\theta^{(\ell+1)}\big) - \mathcal{L}_{C_n}^* \\ &\leq \big(1-\tau\mu_{C_n}\big)\Big[\mathcal{L}_{C_n}\big(\theta^{(\ell)}\big) - \mathcal{L}_{C_n}^*\Big].\end{aligned} \tag{A.15}$$

Iterating (A.15) proves (A.12). This completes the proof. □

Theorem A.1. Reduced local sufficient-iteration bound. Fix any relative accuracy level $\xi \in (0,1)$. Suppose that the assumptions of Lemma A.1 hold for both the early stage-wise weighted surrogates and the corresponding uniform-fitting surrogate, with the same admissible step size τ. Under Assumption A.1, a sufficient number of gradient steps for CogGen at any early stage $n \leq n_0$ is

$$\ell_{\mathrm{CogGen}}^{\mathrm{suf}} = \left\lceil \frac{\log(1/\xi)}{\tau\mu_{\mathrm{early}}} \right\rceil. \tag{A.16}$$

For uniform fitting, the corresponding sufficient number of steps is

$$\ell_{\mathrm{uniform}}^{\mathrm{suf}} = \left\lceil \frac{\log(1/\xi)}{\tau\mu_{\mathrm{uniform}}} \right\rceil. \tag{A.17}$$

Before integer rounding, the CogGen sufficient-iteration bound is strictly smaller than the uniform-fitting bound. After integer rounding,

$$\ell_{\mathrm{CogGen}}^{\mathrm{suf}} \leq \ell_{\mathrm{uniform}}^{\mathrm{suf}}. \tag{A.18}$$

Proof. If

$$\mathcal{L}_{C_n}(\theta^{(0)}) - \mathcal{L}_{C_n}^* = 0, \tag{A.19}$$

then the desired relative accuracy has already been reached. Otherwise, from (A.12),

$$\frac{\mathcal{L}_{C_n}(\theta^{(\ell)}) - \mathcal{L}_{C_n}^*}{\mathcal{L}_{C_n}(\theta^{(0)}) - \mathcal{L}_{C_n}^*} \leq \big(1-\tau\mu_{C_n}\big)^{\ell}. \tag{A.20}$$

Using $1 - x \leq e^{-x}$,

$$\big(1-\tau\mu_{C_n}\big)^{\ell} \leq \exp\big(-\tau\mu_{C_n}\ell\big). \tag{A.21}$$

Thus, to make the relative objective gap no larger than ρ, it suffices to require

$$\ell \geq \frac{\log(1/\xi)}{\tau\mu_{C_n}}. \tag{A.22}$$

For all early stages $n \leq n_0$, $\mu_{C_n} \geq \mu_{\mathrm{early}}$. Hence (A.16) is sufficient for CogGen at any early stage. For uniform fitting, applying the same argument with μ_{uniform} gives (A.17). Since $\mu_{\mathrm{early}} > \mu_{\mathrm{uniform}}$,

$$\frac{\log(1/\xi)}{\tau\mu_{\mathrm{early}}} < \frac{\log(1/\xi)}{\tau\mu_{\mathrm{uniform}}}. \tag{A.23}$$

Therefore, the unrounded sufficient-iteration bound of CogGen is strictly smaller than that of uniform fitting. Since the ceiling function is monotone nondecreasing,

$$\ell_{\mathrm{CogGen}}^{\mathrm{suf}} \leq \ell_{\mathrm{uniform}}^{\mathrm{suf}}. \tag{A.24}$$

This proves the stated local sufficient-iteration comparison. □

B. Cumulative Noise-Amplification Bound of CogGen

We next analyze how stage-wise measurement weighting affects the accumulation of measurement noise during iterative reconstruction. In conventional FU-DGMs, prolonged fitting of noisy and ill-posed measurements may imprint measurement noise into the reconstructed image. The following analysis shows that CogGen can reduce the effective noise component injected during early updates, leading to a smaller upper bound on cumulative noise amplification under local stability conditions. Consider the stage-wise image-domain surrogate

$$L_{C_n}(x) = \frac{\|S_{C_n}(Ax-y)\|_2}{\|S_{C_n}y\|_2}. \tag{A.25}$$

Let

$$W_{C_n} = S_{C_n}^H S_{C_n}. \tag{A.26}$$

On a local neighborhood where $S_{C_n}(Ax-y) \neq 0$, the gradient direction is proportional to

$$A^H W_{C_n}(Ax - y). \tag{A.27}$$

To expose the noise-propagation mechanism, we consider the following rescaled Landweber-type surrogate dynamics:

$$x_{C_n}^{(\ell+1)} = x_{C_n}^{(\ell)} - \alpha_{C_n}^{(\ell)} A^H W_{C_n}\big(Ax_{C_n}^{(\ell)} - y\big). \tag{A.28}$$

The rescaling factor induced by the non-squared normalized fidelity is

$$\alpha_{C_n}^{(\ell)} = \frac{\tau}{\left\| S_{C_n}\big(Ax_{C_n}^{(\ell)} - y\big)\right\|_2 \|S_{C_n} y\|_2}. \tag{A.29}$$

This surrogate is used to describe how the stage-wise weighting operator modulates the noise injected through the data-consistency update.

Let

$$e_{C_n}^{(\ell)} = x_{C_n}^{(\ell)} - x^*. \tag{A.30}$$

Using the noisy observation model

$$y = Ax^* + \varepsilon, \tag{A.31}$$

we obtain the error recursion

$$\begin{aligned} e_{C_n}^{(\ell+1)} = &\left(I - \alpha_{C_n}^{(\ell)} A^H W_{C_n} A\right) e_{C_n}^{(\ell)} \\ &+ \alpha_{C_n}^{(\ell)} A^H W_{C_n} \varepsilon. \end{aligned} \tag{A.32}$$

Define

$$M_{C_n}^{(\ell)} = I - \alpha_{C_n}^{(\ell)} A^H W_{C_n} A. \tag{A.33}$$

Since A is undersampled, strict contraction is not expected on the entire image space. We therefore restrict the analysis to the locally identifiable subspace $\mathcal{X}_{C_n}$. Along this subspace, assume that

$$0 < \alpha_{C_n}^{(\ell)} \leq \alpha_{\max}, \tag{A.34}$$

and

$$\left\| M_{C_n}^{(\ell)}|_{\mathcal{X}_{C_n}} \right\|_2 \leq \rho_{C_n}, \qquad 0 < \rho_{C_n} < 1. \tag{A.35}$$

Let $L(n)$ denote the number of inner iterations at stage C_n.

Lemma A.2. Local cumulative noise-propagation bound. Under the error recursion (A.32) and the local stability conditions (A.34)–(A.35), the noise-induced component of the final reconstruction error satisfies

$$\left\| e_{\text{noise}}^{(T)} \right\|_2 \leq \alpha_{\max} \|A\|_2 \sum_{n=1}^{T} \Gamma_n \left\| W_{C_n} \varepsilon \right\|_2, \tag{A.36}$$

where

$$\Gamma_n = \left(\prod_{k=n+1}^{T} \rho_{C_k}^{L(k)} \right) \left(\sum_{\ell=0}^{L(n)-1} \rho_{C_n}^{\ell} \right). \tag{A.37}$$

Proof. Let e_{noise} denote the component of the error generated by the measurement noise term in (A.32). Taking norms on the locally identifiable subspace and using (A.34)–(A.35) gives, within stage C_n,

$$\left\| e_{\text{noise},C_n}^{(\ell+1)} \right\|_2 \leq \rho_{C_n} \left\| e_{\text{noise},C_n}^{(\ell)} \right\|_2 + \alpha_{\max} \|A\|_2 \left\| W_{C_n} \varepsilon \right\|_2. \tag{A.38}$$

Thus, the cumulative noise injected during the $L(n)$ inner iterations of stage C_n is bounded, at the exit of this stage, by

$$\alpha_{\max} \|A\|_2 \left(\sum_{\ell=0}^{L(n)-1} \rho_{C_n}^{\ell} \right) \left\| W_{C_n} \varepsilon \right\|_2. \tag{A.39}$$

This contribution is further propagated through later stages $C_{n+1}, \ldots, C_T$, which gives the multiplicative factor $\prod_{k=n+1}^{T} \rho_{C_k}^{L(k)}$. Summing the propagated contributions over all stages yields (A.36) with Γ_n defined in (A.37). This completes the proof. □

For a uniform-fitting FU-DGM baseline, all acquired measurements participate in the data-consistency term throughout the inference process. For notational simplicity, this representative uniform-fitting baseline is denoted as

$$W_{C_n}^{\text{uniform}} \equiv I_M, \qquad n = 1, \ldots, T. \tag{A.40}$$

Thus,

$$\left\| W_{C_n}^{\text{uniform}} \varepsilon \right\|_2 = \|\varepsilon\|_2. \tag{A.41}$$

Assumption A.2. Early-stage noise attenuation. For CogGen, the early stages defer a subset of difficult or noise-sensitive measurements. Assume that there exist $n_0 < T$ and $q \in (0, 1)$ such that

$$\left\| W_{C_n} \varepsilon \right\|_2 \leq q \|\varepsilon\|_2, \qquad n = 1, \ldots, n_0, \tag{A.42}$$

and

$$\left\| W_{C_n} \varepsilon \right\|_2 \leq \|\varepsilon\|_2, \qquad n = n_0 + 1, \ldots, T. \tag{A.43}$$

This captures the intended effect of progressive exposure: early updates avoid fitting all noisy measurements at once, while later updates gradually incorporate more complete data information.

Theorem A.2. Smaller cumulative noise-amplification bound. Assume that Lemma A.2 holds. To isolate the noise-injection effect, suppose that CogGen and the representative uniform-fitting baseline are compared under the same stage lengths and comparable local stability factors summarized by Γ_n. Under Assumption A.2, the CogGen noise-amplification bound is

$$\begin{aligned} \mathcal{B}_{\text{CogGen}}(T) = &\alpha_{\max} \|A\|_2 \|\varepsilon\|_2 \\ &\times \left(q \sum_{n=1}^{n_0} \Gamma_n + \sum_{n=n_0+1}^{T} \Gamma_n \right), \end{aligned} \tag{A.44}$$

whereas the corresponding uniform-fitting bound is

$$\begin{aligned} \mathcal{B}_{\text{uniform}}(T) = &\alpha_{\max} \|A\|_2 \|\varepsilon\|_2 \\ &\times \left(\sum_{n=1}^{n_0} \Gamma_n + \sum_{n=n_0+1}^{T} \Gamma_n \right). \end{aligned} \tag{A.45}$$

If

$$\sum_{n=1}^{n_0} \Gamma_n > 0, \tag{A.46}$$

then

$$\mathcal{B}_{\mathrm{CogGen}}(T) < \mathcal{B}_{\mathrm{uniform}}(T). \tag{A.47}$$

Consequently,

$$\left\| e^{(T)}_{\mathrm{noise,CogGen}} \right\|_2 \le \mathcal{B}_{\mathrm{CogGen}}(T) < \mathcal{B}_{\mathrm{uniform}}(T). \tag{A.48}$$

Proof. By Lemma A.2 and Assumption A.2,

$$\begin{aligned} \left\| e^{(T)}_{\mathrm{noise,CogGen}} \right\|_2 \le & \alpha_{\max} \|A\|_2 \|\varepsilon\|_2 \\ & \times \left(q \sum_{n=1}^{n_0} \Gamma_n + \sum_{n=n_0+1}^{T} \Gamma_n \right). \end{aligned} \tag{A.49}$$

This gives (A.44). For uniform fitting, using (A.41) in Lemma A.2 gives

$$\begin{aligned} \left\| e^{(T)}_{\mathrm{noise,uniform}} \right\|_2 \le & \alpha_{\max} \|A\|_2 \|\varepsilon\|_2 \\ & \times \left(\sum_{n=1}^{n_0} \Gamma_n + \sum_{n=n_0+1}^{T} \Gamma_n \right), \end{aligned} \tag{A.50}$$

which gives (A.45). Since $q < 1$ and (A.46) holds,

$$\begin{aligned} & \mathcal{B}_{\mathrm{uniform}}(T) - \mathcal{B}_{\mathrm{CogGen}}(T) \\ & = \alpha_{\max} \|A\|_2 \|\varepsilon\|_2 (1-q) \sum_{n=1}^{n_0} \Gamma_n > 0. \end{aligned} \tag{A.51}$$

Therefore, $\mathcal{B}_{\mathrm{CogGen}}(T) < \mathcal{B}_{\mathrm{uniform}}(T)$, and (A.48) follows from (A.49). □

Therefore, when early-stage weighting reduces the effective noise component and the local stability factors are comparable, CogGen admits a smaller upper bound on the cumulative noise-amplification term than the corresponding uniform fitting. This explains why CogGen can suppress noise imprinting and preserve image details more effectively during finite-budget reconstruction.

REFERENCES

[1] S. Hussain, I. Mubeen, N. Ullah, S. S. U. D. Shah, B. A. Khan, M. Zahoor, R. Ullah, F. A. Khan, and M. A. Sultan, "Modern diagnostic imaging technique applications and risk factors in the medical field: a review," *BioMed research international*, vol. 2022, no. 1, p. 5164970, 2022.

[2] S. De Pietro, G. Di Martino, M. Caroprese, A. Barillaro, S. Cocozza, R. Pacelli, R. Cuocolo, L. Ugga, F. Briganti, A. Brunetti *et al.*, "The role of mri in radiotherapy planning: a narrative review "from head to toe"," *Insights into Imaging*, vol. 15, no. 1, p. 255, 2024.

[3] A. Kits, J. Al-Saadi, F. De Luca, F. Janzon, M. V. Mazya, J. Lundberg, T. Sprenger, S. Skare, and A. F. Delgado, "2.5-minute fast brain mri with multiple contrasts in acute ischemic stroke," *Neuroradiology*, vol. 66, no. 5, pp. 737–747, 2024.

[4] R. Guo, S. Weingärtner, P. Šiurytė, C. T Stoeck, M. Füetterer, A. E Campbell-Washburn, A. Suinesiaputra, M. Jerosch-Herold, and R. Nezafat, "Emerging techniques in cardiac magnetic resonance imaging," *Journal of Magnetic Resonance Imaging*, vol. 55, no. 4, pp. 1043–1059, 2022.

[5] C. M. Sandino, J. Y. Cheng, F. Chen, M. Mardani, J. M. Pauly, and S. S. Vasanawala, "Compressed sensing: From research to clinical practice with deep neural networks: Shortening scan times for magnetic resonance imaging," *IEEE signal processing magazine*, vol. 37, no. 1, pp. 117–127, 2020.

[6] C. Curatolo, "Recent advances in parallel imaging for mri: Wave-caipi technique," *Journal of Advanced Health Care*, vol. 4, no. 1, 2022.

[7] M. Tavakkoli and M. Noseworthy, "A review on accelerated magnetic resonance imaging techniques: Parallel imaging, compressed sensing, and machine learning," *Critical Reviews™ in Biomedical Engineering*, 2025.

[8] L. Ruthotto and E. Haber, "An introduction to deep generative modeling," *GAMM-Mitteilungen*, vol. 44, no. 2, p. e202100008, 2021.

[9] M. Suzuki and Y. Matsuo, "A survey of multimodal deep generative models," *Advanced Robotics*, vol. 36, no. 5-6, pp. 261–278, 2022.

[10] E. M. Eksioglu, "Decoupled algorithm for mri reconstruction using nonlocal block matching model: Bm3d-mri," *Journal of Mathematical Imaging and Vision*, vol. 56, pp. 430–440, 2016.

[11] M. Schloegl, M. Holler, A. Schwarzl, K. Bredies, and R. Stollberger, "Infimal convolution of total generalized variation functionals for dynamic mri," *Magnetic Resonance in Medicine*, vol. 78, no. 1, pp. 142–155, 2016.

[12] R. Cohen, M. Elad, and P. Milanfar, "Regularization by denoising via fixed-point projection (red-pro)," *SIAM Journal on Imaging Sciences*, vol. 14, no. 3, pp. 1374–1406, 2021.

[13] Q. Zhu, Z.-X. Cui, Y. Liu, J. Cheng, K. Zhao, H. Wang, Y. Zhu, and D. Liang, "Characteristic-constrained accelerating mr t1rho mapping with blockwise infimal convolution of matrix elastic-net regularization," *Medical Physics*, vol. 50, no. 4, pp. 2224–2238, 2023.

[14] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1310–1321, 2018.

[15] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing mri reconstruction using a generative adversarial network with a cyclic loss," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1488–1497, 2018.

[16] Y. Li, J. Li, F. Ma, S. Du, and Y. Liu, "High quality and fast compressed sensing mri reconstruction via edge-enhanced dual discriminator generative adversarial network," *Magnetic Resonance Imaging*, vol. 77, pp. 124–136, 2021.

[17] G. Oh, B. Sim, H. Chung, L. Sunwoo, and J. C. Ye, "Unpaired deep learning for accelerated mri using optimal transport driven cyclegan," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1285–1296, 2020.

[18] J. Yoo, K. H. Jin, H. Gupta, J. Yerly, M. Stuber, and M. Unser, "Time-dependent deep image prior for dynamic mri," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3337–3348, 2021.

[19] M. Z. Darestani and R. Heckel, "Accelerated mri with un-trained neural networks," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 724–733, 2021.

[20] L. Shen, J. Pauly, and L. Xing, "Nerp: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 1, pp. 770–782, 2022.

[21] Q. Zhu, B. Liu, Z.-X. Cui, C. Cao, X. Yan, Y. Liu, J. Cheng, Y. Zhou, Y. Zhu, H. Wang, H. Zeng, and D. Liang, "Pearl: Cascaded self-supervised cross-fusion learning for parallel mri acceleration," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 5, pp. 3086–3097, 2025.

[22] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9446–9454.

[23] A. Molaei, A. Aminimehr, A. Tavakoli, A. Kazerouni, B. Azad, R. Azad, and D. Merhof, "Implicit neural representation in medical imaging: A comparative survey," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 2381–2391.

[24] A. Benfenati, A. Catozzi, G. Franchini, and F. Porta, "Early stopping strategies in deep image prior," *Soft Computing*, vol. 29, no. 8, pp. 4153–4174, 2025.

[25] R. Heckel and P. Hand, "Deep decoder: Concise image representations from untrained non-convolutional networks," *arXiv preprint arXiv:1810.03982*, 2018.

[26] J. Xu, Y. Liu, Y. Yang, Z.-X. Cui, J. Cheng, Q. Zhu, N. Zhang, Y. Zhou, D. Liang, and Y. Zhu, "Self-supervised deep unrolled model with implicit neural representation regularization for accelerating mri reconstruction," *arXiv preprint arXiv:2510.06611*, 2025.

[27] J. Liu, Y. Sun, X. Xu, and U. S. Kamilov, "Image restoration using total variation regularized deep image prior," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Ieee, 2019, pp. 7715–7719.

[28] I. Alkhouri, S. Liang, E. Bell, Q. Qu, R. Wang, and S. Ravishankar, "Image reconstruction via autoencoding sequential deep image prior," *Advances in Neural Information Processing Systems*, vol. 37, pp. 18 988–19 012, 2024.

[29] H. Wang, T. Li, Z. Zhuang, T. Chen, H. Liang, and J. Sun, "Early stopping for deep image prior," *arXiv preprint arXiv:2112.06074*, 2021.

[30] J. Sweller, "The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion," *Educational Psychology Review*, vol. 35, no. 4, p. 95, 2023.
[31] O. Chen, F. Paas, and J. Sweller, "A cognitive load theory approach to defining and measuring task complexity through element interactivity," *Educational Psychology Review*, vol. 35, no. 2, p. 63, 2023.
[32] T.-H. Kim and J. Choi, "Screenernet: Learning self-paced curriculum for deep neural networks," *arXiv preprint arXiv:1801.00904*, 2018.
[33] D. Zhang, J. Han, L. Zhao, and D. Meng, "Leveraging prior-knowledge for weakly supervised object detection under a collaborative self-paced curriculum learning framework," *International Journal of Computer Vision*, vol. 127, no. 4, pp. 363–380, 2019.
[34] Y.-W. Choi, H.-B. Shin, and S.-W. Lee, "Brain-guided self-paced curriculum learning for adaptive human–machine interfaces," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2025.
[35] G. Jagatap and C. Hegde, "Phase retrieval using untrained neural network priors," in *NeurIPS 2019 Workshop on Solving Inverse Problems with Deep Networks*, 2019.
[36] V. Saragadam, R. Balestriero, A. Veeraraghavan, and R. G. Baraniuk, "Deeptensor: Low-rank tensor decomposition with deep network priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
[37] R. Heckel and M. Soltanolkotabi, "Denoising and regularization via exploiting the structural bias of convolutional generators," 2020.
[38] N. Buskulic, Y. Quéau, and J. Fadili, "Convergence guarantees of overparametrized wide deep inverse prior," in *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer, 2023, pp. 406–417.
[39] N. Buskulic, J. Fadili, and Y. Quéau, "Convergence and recovery guarantees of unsupervised neural networks for inverse problems," *Journal of Mathematical Imaging and Vision*, vol. 66, no. 4, pp. 584–605, 2024.
[40] Y. Liu, D. Li, X. Fu, X. Lu, J. Huang, and Z.-J. Zha, "Uhd-processer: Unified uhd image restoration with progressive frequency learning and degradation-aware prompts," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 23 121–23 130.
[41] X. Zhou, O. Wu, and M. Li, "Investigating the sample weighting mechanism using an interpretable weighting framework," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 5, pp. 2041–2055, 2023.
[42] J. Shu, X. Yuan, D. Meng, and Z. Xu, "Cmw-net: Learning a class-aware sample weighting mapping for robust deep learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 10, pp. 11 521–11 539, 2023.
[43] O. Wu and R. Yao, "Data optimization in deep learning: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 37, no. 5, pp. 2356–2375, 2025.
[44] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 41–48.
[45] M. Kumar, B. Packer, and D. Koller, "Self-paced learning for latent variable models," vol. 23, 2010.
[46] H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 394–405, 2018.
[47] F. Knoll, J. Zbontar, A. Sriram, M. J. Muckley, M. Bruno, A. Defazio, M. Parente, K. J. Geras, J. Katsnelson, H. Chandarana *et al.*, "fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning," *Radiology: Artificial Intelligence*, vol. 2, no. 1, p. e190007, 2020.
[48] Q. Y. Zhu, W. Wang, J. Cheng, and X. Peng, "Incorporating reference guided priors into calibrationless parallel imaging reconstruction," *Magnetic Resonance Imaging*, vol. 57, pp. 347–358, 2019.
[49] X. Peng, L. Ying, Q. G. Liu, Y. J. Zhu, Y. Y. Liu, X. B. Qu, X. Liu, H. R. Zheng, and D. Liang, "Incorporating reference in parallel imaging and compressed sensing," *Magnetic Resonance in Medicine*, vol. 73, no. 4, pp. 1490–1504, 2015.
[50] E. M. Eksioglu, "Decoupled algorithm for mri reconstruction using nonlocal block matching model: Bm3d-mri," *Journal of Mathematical Imaging and Vision*, vol. 56, no. 3, pp. 430–440, 2016.
[51] M. A. Sultan, C. Chen, Y. Liu, K. Gil, K. Zareba, and R. Ahmad, "An unsupervised method for mri recovery: deep image prior with structured sparsity," *Magnetic Resonance Materials in Physics, Biology and Medicine*, pp. 1–13, 2025.
[52] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magnetic resonance in medicine*, vol. 84, no. 6, pp. 3172–3191, 2020.
[53] C. Wang, L. Guo, Y. Wang, H. Cheng, Y. Yu, and B. Wen, "Progressive divide-and-conquer via subsampling decomposition for accelerated mri," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 25 128–25 137.
[54] F. Yang, K. He, L. Yang, H. Du, J. Yang, B. Yang, and L. Sun, "Learning interpretable decision rule sets: A submodular optimization approach," *Advances in Neural Information Processing Systems*, vol. 34, pp. 27 890–27 902, 2021.
[55] M. El Halabi, S. Srinivas, and S. Lacoste-Julien, "Data-efficient structured pruning via submodular optimization," *Advances in Neural Information Processing Systems*, vol. 35, pp. 36 613–36 626, 2022.
[56] M. Shi, Z. Yang, and W. Wang, "Greedy guarantees for minimum submodular cost submodular/non-submodular cover problem," *Journal of Combinatorial Optimization*, vol. 45, no. 1, p. 8, 2023.
[57] M. Shi, Q. Zhu, B. Liu, and Y. Li, "Weak submodularity implies localizability: Local search for constrained non-submodular function maximization," *Discrete Mathematics*, vol. 348, no. 2, p. 114287, 2025.
[58] E. Golikov, E. Pokonechnyy, and V. Korviakov, "Neural tangent kernel: A survey," *arXiv preprint arXiv:2208.13614*, 2022.
[59] Q. Xiao, S. Lu, and T. Chen, "A generalized alternating method for bilevel learning under the Polyak–Łojasiewicz condition," *arXiv preprint arXiv:2306.02422*, 2023.
[60] A. B. Nejma, "Polyak–Łojasiewicz inequality is essentially no more general than strong convexity for c^2 functions," *arXiv preprint arXiv:2512.05285*, 2025.