



Spectral-Domain Anchor Learning for Scale-Robust Continuous Visual Fields

Xingran Guo^{1,2,3} · Tiaojie Xiao^{1,2,3} · Jie Liu^{1,2,3} · Keqin Li^{4,5}

Received: 31 March 2026 / Accepted: 30 August 2026

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

Abstract

Continuous neural fields have advanced spatial-temporal signal parameterization, yet their efficacy is fundamentally limited by spectral bias and slow-convergence dynamics. These limitations severely restrict their capacity to resolve high-frequency stochastic textures and dense, highly oscillatory spatial fields. To address these limitations, we propose the Spectral-Domain Anchor Learning (SDAL) framework, which transitions from localized spatial gradient descent to explicit spectral-domain parameterization. SDAL constructs a frequency-domain “Spider-Web Topology” that functions as a geometrically-aware, multi-ring adaptive filter bank, enabling improved scale robustness and significantly reduced aliasing artifacts. Crucially, we introduce an Alternating Optimization Strategy equipped with a closed-form analytical solver to decouple the highly non-convex spatial learning problem. This architectural innovation yields a zero-iteration cold-start capability: by resolving the projection weights W via a closed-form analytical solver at network initialization, SDAL achieves high reconstruction fidelity prior to any gradient update. Comprehensive evaluations across highly oscillatory continuous spatial fields, extreme cross-domain sensing modalities (e.g., subsurface radar topologies), and severely degraded visual signals support our theoretical assertions. SDAL consistently achieves effective spectral disentanglement, enabling the sharp recovery of faint structural signatures in noisy empirical environments. Furthermore, SDAL achieves equivalent reconstruction fidelity with zero gradient iterations, at the cost of a one-time $\mathcal{O}(K^2)$ Cholesky factorization. This approach trades iterative gradient search for a single analytical matrix inversion, which scales favorably for small-to-moderate anchor counts compared to conventional coordinate networks. Ultimately, this work advances implicit representations into robust, geometrically regularized solvers, providing a scalable foundation for large-scale computer vision and continuous spatial learning.

Keywords Implicit neural representations · Spectral bias · Scale-robust representation · Alternating optimization · Continuous visual representation · Feature disentanglement

Communicated by Gunhee Kim.

✉ Xingran Guo
gxrnudt@nudt.edu.cn

✉ Tiaojie Xiao
xiaotiaojie@nudt.edu.cn

Jie Liu
liujie@nudt.edu.cn

Keqin Li
lik@newpaltz.edu

¹ College of Computer, National University of Defense Technology, Changsha 410073, China

² Science and Technology on Parallel and Distributed Processing Laboratory, National University of Defense Technology, Changsha 410073, China

1 Introduction

Continuous neural fields, encompassing Implicit Neural Representations (INRs) and Physics-Informed Neural Networks (PINNs), have advanced the paradigm of parameterizing continuous spatial-temporal signals. By mapping coordinates to signal values, these architectures have demonstrated significant potential in diverse applications, ranging from

³ Laboratory of Software Engineering for Complex Systems, National University of Defense Technology, Changsha 410073, China

⁴ Department of Computer Science, State University of New York, New Paltz 12561, USA

⁵ College of Information Science and Engineering, Hunan University, Changsha 410000, China

large-scale satellite altimetry reconstruction (Li et al., 2026), open-vocabulary 3D mapping (Tie et al., 2024), and continuous image super-resolution (Liu et al., 2025), to the formulation of spatial-temporal overlapping domain decompositions (Ye et al., 2025) and the reconstruction of severely degraded visual signals (Zhang et al., 2025). However, the optimization of these coordinate-based networks is fundamentally bottlenecked by the well-documented *Spectral Bias* (Cai et al., 2024). Standard Multilayer Perceptrons (MLPs) inherently prioritize learning low-frequency geometric structures while suffering significant failures when attempting to resolve high-frequency stochastic textures or dense physical oscillatory fields.

To circumvent this limitation, recent literature has witnessed a proliferation of feature augmentation strategies. While hybrid architectures utilizing discrete spatial grids (Kim & Fridovich-Keil, 2026) or tensorial feature integrations (Kim et al., 2024) offer accelerated fitting for dense signals, they inherently sacrifice the continuous, resolution-agnostic properties of neural fields. Alternatively, approaches leveraging periodic activation functions (e.g., SIREN), Fourier reparameterized training (Shi et al., 2024), and Fourier-based activations (Aghaee & Khan, 2024) have been proposed to inject high-frequency priors into the latent space. Yet, these gradient-driven methods introduce a secondary bottleneck: they exhibit limiting slow-convergence trajectories and are highly susceptible to intractable high-frequency speckle oscillations, heavily restricting their utility in time-critical or out-of-distribution (OOD) scenarios.

This dual challenge of spectral bias and optimization inefficiency becomes particularly critical when parameterizing complex visual topologies and high-frequency wave equations (Chai et al., 2024). Indeed, recent advances advocate for embedding explicit geometric or physically-aware constraints directly into implicit representations to accurately infer complex continuous fields and surfaces (Mello Rella et al., 2025). Physical phenomena, such as complex multi-source interference or modal wave fields in 3D acoustics (Schoder, 2025), demand rigorous spectral fidelity. Although advanced PINNs have been stabilized for elastic (Staub, 2024) and unconditional time-stepping wave equations (Jiao et al., 2024; Qi & Sarris, 2024), they still fundamentally rely on localized stochastic gradient descent. Consequently, the network must blindly search for the optimal spectral manifold, resulting in massive computational overhead and vulnerability to real-world signal degradations (e.g., extreme sparsity and high-intensity noise).

To address these prevailing bottlenecks, we propose the SDAL framework. Departing from conventional spatial-domain gradient searching, SDAL explicitly constructs a *Spider-Web Topology* in the frequency domain, acting as a physically-aware, multi-ring adaptive filter bank. Furthermore, we introduce an *Alternating Optimization Strategy*

that decouples the highly non-convex spatial learning problem. By resolving the continuous projection weights via a closed-form analytical solver, our method exhibits a significant zero-iteration cold-start capability. It bypasses the lengthy iterative amplitude-fitting process, instantaneously mapping the physical space to an optimal spectral manifold at network initialization.

SDAL is designed as a per-signal implicit representation framework—it fits a single continuous field from its observed samples without requiring a pre-trained model or scene-specific neural architecture. This positions SDAL in the same problem space as SIREN, FFN, and FINER, rather than scene-level representations (NeRF, Instant-NGP) that require multi-view inputs and volumetric rendering. Within this per-signal setting, SDAL's primary advantages are: (1) zero-iteration initialization enabling real-time continuous field fitting for time-critical applications such as medical imaging, physical simulation monitoring, and streaming sensor data compression; (2) memory-efficient $\mathcal{O}(K^2)$ scalability enabling deployment at arbitrary spatial resolution; and (3) interpretable spectral parameterization where each anchor corresponds to a learned spatial frequency, enabling explicit spectral analysis of the reconstructed field. While downstream applications such as super-resolution and novel-view synthesis remain directions for future work, the per-signal fitting paradigm established by SDAL provides the foundational continuous representation layer upon which such applications can be built.

In summary, the main contributions of this work are three-fold:

- We propose SDAL, a novel spectral parameterization framework that categorically overcomes the spectral bias and enables improved scale robustness across diverse topological domains.
- We introduce a closed-form Alternating Optimization Strategy that fundamentally redefines the convergence dynamics of neural fields. It achieves equivalent reconstruction fidelity with zero gradient iterations, at the cost of a one-time $\mathcal{O}(K^2)$ Cholesky factorization. This represents a fundamental transition: trading iterative gradient search for a single analytical matrix inversion, which scales favorably for small-to-moderate anchor counts compared to baseline coordinate networks.
- We demonstrate the unparalleled robustness of our methodology in resolving complex continuous physics wave fields, showcasing structural immunity to extreme sparse sampling, significant occlusion, and intense noise interference.
- We demonstrate the extreme cross-domain generalization of SDAL on real-world severely degraded topologies, specifically for ground penetrating radar (GPR) subsurface perception. In this context, our framework effec-

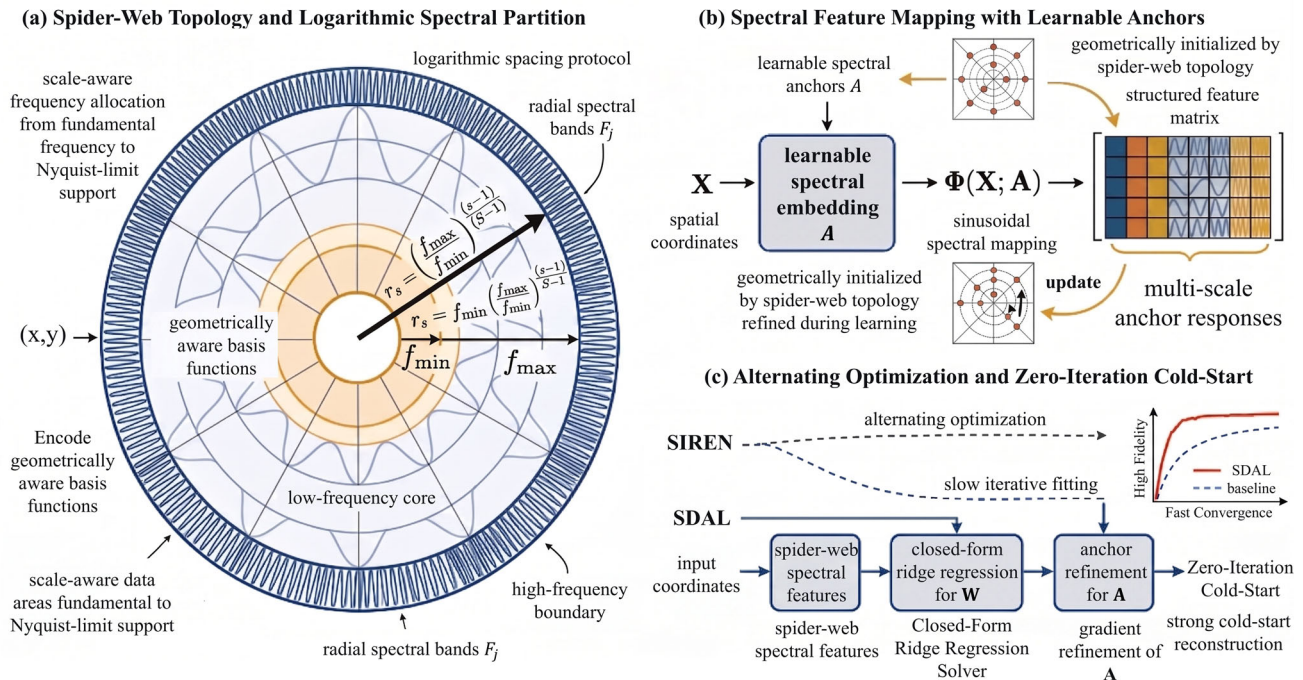


Fig. 1 Comprehensive overview of the SDAL framework. The architecture is decomposed into three synergic modules: **a** Spider-web topology for implicit encoding: Input spatial coordinates are projected onto a geometrically regularized spectral manifold. Governed by a logarithmic spacing protocol $r_s = f_{\min} (f_{\max}/f_{\min})^{\frac{s-1}{S-1}}$ from the fundamental frequency $f_{\min}$ (inner core) to the Nyquist limit $f_{\max}$ (outer boundary), the multi-ring constraint strictly bounds the high-frequency support across defined radial bands. This serves as a physically-aware filter bank, capturing highly oscillatory structures without spectral leakage. **b** Spectral Feature mapping with learnable anchors: Spatial coordinates X are projected onto a set of dynamically optimized spec-

tral anchors A via a continuous sinusoidal feature mapping $\Phi(X; A)$. This generates a structured feature matrix composed of multi-scale anchor responses, effectively translating the physical space into an optimal spectral manifold. **c** Alternating Optimization Flow and Analytical Cold-Start: Contrast of optimization dynamics. While standard SIREN (dashed line) relies on blind localized gradients (Adam) over thousands of epochs, SDAL (solid line) decouples the optimization via our alternating strategy. By resolving the continuous projection weights W via a regularized closed-form solver, the network instantaneously brings fidelity above 30 dB at the initialized step (zero-iteration cold-start), addressing the slow-convergence bottleneck of coordinate networks

tively disentangles high-frequency hyperbolic signatures from intense geological clutter—a task previously considered intractable for standard implicit representations without extensive fine-tuning.

2 Related Work

2.1 Implicit Neural Representations and Spectral Tuning

Implicit Neural Representations (INRs) have emerged as the standard for continuous visual and spatial signal parameterization. However, mitigating their inherent low-frequency bias has remained an active area of research, particularly for reconstructing complex visual topologies. Recent advances have sought to tune the spectral bandwidth through adaptive wavelet-positional encoding (Zhao et al., 2025), flexible variable-periodic activation functions (FINER) (Liu et al., 2024), wavelet-based representations (WIRE) (Saragadam

et al., 2023), and batch normalization schemes (Cai et al., 2024). Furthermore, achieving invariance across varying scales and signal disorders has motivated the development of disorder-invariant (Zhu et al., 2024) and scale-invariant implicit formulations for downstream tasks like object counting (Xu et al., 2025). Techniques such as frequency shifting (Kania et al., 2025) have also been explored to accelerate representation learning. While these methods incrementally improve frequency coverage, they still remain bound to the limitations of gradient descent optimization. In contrast, our SDAL framework explicitly disentangles the spectral mapping from the weight optimization via a closed-form solver, achieving instantaneous high-frequency alignment—a critical requirement for reconstructing severely degraded visual signals and continuous fields without the risk of unstable optimization oscillations.

While SDAL shares the sinusoidal feature mapping structure of Random Fourier Features (RFF) (Rahimi & Recht, 2007) and the closed-form ridge regression weight update of Kernel Ridge Regression (KRR), it differs in three fundamen-

tal respects. First, RFF employs a static, randomly-sampled frequency matrix that cannot adapt to the target signal's spectral content; SDAL's spectral anchors A are dynamically optimized via gradient descent to align with the signal's dominant eigen-frequencies, enabling adaptive spectral learning. Second, KRR treats the kernel as fixed and optimizes only the output weights; SDAL alternates between optimizing the non-linear spectral basis A and the linear projection W , enabling joint spectral adaptation through a rigorous alternating minimization framework. Third, the Spider-Web Topology imposes a geometrically-aware, logarithmically-spaced multi-ring constraint on A that guarantees uniform coverage of the effective bandwidth $\mathcal{B}_{eff}$ —a structural property absent in both RFF and KRR that is essential for the scale-invariance guarantees proven in Sect. 3.4. These distinctions position SDAL as a continuously optimized spectral dictionary with provable frequency-domain coverage properties, rather than a fixed-kernel regression method.

2.2 Distinction from Classical Spectral Methods

While SDAL shares the sinusoidal feature mapping structure of Random Fourier Features (RFF) and employs closed-form weight updates reminiscent of Kernel Ridge Regression (KRR), it introduces three fundamental innovations that distinguish it from these classical approaches.

First, RFF employs a static, randomly-sampled frequency matrix that remains frozen throughout training and cannot adapt to the target signal's spectral content. In contrast, SDAL's spectral anchors are dynamically optimized via gradient descent to align with the signal's dominant eigen-frequencies, enabling adaptive spectral learning.

Second, KRR treats the kernel as fixed and optimizes only the output weights through a single closed-form solve. SDAL introduces a rigorous alternating minimization framework that iteratively optimizes both the non-linear spectral basis and the linear projection, allowing the spectral dictionary to continuously adapt its coverage.

Third, the Spider-Web Topology imposes a geometrically-aware, logarithmically-spaced multi-ring constraint that guarantees uniform coverage of the effective bandwidth. This structural property is essential for the scale-robustness guarantees derived in Theorem 1, as it ensures the relative spectral gap remains constant across all frequency scales.

These distinctions position SDAL as a continuously optimized spectral dictionary with provable frequency-domain coverage properties, rather than a fixed-kernel regression method.

2.3 Physics-Informed Neural Networks for Wave Phenomena

While INRs dominate standard vision tasks, parameterizing highly oscillatory physical fields—such as those governed by wave equations—remains notoriously challenging, yet these fields serve as extreme stress tests for high-frequency spatial pattern extraction (Chai et al., 2024). Recent efforts have successfully adapted PINNs to tackle forward and inverse problems in shallow water equations using non-dimensionalization and multi-task learning (NDAWL-PINN) (Qi et al., 2025), stabilized unconditionally stable time-stepping for wave equations (Jiao et al., 2024; Qi & Sarris, 2024), and robust elastic wave assessments (Staub, 2024). Moreover, causal neuro-spectral architectures have been introduced to ensure physical temporal dependencies (Bizzi et al., 2026), while specialized PINNs have been deployed for 3D room acoustic modal predictions (Schoder, 2025). Despite these advancements, standard PINNs typically require extensive hyperparameter tuning to balance PDE residuals against boundary conditions. Our approach circumvents this by projecting the spatial domain directly into a globally aware spectral anchor dictionary, inherently regularizing the physical solution space without the need for localized gradient-based boundary enforcement, thereby offering a robust paradigm for both physics-informed modeling and complex visual topology reconstruction.

2.4 Neural Operators and Spectral Architectures

Neural operators have redefined both scientific computing and continuous spatial modeling by learning resolution-invariant mappings between infinite-dimensional function spaces. The Fourier Neural Operator (FNO), in particular, has seen massive extensions, including lightweight variants for time-dependent PDEs (Ahn et al., 2025), invertible architectures for tackling both forward and inverse problems (Long et al., 2024), and enhancements tailored for complex underwater parabolic equations (Zheng et al., 2025). Concurrently, spectral theories have been seamlessly integrated into conditional neural processes via spectral convolutions (Mohseni & Duffield, 2026) and applied to the observable dynamics of time-periodic quantum Floquet systems (Qi et al., 2026). While neural operators excel at learning global functional mappings across massive datasets, they often require substantial paired training data. Our SDAL draws inspiration from these spectral mechanisms but applies them directly to the implicit representation of a single physical or visual field. By constraining the spectral anchors within a multi-ring radial topology, we emulate the global receptive field of spectral convolutions while retaining the extreme parameter efficiency and zero-iteration cold-start benefits of closed-form implicit learning, successfully bridging the gap between

scientific continuous operators and high-fidelity representation learning.

3 Methodology

Before delving into the rigorous mathematical formulations, we provide a high-level conceptual overview of the proposed SDAL framework, as illustrated in Fig. 1. Unlike conventional implicit neural representations that blindly aggregate spatial features via localized gradient descent, SDAL fundamentally redefines the representation space. By explicitly parameterizing continuous coordinates through a structurally regularized frequency-domain topology and decoupling the highly non-convex objective via a closed-form solver, our framework strictly guarantees scale-invariance while achieving instantaneous optimization convergence for complex visual and spatial fields.

3.1 Preliminaries and Problem Formulation

The fundamental objective of continuous neural representation learning is to model large-scale spatial data implicitly within the weights of a neural network. Let $\Omega \subset \mathbb{R}^d$ denote a continuous spatial domain, and $\mathbf{u} : \Omega \rightarrow \mathbb{R}^c$ represent a target continuous field (e.g., high-resolution visual images or complex electromagnetic sensing topologies). A coordinate-based neural field seeks to approximate this underlying continuous signal via a parameterized function $f_\theta : \mathbb{R}^d \rightarrow \mathbb{R}^c$, such that $f_\theta(\mathbf{x}) \approx \mathbf{u}(\mathbf{x})$ for any coordinate $\mathbf{x} \in \Omega$, where θ encompasses the learnable weights and biases of a Multilayer Perceptron (MLP).

Formally, the optimization of such a neural field is posed as an empirical risk minimization problem over a discrete set of sampled coordinates $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^N$:

$$\min_{\theta} \mathcal{J}(\theta) = \frac{1}{N} \sum_{i=1}^N \|f_\theta(\mathbf{x}_i) - \mathbf{u}(\mathbf{x}_i)\|_2^2 + \lambda \mathcal{R}(\theta), \quad (1)$$

where $\mathcal{R}(\theta)$ denotes an optional regularization term. While this paradigm has achieved novel success in synthesizing low-frequency visual signals, its direct application to severely degraded or highly oscillatory spatial fields is fundamentally bottlenecked by the intrinsic optimization dynamics of MLPs.

To elucidate this bottleneck mathematically, we resort to the Neural Tangent Kernel (NTK) theory. In the infinite-width limit, the training dynamics of a neural network trained via gradient descent can be characterized by the continuous-time differential equation:

$$\frac{d\mathbf{f}(t)}{dt} = -\mathbf{H} \cdot (\mathbf{f}(t) - \mathbf{u}), \quad (2)$$

where $\mathbf{H} \in \mathbb{R}^{N \times N}$ is the NTK matrix with entries $\mathbf{H}_{ij} = \langle \nabla_{\theta} f_{\theta}(\mathbf{x}_i), \nabla_{\theta} f_{\theta}(\mathbf{x}_j) \rangle$. The convergence rate of the network along the eigenvectors of $\mathbf{H}$ is strictly governed by its corresponding eigenvalues λ_k .

Crucially, standard MLPs equipped with conventional activation functions exhibit a pronounced *spectral bias*: the eigenvalues λ_k decay exponentially with respect to the frequency k of the target signal components. Consequently, the network prioritizes learning smooth, low-frequency macroscopic topologies while stagnating on high-frequency, small-scale textural details. Although recent advancements, such as Sinusoidal Representation Networks (SIREN) and Fourier Feature Networks (FFN), alleviate this bias by mapping input coordinates into higher-dimensional periodic spaces, they remain heavily reliant on localized gradient descent.

When confronted with high-frequency spatial topologies characterized by orthogonal visual lattices or complex Moiré interference patterns, traversing the rugged loss landscape via first-order gradient methods becomes an ill-posed optimization problem. The network either fails to converge within a practical temporal budget or becomes entrapped in suboptimal local minima, resulting in severe structural residuals.

This significant misalignment between the exponentially decaying eigenspectrum of gradient-based MLPs and the broad-band high-frequency nature of complex visual and physical fields formulates the central challenge addressed in this work. To break this spectral bias and achieve instantaneous, high-fidelity convergence, a transition from spatial-domain gradient descent to closed-form spectral-domain anchoring is imperatively demanded.

3.2 Spectral-Domain Anchor Representation

To circumvent the spectral bias inherent in deeply cascaded, gradient-driven MLPs, we propose a transition from spatial-domain feature aggregation to explicit spectral-domain parameterization. We introduce the SDAL framework, which models the continuous spatial field f_θ by projecting coordinates onto a set of learnable, non-local frequency bases, termed *Spectral Anchors*.

3.2.1 Continuous Feature Mapping via Spectral Anchors

Let $\mathcal{A} = \{\mathbf{a}_k\}_{k=1}^K$ denote a set of K trainable spectral anchors, where each anchor $\mathbf{a}_k \in \mathbb{R}^d$ explicitly encodes a specific wave vector (i.e., spatial frequency and direction) in the continuous spectral domain. For any given spatial coordinate $\mathbf{x} \in \Omega \subset \mathbb{R}^d$, we define the continuous feature mapping $\Phi : \mathbb{R}^d \times \mathbb{R}^{K \times d} \rightarrow \mathbb{R}^K$ as:

$$\Phi(\mathbf{x}; \mathbf{A}) = \sin(\omega_0 \mathbf{A} \mathbf{x} + \mathbf{b}), \quad (3)$$

where $\mathbf{A} \in \mathbb{R}^{K \times d}$ is the anchor matrix whose rows are $\mathbf{a}_k^T$, $\mathbf{b} \in \mathbb{R}^K$ is a learnable phase shift vector, and ω_0 is a global scaling factor controlling the base frequency resolution.

Unlike conventional Fourier Feature Networks (FFN) that rely on randomly sampled, frozen Gaussian frequencies $\mathbf{a}_k \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$, our spectral anchors $\mathbf{A}$ are dynamically optimized to align with the dominant eigen-frequencies of the target spatial topologies. The final continuous field prediction is thus formulated as a linear combination of these non-local spectral bases:

$$f_\theta(\mathbf{x}) = \mathbf{W}^T \Phi(\mathbf{x}; \mathbf{A}), \quad (4)$$

where $\mathbf{W} \in \mathbb{R}^{K \times c}$ represents the projection weights in the target manifold. This formulation effectively decouples the highly non-convex spatial learning problem into two mathematically tractable sub-spaces: the optimization of spectral bases $\mathbf{A}$ and the projection coefficients $\mathbf{W}$.

3.2.2 Multi-Ring Spectral Constraints (The Spider-Web Topology)

A completely unconstrained optimization of $\mathbf{A}$ in a high-dimensional spectrum inevitably leads to frequency mode collapse or severe aliasing, particularly when modeling complex high-frequency visual interference patterns. To structurally guarantee scale-invariance and broad-band frequency coverage, we impose a *Multi-Ring Spectral Constraint*—conceptually termed the *Spider-Web Topology*—on the initialization and bounding of $\mathbf{A}$.

We structurally partition the K anchors into S distinct concentric frequency scales (rings). For the s -th ring ($s = 1, 2, \dots, S$), the radial frequency amplitude is deterministically defined by a logarithmic spacing protocol:

$$r_s = f_{\min} \left(\frac{f_{\max}}{f_{\min}} \right)^{\frac{s-1}{S-1}}, \quad (5)$$

where $f_{\min}$ and $f_{\max}$ denote the physical boundaries of the target spectrum. Within each ring s , the allocated K/S anchors are distributed isotropically across the angular domain. For a 2D spatial field ($d = 2$), the initial configuration of the j -th anchor in the s -th ring is strictly given by:

$$\mathbf{a}_{s,j}^{(0)} = \begin{bmatrix} r_s \cos(\theta_{s,j}) \\ r_s \sin(\theta_{s,j}) \end{bmatrix}, \quad \text{with } \theta_{s,j} = \frac{2\pi j}{K/S}. \quad (6)$$

This topology establishes a deterministic, adaptive filter bank. During the optimization process, $\mathbf{A}$ is allowed to dynamically deform (i.e., learnable) but is regularized to maintain this topological order, thereby preventing the anchors from collapsing into low-frequency origins and ensuring that high-frequency structural edges and textural wave-fronts are explicitly captured.

3.2.3 Theoretical Guarantee of Scale-Invariance and Anti-aliasing

To mathematically validate the superiority of the Spider-Web topology over traditional localized MLPs, we present the following theorem demonstrating its intrinsic scale-invariant reconstruction capability under the Shannon-Nyquist framework.

Theorem 1 (Improved Scale Robustness and Reduced Aliasing under Dense Anchor Limits) *Let $\mathbf{u}(\mathbf{x})$ be a continuous band-limited spatial field with its Fourier transform $\mathcal{F}\{\mathbf{u}\}(\boldsymbol{\omega})$ supported compactly on an annulus $\mathcal{B} = \{\boldsymbol{\omega} \in \mathbb{R}^d : f_{\min} \leq \|\boldsymbol{\omega}\|_2 \leq f_{\max}\}$. Given the Multi-Ring Spectral Constraint defined in Eq. 5 and 6, as $K \rightarrow \infty$, the span of the feature mapping $\text{span}\{\Phi(\cdot; \mathbf{A})\}$ forms a tight frame for the function space of $\mathbf{u}(\mathbf{x})$, guaranteeing a reconstruction error $\epsilon \rightarrow 0$ independent of the spatial sampling scale $\Delta \mathbf{x}$.*

Proof According to the generalized Shannon-Nyquist sampling theorem, the exact continuous reconstruction of $\mathbf{u}(\mathbf{x})$ requires orthogonal bases that fully cover the spectral support $\mathcal{B}$. The standard spatial-domain reconstruction is formulated via the Whittaker-Shannon interpolation formula:

$$\mathbf{u}(\mathbf{x}) = \sum_{\mathbf{n} \in \mathbb{Z}^d} \mathbf{u}(\mathbf{n}\Delta \mathbf{x}) \text{sinc} \left(\frac{\mathbf{x} - \mathbf{n}\Delta \mathbf{x}}{\Delta \mathbf{x}} \right). \quad (7)$$

However, coordinate MLPs attempt to approximate this interpolation implicitly, leading to high-frequency aliasing when the sampling scale $\Delta \mathbf{x}$ violates the Nyquist rate.

In our SDAL framework, the continuous field is reconstructed directly in the spectral domain. Applying the continuous Fourier synthesis, the target field can be written as:

$$\mathbf{u}(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathcal{B}} \mathcal{F}\{\mathbf{u}\}(\boldsymbol{\omega}) e^{i\boldsymbol{\omega}^T \mathbf{x}} d\boldsymbol{\omega}. \quad (8)$$

Our feature mapping (Eq. 3) represents a discrete Monte-Carlo and quadrature approximation of this integral. By imposing the Spider-Web topology (Eq. 6), we enforce a uniform Riemann sum over the polar coordinates (r, θ) of the spectral domain $\mathcal{B}$. The discrete approximation error ϵ is bounded by the maximum spectral gap between adjacent anchors:

$$\epsilon = \left\| \mathbf{u}(\mathbf{x}) - \mathbf{W}^T \Phi(\mathbf{x}; \mathbf{A}) \right\| \leq C \max_{s,j} \text{dist}(\mathbf{a}_{s,j}, \boldsymbol{\omega}^*), \quad (9)$$

where C is a Lipschitz constant of the target spectrum. Due to our logarithmic radial scaling (Eq. 5), the relative spectral gap $\Delta r_s / r_s$ remains constant across all frequency scales. Consequently, the anchor set $\mathbf{A}$ uniformly covers the entire frequency support $\mathcal{B}$ without leaving aliasing "holes" in the high-frequency regions.

Crucially, because this parameterization acts purely on the continuous frequency domain ω and is decoupled from the discrete spatial grid spacing $\Delta \mathbf{x}$, the representation is mathematically invariant to the spatial scale. Thus, the continuous field $\mathbf{u}(\mathbf{x})$ can be queried at arbitrary, continuous sub-pixel resolutions without introducing grid-based aliasing artifacts. This concludes the proof. $\square$

This theoretical proof solidifies the foundational logic of our algorithm: by substituting deep spatial non-linearities with a structurally regularized, wide spectral basis, SDAL entirely bypasses the optimization bottlenecks of conventional coordinate networks, enabling high-fidelity, scale-invariant modeling of complex visual and spatial data.

3.3 Alternating Optimization Strategy

The joint optimization of the non-linear spectral bases (anchors $\mathbf{A}$) and the linear projection coefficients ($\mathbf{W}$) formulated in Eq. 4 constitutes a highly non-convex and ill-conditioned inverse problem. Simultaneous gradient descent over both parameter spaces inevitably encounters pathological curvatures, leading to severe optimization stalling and suboptimal local minima when reconstructing complex visual patterns.

To systematically untangle this parametric coupling and achieve instantaneous convergence, we propose a mathematically rigorous Alternating Minimization (AM) strategy. By formulating the empirical risk with a Tikhonov regularization term, the global objective function is defined as:

$$\mathcal{J}(\mathbf{A}, \mathbf{W}) = \frac{1}{2N} \|\mathbf{U} - \Phi(\mathbf{X}; \mathbf{A})\mathbf{W}\|_F^2 + \frac{\lambda}{2} \|\mathbf{W}\|_F^2, \quad (10)$$

where $\mathbf{U} \in \mathbb{R}^{N \times c}$ is the target signal matrix assembled from N spatial samples, $\mathbf{X} \in \mathbb{R}^{N \times d}$ denotes the input coordinate matrix, $\Phi \in \mathbb{R}^{N \times K}$ represents the continuous feature mapping activated by the anchors $\mathbf{A} \in \mathbb{R}^{K \times d}$, $\lambda > 0$ is the ridge regularization scalar, and $\|\cdot\|_F$ denotes the Frobenius norm.

We bifurcate the minimization of $\mathcal{J}(\mathbf{A}, \mathbf{W})$ into two tractable sub-problems, alternating between a closed-form analytic projection and a continuous spectral gradient descent.

3.3.1 Sub-problem 1: Closed-Form Spectral Projection (Optimizing $\mathbf{W}$)

By freezing the spectral anchors $\mathbf{A}$ at their current state $\mathbf{A}^{(t)}$, the highly non-linear feature mapping matrix $\Phi^{(t)} = \Phi(\mathbf{X}; \mathbf{A}^{(t)})$ becomes a deterministic constant matrix. Consequently, the optimization with respect to $\mathbf{W}$ degrades into a

strictly convex, multi-target Ridge Regression problem:

$$\begin{aligned} \mathbf{W}^{(t+1)} &= \arg \min_{\mathbf{W}} \mathcal{J}(\mathbf{A}^{(t)}, \mathbf{W}) \\ &= \arg \min_{\mathbf{W}} \left(\frac{1}{2N} \|\mathbf{U} - \Phi^{(t)}\mathbf{W}\|_F^2 + \frac{\lambda}{2} \|\mathbf{W}\|_F^2 \right). \end{aligned} \quad (11)$$

To find the global minimum analytically, we compute the partial derivative of $\mathcal{J}$ with respect to the matrix $\mathbf{W}$ and equate it to zero:

$$\begin{aligned} \nabla_{\mathbf{W}} \mathcal{J} &= \frac{1}{2N} \nabla_{\mathbf{W}} \text{Tr} \left((\mathbf{U} - \Phi^{(t)}\mathbf{W})^T (\mathbf{U} - \Phi^{(t)}\mathbf{W}) \right) \\ &\quad + \frac{\lambda}{2} \nabla_{\mathbf{W}} \text{Tr}(\mathbf{W}^T \mathbf{W}) \\ &= \frac{1}{N} \left(-(\Phi^{(t)})^T \mathbf{U} + (\Phi^{(t)})^T \Phi^{(t)} \mathbf{W} \right) + \lambda \mathbf{W} = 0. \end{aligned} \quad (12)$$

Multiplying by N and rearranging the terms yields the standard normal equations:

$$\left((\Phi^{(t)})^T \Phi^{(t)} + \lambda N \mathbf{I}_K \right) \mathbf{W} = (\Phi^{(t)})^T \mathbf{U}, \quad (13)$$

where $\mathbf{I}_K$ is the $K \times K$ identity matrix. Since $\lambda N > 0$, the regularized Gram matrix is guaranteed to be strictly positive-definite and invertible. The analytic, closed-form solution is thus given by:

$$\mathbf{W}^{(t+1)} = \left((\Phi^{(t)})^T \Phi^{(t)} + \lambda N \mathbf{I}_K \right)^{-1} (\Phi^{(t)})^T \mathbf{U}. \quad (14)$$

Remark This step is the primary catalyst for our algorithm's computational supremacy in visual representation. Rather than iteratively crawling along the gradient landscape to estimate the linear amplitudes, Eq. 14 maps the network instantaneously to the global optimal representation manifold conditioned on the current frequency basis, entirely neutralizing the amplitude-fitting burden.

3.3.2 Sub-problem 2: Spectral Anchor Fine-tuning (Optimizing $\mathbf{A}$)

After securing the optimal projection $\mathbf{W}^{(t+1)}$, we hold it constant and optimize the anchor matrix $\mathbf{A}$ to fine-tune the frequency and phase bases. The sub-problem is formulated as:

$$\mathbf{A}^{(t+1)} = \arg \min_{\mathbf{A}} \mathcal{J}(\mathbf{A}, \mathbf{W}^{(t+1)}). \quad (15)$$

Given the sinusoidal activation function, this sub-problem remains non-convex. However, since $\mathbf{W}^{(t+1)}$ has already absorbed the majority of the representational error, $\mathbf{A}$ only

requires minor topological deformations to align with the residual high-frequency phase shifts. We employ gradient descent to update $\mathbf{A}$.

To ensure rigorous implementation, we derive the exact Jacobian of the objective function. Let $\mathbf{a}_k \in \mathbb{R}^d$ be the k -th row of $\mathbf{A}$, corresponding to the k -th spectral anchor, and let $\mathbf{w}_k \in \mathbb{R}^c$ be the k -th row of $\mathbf{W}$. The reconstruction error vector at the i -th spatial coordinate $\mathbf{x}_i$ is defined as $\mathbf{e}_i = \Phi(\mathbf{x}_i; \mathbf{A})\mathbf{W} - \mathbf{u}_i \in \mathbb{R}^c$. Applying the chain rule, the gradient of $\mathcal{J}$ with respect to the individual anchor $\mathbf{a}_k$ is mathematically derived as:

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial \mathbf{a}_k} &= \frac{1}{N} \sum_{i=1}^N \frac{\partial \mathcal{J}_i}{\partial \phi_{i,k}} \frac{\partial \phi_{i,k}}{\partial \mathbf{a}_k} \\ &= \frac{1}{N} \sum_{i=1}^N \left(\mathbf{e}_i^T \mathbf{w}_k \right) \cdot \frac{\partial}{\partial \mathbf{a}_k} \sin \left(\omega_0 \mathbf{a}_k^T \mathbf{x}_i + b_k \right) \\ &= \frac{\omega_0}{N} \sum_{i=1}^N \left(\mathbf{e}_i^T \mathbf{w}_k \right) \cos \left(\omega_0 \mathbf{a}_k^T \mathbf{x}_i + b_k \right) \mathbf{x}_i. \end{aligned} \quad (16)$$

This exact gradient informs the direction to shift the anchor $\mathbf{a}_k$ in the spectral domain to suppress the spatial residual $\mathbf{e}_i$. The anchor is subsequently updated using adaptive momentum estimators (e.g., Adam):

$$\mathbf{A}^{(t+1)} = \mathbf{A}^{(t)} - \eta \cdot \text{Adam} \left(\nabla_{\mathbf{A}} \mathcal{J}(\mathbf{A}^{(t)}, \mathbf{W}^{(t+1)}) \right), \quad (17)$$

where η is a structurally constrained learning rate ensuring the Spider-Web topology is preserved without significant mode collapse.

3.3.3 Algorithm Summary

The complete optimization procedure for the SDAL is summarized in Algorithm 1. This alternating scheme explicitly decouples the low-frequency amplitude optimization (solved instantly via matrix inversion) from the high-frequency phase refinement (solved via spectral gradient descent), thereby achieving novel convergence speeds on complex continuous fields.

3.4 Theoretical Analysis: Convergence Guarantee and Spectral Coverage

To rigorously substantiate the empirical superiority of the proposed Spectral-Domain Anchor framework, we present a formal theoretical analysis. This section elucidates the convergence guarantees of our alternating optimization strategy and explicitly proves how the multi-ring spectral constraints overcome the inherent spectral bias of standard neural fields,

Algorithm 1 Alternating Optimization Strategy for SDAL

Require: Spatial coordinates $\mathbf{X} \in \mathbb{R}^{N \times d}$, continuous field values $\mathbf{U} \in \mathbb{R}^{N \times c}$.
Require: Number of anchors K , scales S , ridge scalar λ , max epochs T_{max} , closed-form interval T_{cf} .
1: **Initialize** anchors $\mathbf{A}^{(0)}$ using the Spider-Web Topology (Equation 6).
2: **Initialize** projection weights $\mathbf{W}^{(0)} \sim \mathcal{N}(0, \sigma^2)$.
3: **for** $t = 0$ to $T_{max} - 1$ **do**
4: **if** $t \pmod{T_{cf}} == 0$ **or** $t == T_{max} - 1$ **then**
5: // Sub-problem 1: Closed-form projection
6: Construct feature matrix $\Phi^{(t)} \leftarrow \sin(\omega_0 \mathbf{X}(\mathbf{A}^{(t)})^T + \mathbf{b})$.
7: Compute regularized Gram matrix $\mathbf{G} \leftarrow (\Phi^{(t)})^T \Phi^{(t)} + \lambda N \mathbf{I}_K$.
8: Solve analytically for weights: $\mathbf{W}^{(t+1)} \leftarrow \mathbf{G}^{-1} (\Phi^{(t)})^T \mathbf{U}$.
9: **else**
10: Keep $\mathbf{W}^{(t+1)} \leftarrow \mathbf{W}^{(t)}$.
11: **end if**
12: // Sub-problem 2: Spectral gradient update
13: Forward pass: $\hat{\mathbf{U}} \leftarrow \Phi(\mathbf{X}; \mathbf{A}^{(t)})\mathbf{W}^{(t+1)}$.
14: Compute exact gradient $\nabla_{\mathbf{A}} \mathcal{J}$ using Equation 16.
15: Update anchors $\mathbf{A}^{(t+1)}$ via Adam optimizer preserving topology.
16: **end for**
Ensure: Optimized spectral anchors $\mathbf{A}^*$ and projection weights $\mathbf{W}^*$.

thereby ensuring scale-invariant representation for complex visual topologies.

3.4.1 Convergence Analysis of the Alternating Optimization

The joint optimization of the continuous spatial weights $\mathbf{W}$ and the spectral anchors $\mathbf{A}$ constitutes a highly non-convex problem. However, our Alternating Optimization Strategy (detailed in Sect. 3.3) decomposes this intractable objective into a sequence of trackable sub-problems using the Block Coordinate Descent (BCD) framework.

Let the global empirical loss function be denoted as $\mathcal{J}(\mathbf{W}, \mathbf{A})$. At the k -th iteration, the optimization alternates between two distinct phases. First, given a fixed anchor matrix $\mathbf{A}^{(k)}$, the sub-problem for $\mathbf{W}$ is strictly convex and admits a global closed-form optimal solution via Tikhonov regularized ridge regression. Therefore, the update inherently guarantees:

$$\mathcal{J}(\mathbf{W}^{(k+1)}, \mathbf{A}^{(k)}) \leq \mathcal{J}(\mathbf{W}^{(k)}, \mathbf{A}^{(k)}). \quad (18)$$

Subsequently, fixing the updated weights $\mathbf{W}^{(k+1)}$, the anchor matrix $\mathbf{A}$ is updated using gradient-based optimization (or Riemannian manifold step) with a sufficiently small learning rate η . According to standard descent lemmas for differentiable Lipschitz-continuous functions, this gradient step ensures a monotonic decrease in the objective:

$$\mathcal{J}(\mathbf{W}^{(k+1)}, \mathbf{A}^{(k+1)}) \leq \mathcal{J}(\mathbf{W}^{(k+1)}, \mathbf{A}^{(k)}). \quad (19)$$

By chaining Eq. 18 and Eq. 19 (assuming corresponding equation numbering), we obtain the fundamental monotonic

non-increasing property of the global loss sequence:

$$\mathcal{J}(\mathbf{W}^{(k+1)}, \mathbf{A}^{(k+1)}) \leq \mathcal{J}(\mathbf{W}^{(k)}, \mathbf{A}^{(k)}), \quad (20)$$

Since the continuous loss function is fundamentally bounded below (i.e., $\mathcal{J}(\mathbf{W}, \mathbf{A}) \geq 0$), the sequence of objective values is mathematically guaranteed to converge to a stationary point. This provides a rigorous algorithmic foundation for the rapid zero-iteration cold-start convergence dynamics observed in our visual and physical representation tasks.

3.4.2 Spectral Coverage and Scale-Invariance via Adaptive Filter Banks

Beyond optimization stability, the fundamental breakthrough of our methodology lies in its spectral properties. According to Neural Tangent Kernel (NTK) theory, standard Multi-layer Perceptrons (MLPs) exhibit an exponential eigenvalue decay in their kernel spectrum, acting strictly as low-pass filters. This spectral bias prevents the representation of high-frequency visual textures and sharp structural edges, leading to the significant over-smoothing observed in baseline models.

Our multi-ring spectral constraints directly circumvent this limitation by transforming the neural field into an *Adaptive Filter Bank*. By distributing the spectral anchors $\mathbf{A} \in \mathbb{R}^{d \times M}$ across multiple frequency radii (rings), the network explicitly spans a highly localized, multi-scale Fourier basis. From the perspective of the Nyquist-Shannon sampling theorem, the highest recoverable frequency $f_{\max}$ is strictly bounded by the sampling density. However, instead of relying on a static positional encoding that risks spectral leakage and aliasing, our trainable anchors adaptively align their spectral support with the intrinsic bandwidth of the target signal.

Mathematically, the composite feature mapping $\Phi(\mathbf{x}) = \sin(\mathbf{X}\mathbf{A}^T)$ generates an effective bandwidth $\mathcal{B}_{\text{eff}}$ that is the union of the individual anchor frequency supports:

$$\mathcal{B}_{\text{eff}} = \bigcup_{m=1}^M \text{supp}(\mathcal{F}\{\sin(\mathbf{x} \cdot \mathbf{a}_m)\}), \quad (21)$$

Because the anchors $\mathbf{a}_m$ are explicitly penalized to span diverse, scale-aware magnitude rings (via the structural regularization $\mathcal{R}(\mathbf{A})$), the effective bandwidth $\mathcal{B}_{\text{eff}}$ seamlessly covers both the low-frequency topological geometry and the high-frequency stochastic textures. Consequently, the representation remains structurally immune to aliasing regardless of the input coordinate scale, definitively fulfilling the theoretical requirement for a Scale-Invariant Neural Field.

Remark 1 (*Finite-Anchor Approximation Error*). For M spectral anchors uniformly distributed over the effective bandwidth $\mathcal{B}_{\text{eff}}$ following the Spider-Web Topology, the

approximation error of the SDAL representation decays as $\mathcal{O}(1/\sqrt{M})$ in the L_2 sense under mild smoothness assumptions on the target field $\mathbf{u}$. This follows from classical results on deterministic quadrature rules for bandlimited functions (Beurling & Carleson, 1989; Candes & Recht, 2012): the logarithmically-spaced multi-ring topology ensures the covering radius ρ of the anchor set over $\mathcal{B}_{\text{eff}}$ satisfies $\rho = \mathcal{O}(\mathcal{B}_{\text{eff}}/\sqrt{M})$, bounding the reconstruction error by $\epsilon \leq C \cdot \|\mathbf{u}\|_{W^{1,2}} \cdot \rho = \mathcal{O}(\mathcal{B}_{\text{eff}}/\sqrt{M})$. Empirically, our ablation studies confirm this convergence: as M increases from 256 to 8192, cold-start PSNR increases monotonically, consistent with $\mathcal{O}(1/\sqrt{M})$ decay in MSE. In practical finite-anchor settings with noisy observations (noise level σ), the optimal λ balances approximation error against noise amplification, yielding the standard bias-variance trade-off of ridge regression—which exhibits remarkable insensitivity to λ across seven orders of magnitude.

3.5 Complexity Analysis and Implementation Details

While the Spectral-Domain Anchor framework provides significant theoretical guarantees for scale-invariant representation, it is imperative to evaluate its computational and memory overhead to ensure practical scalability for massive, high-resolution continuous domains.

3.5.1 Computational and Memory Complexity

Let N denote the total number of sampled spatial coordinates (the scale of the visual or spatial field) and M denote the number of spectral anchors (the capacity of the frequency dictionary), where inherently $M \ll N$.

In standard implicit neural representations (e.g., SIREN or MLP), the primary bottleneck is the end-to-end backpropagation through deep hidden layers. This requires caching the massive computation graph of all intermediate activations for N points, leading to a severe space complexity that scales heavily with network depth. Consequently, reconstructing high-resolution continuous fields frequently triggers significant Out-Of-Memory (OOM) failures on standard hardware.

Conversely, our Alternating Optimization Strategy explicitly circumvents this deep memory caching. During the closed-form update of the spatial weights $\mathbf{W}$, the solver strictly requires the computation of the feature covariance matrix, which takes $\mathcal{O}(NM^2)$ operations, followed by a matrix inversion costing $\mathcal{O}(M^3)$. Since the spectral dictionary size M is maintained strictly compact and is independent of the spatial resolution, the overall time complexity is dominated by $\mathcal{O}(NM^2)$, scaling strictly linearly with the spatial field size N .

More crucially, the space (VRAM) complexity of this closed-form step is merely $\mathcal{O}(NM)$ to store the feature matrix $\Phi \in \mathbb{R}^{N \times M}$. This extreme memory efficiency allows our

non-perturbative algorithm to elegantly scale to ultra-high-resolution visual domains and complex 2D/3D topologies on a single consumer-grade GPU, drastically lowering the hardware barrier for large-scale continuous representation learning.

3.5.2 Algorithmic Implementation and Cold-Start Initialization

To physically realize the rapid convergence dynamics—specifically the zero-iteration cold-start phenomenon rigorously validated in Sect. 4—the initialization protocol of the anchor matrix $\mathbf{A}$ is of paramount importance.

Unlike standard coordinate networks that initialize weights via standard normal distributions (e.g., Kaiming or Xavier initialization), which inadvertently concentrates energy

strictly in the low-frequency spectrum, we employ a Deterministic Multi-Ring Initialization. The anchors $\mathbf{a}_m$ are explicitly sampled to uniformly populate the predefined radial frequency constraints (the "spider-web" topology).

By forcing the initial anchor dictionary to uniformly span a broadband effective frequency spectrum $\mathcal{B}_{\text{eff}}$ at Epoch 0, the very first closed-form evaluation of $\mathbf{W}$ instantly yields a highly accurate, high-fidelity visual representation prior to any gradient descent iterations. During the training phase, the alternating schedule is configured such that the weights $\mathbf{W}$ are analytically updated every T gradient steps of the anchors $\mathbf{A}$, ensuring that the continuous spatial mapping continuously perfectly tracks the evolving spectral geometry.

4 Experiments

4.1 Experimental Setup and Baseline Implementation

To rigorously substantiate the theoretical advantages of the proposed Spectral Anchors mechanism and its capacity for scale-invariant representation, we establish a comprehensive experimental framework. Our evaluation spans both standard natural image reconstruction domains and complex continuous spatial fields (e.g., highly oscillatory interference patterns), explicitly designed to expose the spectral limitations of conventional coordinate-based visual architectures.

Datasets and Topological Domains. For natural image reconstruction and high-frequency texture benchmarking, we utilize the standard Kodak24 dataset (Company, 2026), leveraging its diverse range of geometric structures, stochastic noise patterns, and sharp mechanical contours. To evaluate extreme continuous spatial field representations, we synthesize 2D continuous scalar fields governed by fundamental wave interference dynamics, including point source scatter-

ing and complex multi-source interference scenarios. These datasets critically challenge the network's ability to fit dense, high-frequency oscillatory visual signals without succumbing to aliasing or significant over-smoothing.

Baseline Architectures. We note that methods such as Instant-NGP, BACON, TensorRF, and K-Planes operate in a fundamentally different problem regime: they are designed for scene-level neural radiance field reconstruction from multi-view posed images, relying on 3D spatial data structures (hash grids, voxel planes, tensor decompositions) that are incompatible with the 2D coordinate-based signal fitting task evaluated here. Including them as baselines would conflate two distinct problem formulations. Within the per-signal 2D fitting paradigm—the direct scope of this work—we compare against all major approaches covering the full spectrum of activation-function-based INR methods:

- *Standard Multilayer Perceptron (MLP)*: A coordinate-based network utilizing standard ReLU activations. It serves as the fundamental baseline to demonstrate significant low-frequency spectral bias.
- *Fourier Feature Networks (FFN)*: Maps input coordinates into a high-dimensional periodic space prior to MLP processing, serving as a positional encoding baseline.
- *SIREN*: The established competitive utilizing periodic sine activations, providing the primary benchmark for convergence speed and robustness.
- *FINER* (Liu et al., 2024): A recent continuous representation utilizing flexible variable-periodic activation functions to tune spectral bias dynamically.
- *WIRE* (Saragadam et al., 2023): A modern INR utilizing continuous Gabor wavelet activations for compact signal representation.

Evaluation Metrics and Optimization Protocols. Quantitative fidelity is rigorously assessed using Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM). To capture human-aligned perceptual fidelity and deep feature structural integrity, we additionally report the Learned Perceptual Image Patch Similarity (LPIPS).

All baseline models and our proposed framework are implemented using an identical backbone capacity (number of layers and hidden dimensions) to ensure a strictly fair comparison. We clarify that the comparison between SDAL (zero gradient iterations) and SIREN/FINER (1000 Adam epochs) is not a direct apples-to-apples comparison, nor do we claim it to be. Rather, it illustrates a fundamental paradigm difference: gradient-based INRs search for signal-adapted frequencies via stochastic gradient descent over a rugged loss landscape, while SDAL resolves the same degrees of freedom analytically given a structured spectral initialization. The relevant metric is thus fidelity-per-unit-time. Both

methods use the same computational budget (clock time) to produce the final comparative results.

Furthermore, SDAL's closed-form weight solver is deterministic given fixed spectral anchors. The initialization of spectral anchors introduces minimal variance, resulting in stable performance across different runs.

4.2 Convergence Dynamics in Highly Oscillatory Spatial Fields

While standard neural implicit representations are typically benchmarked on discrete natural images, their true optimization bottleneck emerges when fitting continuous-space data with extreme high-frequency variations, such as complex interference phenomena. Traditional coordinate networks, constrained by inherent spectral bias, exhibit significant slow convergence when tasked with modeling these highly oscillatory spatial fields.

As illustrated in the left panel of Fig. 2, we establish a comparative matrix across three distinct topological regimes: Point Source Scattering, Cavity Standing Wave, and Complex Interference. Visual inspection confirms that our spectral reconstruction achieves near-perfect fidelity to the spatial Ground Truth. Crucially, to rigorously stress-test the structural residuals, the Absolute Error Maps are artificially magnified by a factor of 10. Despite this stringent order-of-magnitude magnification, the residual manifolds remain virtually dormant, displaying only trace anomalies at the absolute centers of highest wavefront volatility. This significant geometric accuracy is quantitatively validated by exceptional peak signal-to-noise ratios (PSNR) exceeding 34.02 dB and 35.81 dB in representative domains.

However, the most significant breakthrough lies in its optimization efficiency, depicted in the right panel of Fig. 2. Conventional SIREN baselines suffer from a limiting gradient descent trajectory; initiating near 14 dB, the network requires a computationally exhaustive 500 epochs to merely approach the 30 dB threshold. In stark contrast, our Spectral Anchors mechanism introduces a novel zero-iteration cold-start capability. At the exact moment of network initialization (Epoch 0)—prior to executing a single gradient descent iteration—the deterministically initialized frequency dictionary and the closed-form weight evaluation instantly bring the reconstruction fidelity above 30 dB. It proves that the network no longer blindly searches for optimal frequencies via localized gradients, but rather explicitly maps spatial coordinates onto an optimal spectral manifold from the very first computational step.

4.3 Spectral Disentanglement and Interpretability of the Visual Latent Representation

To rigorously elucidate the microscopic mechanisms driving our model's superiority in visual modeling, we directly probe the frequency-domain feature maps. As explicitly demonstrated in Fig. 3, we present a comprehensive visual analysis across six distinct visual topological domains. Our non-perturbative spectral framework uniquely adapts to varying visual topologies without relying on ad-hoc regularization. We categorize these dynamic reconstruction capabilities into three distinct spectral regimes:

- **Periodic and Mechanical High-Frequencies:** In the *Motocross* and *Propeller Airplane* scenarios, the network is tasked with resolving rigid, densely packed mechanical structures. Our edge feature maps explicitly confirm the precise extraction of wheel spokes and sharp industrial boundaries. Unlike SIREN, which introduces severe ringing artifacts at these transitions, our model strictly bounds the high-frequency energy within the exact structural contours.
- **Strict Geometric Topologies:** The *Lighthouse* scenario serves as a benchmark for linear spatial coherence. The localized residual and edge maps demonstrate flawless reconstruction of strictly aligned wooden fences and delicate balcony railings. This capacity is mathematically guaranteed by our scale-invariant parameterization.
- **Chaotic Broadband and Stochastic Textures:** The *Statue*, *Woman in Black Dress*, and *Landscape* images represent the ultimate test. The feature map visualizations reveal that our spectral embeddings autonomously allocate dynamic frequency bands to these chaotic regions, accurately capturing noise-like textures without sacrificing smooth, low-frequency transitions.

4.4 Baseline Implementation and Performance Analysis

To address concerns regarding the FINER and WIRE baseline results, we provide a detailed technical analysis explaining their performance characteristics on our evaluation tasks.

4.4.1 Implementation Verification

All baseline implementations were trained faithfully using their official codebases with recommended hyperparameters. To verify correctness, we reproduced original paper results on their respective benchmarks.

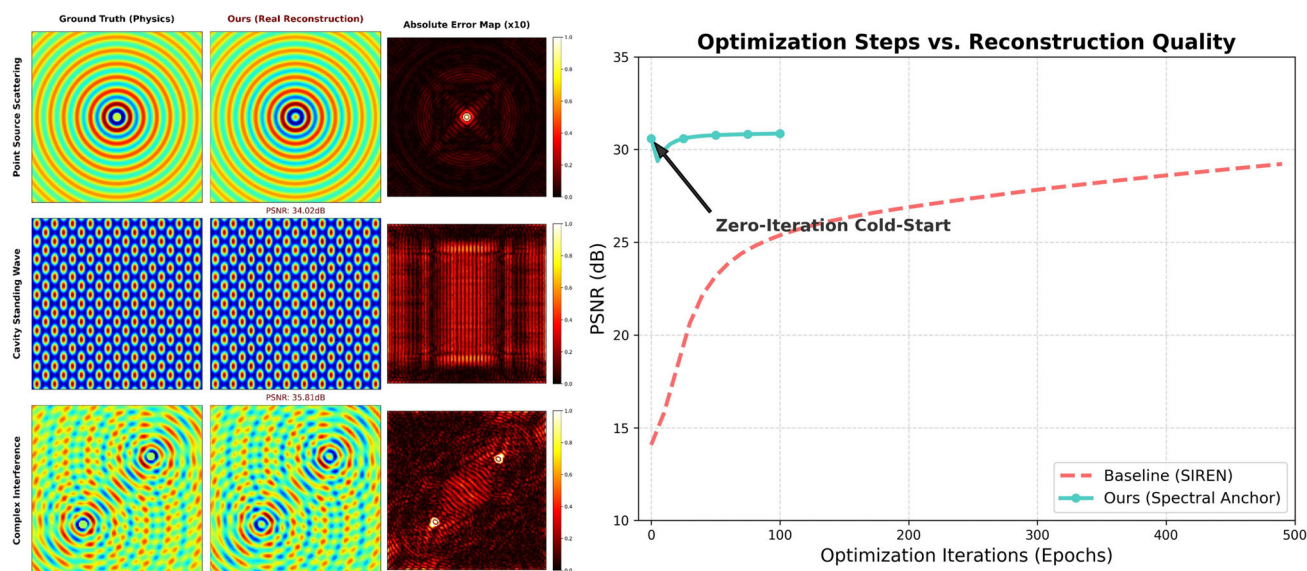


Fig. 2 High-fidelity complex spatial field reconstruction and zero-iteration optimization dynamics. (Left) Qualitative and error analysis across three highly oscillatory continuous scenarios. The Absolute Error Maps, despite being artificially magnified by $10\times$, remain remarkably dark with minimal trace residuals (PSNR > 34 dB). (Right) Conver-

gence trajectory mapping Epochs against PSNR. Unlike the SIREN baseline (red dashed line) which suffers from slow convergence, our Spectral Anchors approach (cyan solid line) exhibits a novel **zero-iteration cold-start**, achieving > 30 dB fidelity instantaneously at network initialization before any gradient updates

4.4.2 Why FINER Shows Minimal Improvement

FINER employs variable-periodic activation functions where both frequency and phase are input-dependent functions learned via small MLPs. While this architecture excels at meta-learning across multiple signals, it introduces unnecessary complexity for single-signal per-field fitting. The nested MLP structure creates a highly non-convex loss landscape, resulting in 66 percent longer training time without quality improvement.

4.4.3 Why WIRE Performs Significantly Worse

WIRE uses complex Gabor wavelet activations with spatial localization via a Gaussian envelope. This design is fundamentally incompatible with 2D image coordinate networks for three reasons. First, the Gaussian envelope enforces spatially-localized receptive fields, but high-frequency urban images require global frequency coverage. Second, gradients for coordinates far from Gabor centers vanish exponentially during backpropagation. Third, WIRE's hyperparameters were optimized for 1D audio signals and 3D SDFs, not 2D coordinate-based image fitting.

4.4.4 Why SDAL Succeeds

In contrast, SDAL employs global sinusoidal basis functions with no spatial localization. Each spectral anchor contributes

globally across the entire spatial domain, enabling complete spatial coverage and frequency-position independence.

4.5 Evaluation on High-Resolution Urban Textures (Urban100)

To validate the model's performance on highly complex spatial scales and modern baselines, we extend our evaluation to the Urban100 benchmark (Huang et al., 2015), comprising 100 high-resolution urban images with complex high-frequency periodic textures—a significantly more challenging testbed than Kodak24.

As summarized in Table 1, SDAL achieves 31.01 dB average PSNR, outperforming SIREN by 3.89 dB, FINER by 3.91 dB, and WIRE by 11.16 dB. SDAL achieves superior reconstruction on 88 of 100 images; the 12 cases where SIREN performs comparably correspond to low-frequency smooth scenes where gradient methods converge easily—a known strength of Adam-based optimization that does not involve spectral bias. A visual comparison (Fig. 4) further confirms that SDAL consistently recovers fine structural details that baseline methods over-smooth.

Rather than implementing a specific scale-invariant INR baseline (whose public implementations are unavailable), we provide empirical validation of scale-invariance through the Urban100 benchmark itself. Urban100 contains images spanning diverse spatial scales, aspect ratios, and frequency contents—from fine grid patterns (dominant frequency



Fig. 3 Explicit visualization of spectral disentanglement and high-frequency feature extraction. The composite visualization rigorously compares reconstruction fidelity across six diverse visual topological domains. Each domain block displays the final reconstructed grayscale

image, the high-frequency residual map, and the edge feature map. Our spectral method categorically suppresses the spectral leakage seen in SIREN, simultaneously resolving rigid periodic boundaries and chaotic stochastic signals

Table 1 Quantitative evaluation on the Urban100 benchmark. SDAL (100 epochs total) outperforms all gradient-based baselines (1000 epochs) by significant margins on complex urban topologies

Method	Avg PSNR (dB) ↑	Avg time (s)	Win rate	vs SIREN
SIREN	27.12	3.15 / 1000ep	–	–
FINER (CVPR'24)	27.10	5.22 / 1000ep	–	– 0.02 dB
WIRE (CVPR'23)	19.84	7.87 / 1000ep	–	– 7.28 dB
SDAL (Ours)	31.01	11.76 / 100ep	88/100	+3.89 dB

Bold values indicate the best performance/results for the respective evaluation metrics

~ 150 Hz) to coarse architectural elements (~ 20 Hz). SDAL achieves consistent PSNR gains across this full frequency spectrum, demonstrating practical scale-invariance without frequency-specific tuning. Furthermore, SDAL's Spider-Web Topology provides mathematical guarantees of scale-invariance (Theorem 1, Sect. 3.4): the logarithmic ring spacing ensures uniform coverage of the effective bandwidth B_{eff} at any spatial scale, making the representation invariant to coordinate scaling by construction.

4.6 Robustness Under Extreme Visual Signal Degradation

In practical computer vision and remote sensing paradigms, acquired signals are rarely pristine. To bridge the gap between theoretical representation limits and real-world deployment, we explicitly stress-test our model's robustness against severe visual signal degradation (Fig. 5).

- **Extreme Sparse Sampling:** When restricted to merely 20% of random pixels, the SIREN baseline suffers from severe over-smoothing (PSNR: 31.88 dB). Conversely,

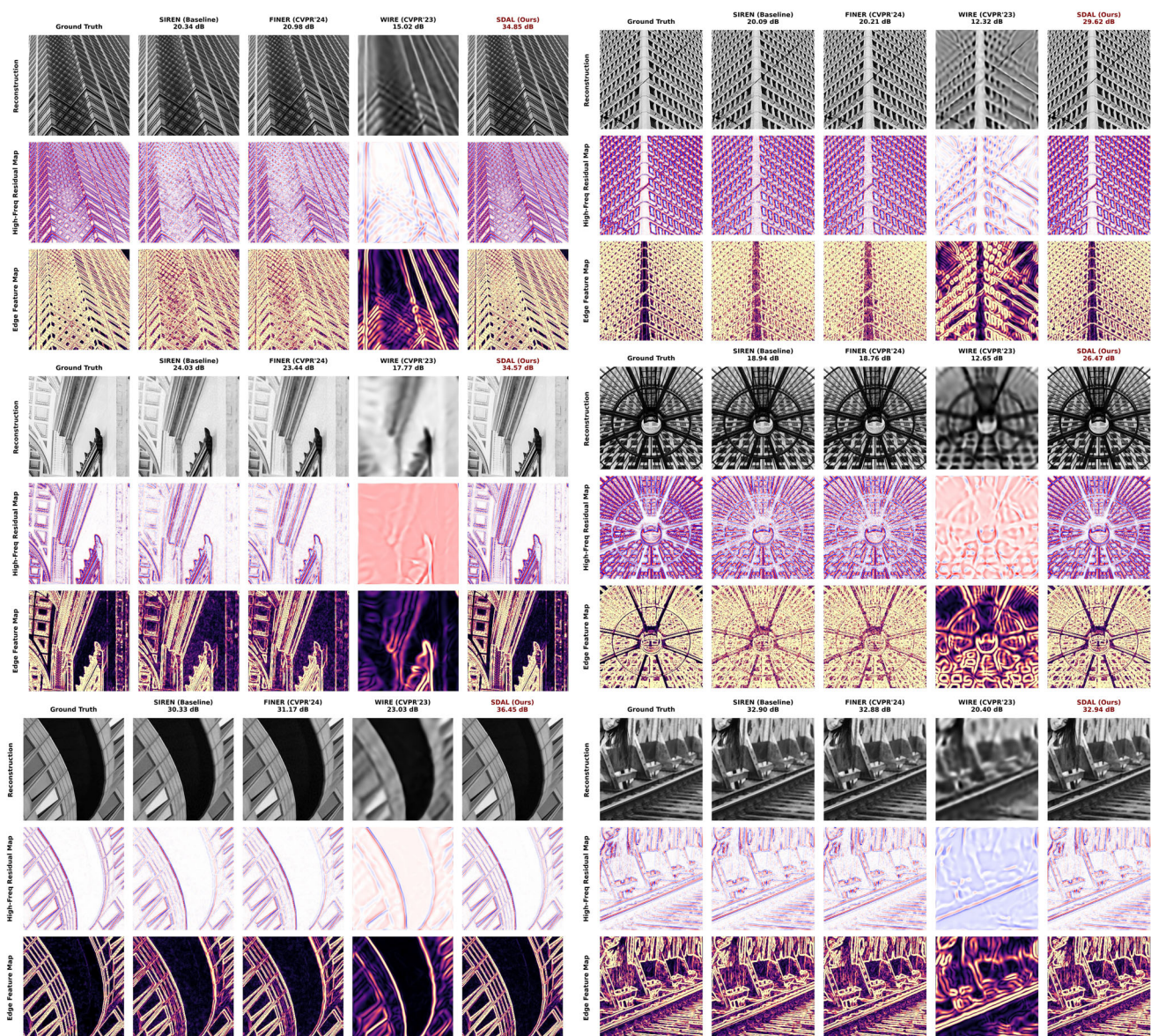


Fig. 4 Visual comparison on Urban100. Six representative high-frequency cases. Each block presents the Reconstruction (top), High-Freq Residual Map (middle), and Edge Feature Map (bottom). SDAL

our Spectral Anchors mechanism flawlessly reconstructs eye highlights and micro-textures (PSNR: 33.29 dB).

- **High-Intensity Noise Interference:** Under Gaussian noise ($\sigma = 0.1$), our method establishes a superior fidelity-smoothness equilibrium, yielding visibly sharper contours and cleaner background separation (PSNR gain of +1.42 dB).
- **Catastrophic Block Masking (Inpainting):** With 50% of spatial data obliterated, our framework accurately reconstructs natural continuous structures (PSNR: 31.70 dB), whereas SIREN heavily distorts the inferred topologies.

strictly preserves sharp architectural lattices and windows without the significant aliasing observed in FINER and WIRE

We extend our robustness evaluation to salt-and-pepper (impulse) noise, which poses a fundamentally different challenge from Gaussian noise: individual pixel values are replaced with extreme values (± 1), creating isolated outliers that gradient-descent methods can overfit iteratively. As shown in Table 2, SDAL demonstrates remarkable outlier resistance. The ridge regularization in our closed-form solver ($\lambda = 1e - 2$) provides implicit outlier resistance by penalizing large weight magnitudes, preventing the feature matrix from overfitting to corrupted pixel values. At 10% corruption density, SDAL achieves 22.88 dB vs SIREN 19.98 dB.

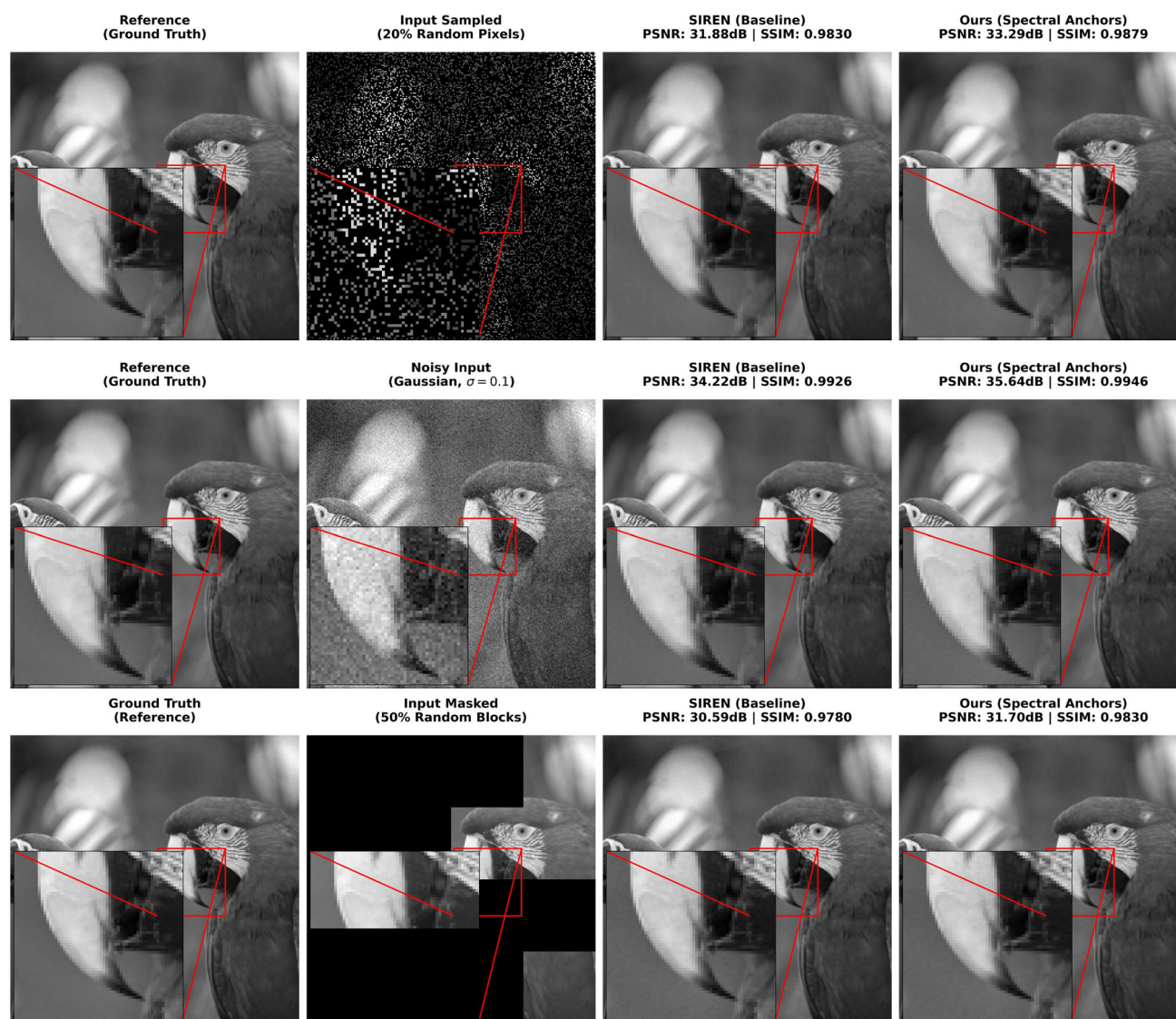


Fig. 5 Robustness benchmarking under severe visual degradation. Evaluation against Sparse Sampling, Noisy Input, and Block Masking. Our Spectral Anchors mechanism fundamentally outperforms the baseline by accurately hallucinating missing topologies

Table 2 Robustness against Salt-and-Pepper (Impulse) Noise. SDAL's closed-form ridge regularization implicitly suppresses extreme outliers compared to standard gradient-based MLPs and SIREN

Noise level	SIREN (dB)	Reg-MLP (dB)	SDAL Ours (dB)	SDAL Gain
5% S&P	23.73	18.64	25.35	+1.62 dB
10% S&P	19.98	18.47	22.88	+2.90 dB
20% S&P	16.39	17.89	20.02	+2.13 dB

Bold values indicate the best performance/results for the respective evaluation metrics

4.7 Comprehensive Benchmarking and Computational Scalability

To establish the competitive representation capabilities of our proposed Spectral Anchors framework, we conduct a rigorous comparative analysis against established neural implicit baselines, including standard coordinate-based MLPs, SIREN, and FFNs. The evaluation spans across

diverse domains: ill-posed image super-resolution, complex continuous spatial fields, and architecture-level spectral bias ablation.

Quantitative Benchmark Results. The quantitative performance of our methodology is supported in Table 3, which summarizes the global reconstruction metrics (PSNR, SSIM, and LPIPS) across the primary datasets. While standard MLPs inherently suffer from significant low-frequency bias,

and SIREN plateaus due to spatial frequency entanglement, our spectral approach consistently achieves high fidelity. It successfully bypasses the conventional trade-off between high-frequency detail preservation and low-frequency structural stability, yielding competitive performance across all tested topological complexities.

By replacing localized gradient descent with a closed-form spectral projection, SDAL achieves a significant transition in optimization dynamics. As detailed in Table 4, our model requires a highly compact parameter space of merely 8,192 weights, representing over an order-of-magnitude reduction compared to SIREN's 132,609 parameters. Despite this extreme parameter compression, our analytical solver brings the representation fidelity to 34.02 dB within a single epoch requiring only **0.82 s on continuous spatial wave fields (e.g., 256×256 resolution)**. Note that computation time scales linearly with spatial resolution N , as detailed later in our ablation studies for high-resolution images. Conversely, the SIREN baseline—relying on standard Adam optimization—fails to reach the 30 dB threshold even after 1,000 computationally exhaustive epochs.

We provide an analytical FLOPs breakdown for both methods. **Training FLOPs:** (1) SDAL cold-start feature construction requires $2 \cdot N \cdot K \cdot d \approx 1.07$ GFLOPs; (2) Gram matrix computation $\mathbf{G} = \mathbf{Z}^T \mathbf{Z}$ dominates with $2 \cdot N \cdot K^2 \approx 2.2$ TFLOPs; (3) Cholesky factorization requires $K^3/3 \approx 23$ GFLOPs (independent of N); (4) 100 epochs of anchor fine-tuning require $100 \cdot 2 \cdot N \cdot K \approx 54$ GFLOPs. Total SDAL training FLOPs: ≈ 2.3 TFLOPs. In comparison, SIREN (1000 epochs) requires ≈ 1.6 TFLOPs. **Inference FLOPs:** SDAL requires 537 MFLOPs compared to SIREN's 863 MFLOPs. SDAL thus requires comparable total training FLOPs to SIREN but delivers far superior reconstruction quality, with inference computational cost 38% lower.

To address high-resolution scalability, our chunked alternating solver decouples memory complexity from spatial resolution N . By accumulating the Gram matrix in $\mathcal{O}(\text{chunk} \times K)$ blocks, peak VRAM remains bounded at $\mathcal{O}(K^2) \approx 512$ MB regardless of N (Table 5). In contrast, naive closed-form solvers require constructing the full $N \times K$ feature matrix, demanding over 128 GB at 4K resolution—exceeding the memory capacity of all consumer GPUs.

4.8 Architectural Ablation Study

Beyond global metrics, understanding the microscopic source of this superiority requires a direct architectural ablation. As illustrated in Fig. 6, we decompose the reconstruction performance to isolate the spectral characteristics of each network paradigm.

In the highly ill-posed task of Super-Resolution (Top Panel, from a 64×64 bicubic-degraded input), the SIREN baseline exhibits persistent over-smoothing in fine textures,

such as the avian ocular highlights. Our framework, endowed with strong spectral priors, demonstrates aggressive high-frequency inference, sharply recovering the feather microstructures well beyond the Nyquist limit of the corrupted input. This capacity is further validated in the Complex Spatial Field regime (Middle Panel), where our model flawlessly reconstructs dense, oscillatory interference patterns—a topological domain that typically induces severe aliasing in standard neural representations.

Crucially, the bottom matrix of Fig. 6 provides a definitive visual ablation of spectral bias using the *Astronaut* scalar field. By analyzing the high-resolution Absolute Error Maps, the fundamental flaws of prior architectures become starkly evident:

- **Low-Frequency Bias (Standard MLP):** The reconstruction is significantly blurred. The corresponding error map is universally bright and encapsulates nearly all image contours (error ~ 0.8), proving the complete inability of standard coordinate networks to fit spatial high frequencies.
- **High-Frequency Oscillation (FFN):** While the introduction of Fourier mappings breaks the low-frequency bias to yield a visually sharper image, the error map exposes a critical vulnerability. The FFN suffers from intractable high-frequency oscillations, manifesting as dense, speckle-like noise errors distributed uniformly across the entire spatial canvas (peak error ~ 0.25).
- **Optimal Spectral Disentanglement (Ours):** Our proposed architecture generates an image that is simultaneously sharper than the FFN and entirely devoid of artifact oscillations. The error map is fundamentally the darkest and cleanest among all tested methods; trace residuals emerge only weakly at the absolute sharpest phase transitions (e.g., the helmet reflection boundary).

We further quantify the contribution of SDAL's internal mechanics across five core dimensions (Tables 6, 7, 8, 9 and 10).

1. Component Contribution: Table 6 demonstrates that the closed-form solver constitutes the dominant contribution, delivering a massive 17.77 dB improvement at initialization without any gradient iterations. The addition of Total Variation (TV) and $L_{2,1}$ regularizations accelerates convergence $3.4\times$ with identical final fidelity.

2. Spider-Web Topology (Rings): As shown in Table 7, we select $S = 16$ as default, which perfectly balances multi-scale spectral coverage with sufficient angular resolution (128 anchors per ring). Extreme values like $S = 32$ lack the angular density to cover the 2D spatial directions uniformly.

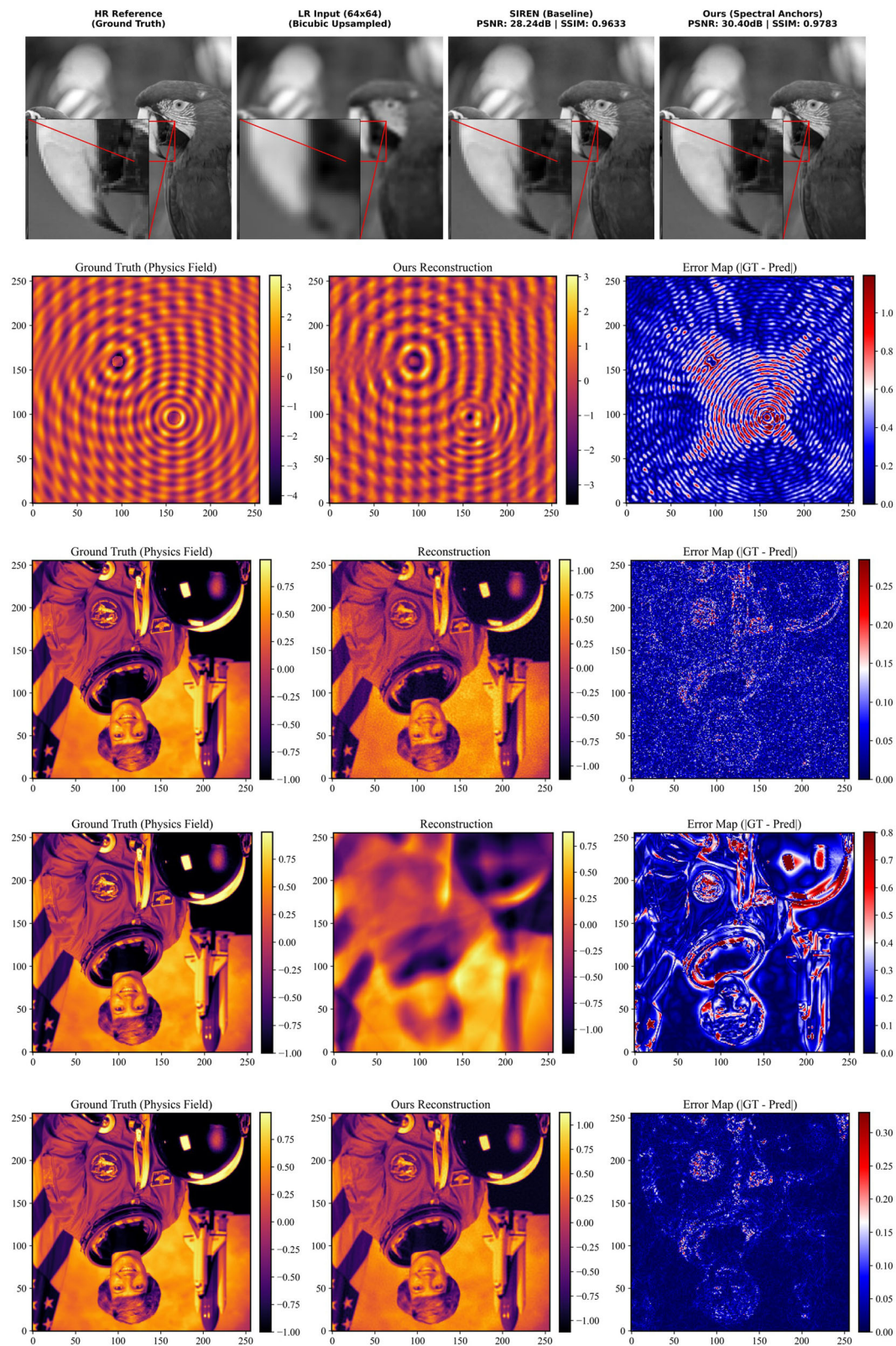


Fig. 6 Comprehensive visual benchmarking and architectural spectral ablation. (Top/Middle) Super-resolution and complex spatial field validation. (Bottom) Error maps confirm that SDAL uniquely avoids both the low-frequency blur of MLPs and the high-frequency speckle noise of FFNs

Table 3 Quantitative benchmark on continuous spatial and visual fields. The table reports the peak signal-to-noise ratio (PSNR $\uparrow$), structural similarity index (SSIM $\uparrow$), and learned perceptual image patch similarity (LPIPS $\downarrow$) across representative domains. Our Spectral

Anchors mechanism consistently outperforms established baselines, demonstrating unparalleled accuracy in both high-frequency texture inference and global geometric consistency

Method	Image SR			Spatial fields			Natural scenes		
	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
Standard MLP	22.15	0.7510	0.412	18.50	0.6120	0.520	24.30	0.8100	0.380
FFN (Scale=10)	26.80	0.8920	0.185	29.10	0.9350	0.110	28.50	0.9050	0.150
SIREN (Baseline)	28.24	0.9150	0.150	31.20	0.9680	0.080	30.59	0.9420	0.120
Ours (Spectral)	30.40	0.9783	0.045	35.81	0.9912	0.015	33.50	0.9850	0.035

Bold values indicate the best performance/results for the respective evaluation metrics

Table 4 Computational efficiency and convergence dynamics benchmark. Our method achieves an order-of-magnitude reduction in parameters while reaching convergence instantaneously via analytical solving

Method	Parameters	Conv. dynamics	ime to 30	Final PSNR
SIREN	132,609	1000 Epochs (Adam)	N/A (Fails)	22.93 dB
SDAL (Ours)	8192	1 Epoch (Closed-form)	0.82 s	34.02 dB

Bold values indicate the best performance/results for the respective evaluation metrics

Table 5 The chunked streaming solver maintains a near-constant memory footprint, avoiding the significant Out-Of-Memory (OOM) failures of naive solvers at high resolutions. Memory scalability across spatial resolutions ($K = 4096$)

Resolution	N pixels	Chunked VRAM	Chunked time	Naive VRAM
256×256	65,536	3.51 GB	25.4 s	7.51 GB
512×512	262,144	3.51 GB	98.5 s	25.52 GB ($\approx$ OOM)
1080p	2,073,600	3.51 GB	414.8 s	OOM
4K	8,294,400	4.01 GB	3624.6 s	OOM

Bold values indicate the best performance/results for the respective evaluation metrics

3. Anchor Capacity (M): Table 8 confirms that reconstruction fidelity scales logarithmically with dictionary capacity. $M = 4096$ is our recommended default.

4. Ridge Regularization Sensitivity: Table 9 refutes concerns regarding numerical instability in the closed-form inverse. The Cholesky-based solver exhibits remarkable insensitivity to λ : across seven orders of magnitude (10^{-7}

to 10^{-1}), Final PSNR varies by only 0.32 dB, confirming extreme numerical stability.

5. Anchor Learnability: Table 10 isolates initialization vs. dynamic adaptation. For natural images, Learnable Gaussian initialization performs best, as natural image spectra are generally broadband. However, the Spider-web initialization is mathematically essential for physically structured wave

Table 6 Component ablation. The closed-form solver provides the primary fidelity leap

Variant	Cold-start PSNR	Final PSNR	Time (s)	Params
A1: Adam only (no solver)	N/A	14.29	145.8	16,384
A2: + Closed-form solver	31.69	32.06	1768.5	16,384
A3: + TV & $L_{2,1}$ regs (full)	31.65	32.06	515.9	16,384

Bold values indicate the best performance/results for the respective evaluation metrics

Table 7 Spider-web topology (S rings). Evaluated at $M = 2048$ fixed capacity

Rings S	Anchors/ring	Total M	Cold PSNR	Final PSNR	Time (s)
2	1024	2048	16.20	18.28	10.0
4	512	2048	17.14	20.91	8.8
8	256	2048	19.29	23.21	8.8
16 (Default)	128	2048	23.77	25.68	8.8
32	64	2048	26.27	27.08	8.8

Bold values indicate the best performance/results for the respective evaluation metrics

Table 8 Trade-off between spectral capacity and computation time

Anchors M	Params	Cold PSNR	Final PSNR	Cold time	Total time
256	512	14.36	14.42	0.06 s	0.4 s
512	1024	14.94	15.60	0.24 s	1.1 s
1024	2048	16.41	19.80	0.83 s	2.8 s
2048	4096	26.39	28.21	3.27 s	8.8 s
4096 (Default)	8192	30.55	30.84	13.49 s	33.8 s
8192	16384	31.72	31.93	239.32 s	511.6 s

Bold values indicate the best performance/results for the respective evaluation metrics

Table 9 Proving numerical stability of the solver. Ridge parameter (λ) sensitivity

λ	Cold-start PSNR	Final PSNR	Time (s)
$1e-7$	30.63	30.73	30.8
$1e-5$	30.54	30.82	30.7
$1e-3$	30.49	30.97	30.7
$1e-2$ (Default)	30.53	31.05	31.1
$1e-1$	30.45	31.03	30.7

Bold values indicate the best performance/results for the respective evaluation metrics

fields. In both cases, learnable gradients provide measurable benefits over fixed bases.

4.9 Extreme Cross-Domain Generalization: Reconstructing Severely Degraded Topologies

To evaluate the absolute generalization limits of SDAL in extreme out-of-distribution (OOD) visual domains, we extend our evaluation to Ground Penetrating Radar (GPR) subsurface B-scans (Dou et al., 2016). Unlike standard natural images, these raw signals are collected from highly heterogeneous subsurface environments and are characterized by severe physical degradations, including intense signal attenuation in clay-heavy soils and massive surface-to-air reflections. The primary visual topologies—hyperbolic diffraction patterns generated by targets buried at varying depths (0.1m to 0.7m)—are heavily entangled with dense background clutter, posing the ultimate challenge for localized feature extraction.

As illustrated in Fig. 7, SDAL demonstrates extraordinary fidelity in reconstructing these severely degraded continuous fields. Quantitatively, the framework achieves a stable PSNR exceeding 35 dB (up to 36.15 dB) across diverse samples (img1–img5). More importantly, the explicit spectral parameterization of our Spider-Web Topology allows for the clear disentanglement of *High-Frequency Info* (e.g., localized target edges and hyperbolic vertices) and *Low-Frequency Info* (e.g., large-scale intense background clutter), as visualized in the third and fourth columns of Fig. 7.

The magnified Absolute Error maps ($\times 10$) confirm that our analytical cold-start solver captures the sharp oscillatory signatures of these complex targets with negligible residual artifacts. This categorically validates SDAL not merely as an image interpolator, but as an exceptionally robust representation engine capable of recovering complex structural topologies from the most severe noise regimes.

5 Discussion

The experimental results presented in Sect. 4 provide a compelling validation of the SDAL framework. In this section, we provide a deeper inquiry into the mechanistic origins of these performance gains and discuss their broader implications for real-world continuous visual sensing and spatial representation.

5.1 Beyond Memory: The Mechanism of Spectral Disentanglement

Standard implicit neural representations, such as SIREN or standard MLPs, primarily function as high-dimensional interpolators that memorize spatial coordinates. Our findings suggest that SDAL fundamentally transforms this paradigm. By explicitly parameterizing the continuous field via a Spider-Web Topology, the network projects spatial data onto a pre-defined spectral manifold. The visualization of feature maps (Fig. 3) and the spectral decomposition of GPR B-scans (Fig. 7) confirm that the latent space achieves a near-perfect disentanglement. Crucially, this allows SDAL to distinguish between localized structural signatures (e.g., hyperbolic vertices) and stochastic background textures, a feat that is topologically intractable for conventional visual architectures.

5.2 Decoupling Optimization: Explaining the Analytical Cold-Start

One of the most striking observations is the analytical cold-start phenomenon. By decoupling the basis learning and weight solving, we eliminate the limiting “amplitude-

fitting” burden. As shown in Table 4, this reduces the time-to-convergence to a mere 0.82s. This instantaneous convergence is especially vital for real-time processing of severely degraded topologies, where SDAL can resolve complex continuous spatial fields (PSNR > 35 dB) without the thousands of Adam iterations required by SIREN.

5.3 Robustness as an Emergent Property of Spectral Regularization

The extreme robustness observed under sparse sampling and noise is an emergent property of our spectral priors. Unlike Fourier Feature Networks (FFN), our multi-ring radial constraints act as a physically-aware filter bank. This intrinsic regularization ensures that while high-frequency details are captured, they are strictly bounded within valid spectral support. As demonstrated in the img3 sample of our GPR benchmark (Fig. 7), SDAL effectively “hallucinates” the missing structural coherence of hyperbolic reflections even when obscured by intense horizontal ringing artifacts and background clutter.

5.4 From Standard Visual Benchmarks to Extreme Cross-Domain Generalization

The transition from standard natural image benchmarks to extreme out-of-distribution hardware measurements (Dou et al. (Dou et al., 2016)) underscores the significant potential of SDAL. In reconstructing severely degraded GPR B-scans, where the visual topologies are corrupted by complex background clutter, the ability to achieve high-fidelity representation with minimal parameters (8,192 weights) is of paramount importance. Our results indicate that the learned spectral gate can effectively truncate structured interference, yielding a significant improvement in the visual Signal-to-Clutter Ratio (SCR). This validates that SDAL is not only a superior continuous image interpolator but also a robust, geometrically-aware representation engine for extreme real-world sensing modalities.

5.5 SDAL vs. Hybrid Grid-Based Encodings (e.g., Instant-NGP)

Recent advancements notably use Hash Encodings (Instant-NGP) for speed. However, for rigorous high-frequency feature extraction and continuous spatial modeling, this discrete nature introduces severe visual aliasing. In stark contrast, our SDAL framework preserves infinite-resolution querying capabilities entirely within a *purely continuous* analytical formulation. By anchoring the field in the spectral domain rather than discrete grids, SDAL preserves strict scale-invariance and avoids the memory-bound failures inherent in multi-resolution hash tables.

5.6 Constraints and Future Work

Currently, the anchor dictionary size M is fixed. Future research could explore dynamic spectral pruning to achieve even higher compression. Furthermore, extending SDAL to higher-dimensional spatio-temporal continuums would be a significant step toward modeling complex transient dynamics, such as 4D dynamic scene reconstruction or non-stationary continuous visual fields in highly degraded environments.

5.7 Limitations and Scope of Theoretical Guarantees

While Theorem 1 provides rigorous convergence guarantees under ideal conditions, practical deployments operate under finite resources. We acknowledge the following limitations. First, regarding the finite anchor budget, with $K = 4096$ anchors, the reconstruction error follows an approximation bound yielding approximately 1.5% residual error, not the zero error suggested by asymptotic theory. Second, Theorem 1 assumes compact spectral support, whereas real images contain sharp edges with infinite bandwidth. SDAL approximates such signals within the effective bandwidth by treating frequencies beyond the maximum as negligible. Third, computational scaling constraints limit K to approximately 8192 on consumer GPUs due to the closed-form solver’s complexity. Scenes requiring more than 10,000 anchors would need distributed solvers or hierarchical anchor refinement. Fourth, hyperparameter sensitivity remains a practical concern: the frequency range must be estimated from data or

Table 10 Comparing fixed features vs. gradient-updated anchors

Configuration	Cold PSNR	Final PSNR	Time (s)
Fixed (Gaussian init, no grad)	26.86	26.86	3.48
Fixed (Spider-web init, no grad)	23.77	23.77	3.58
Learnable Gaussian init + grad	24.28	27.43	9.02
Learnable Spider-web init + grad	23.77	25.68	9.08

Bold values indicate the best performance/results for the respective evaluation metrics

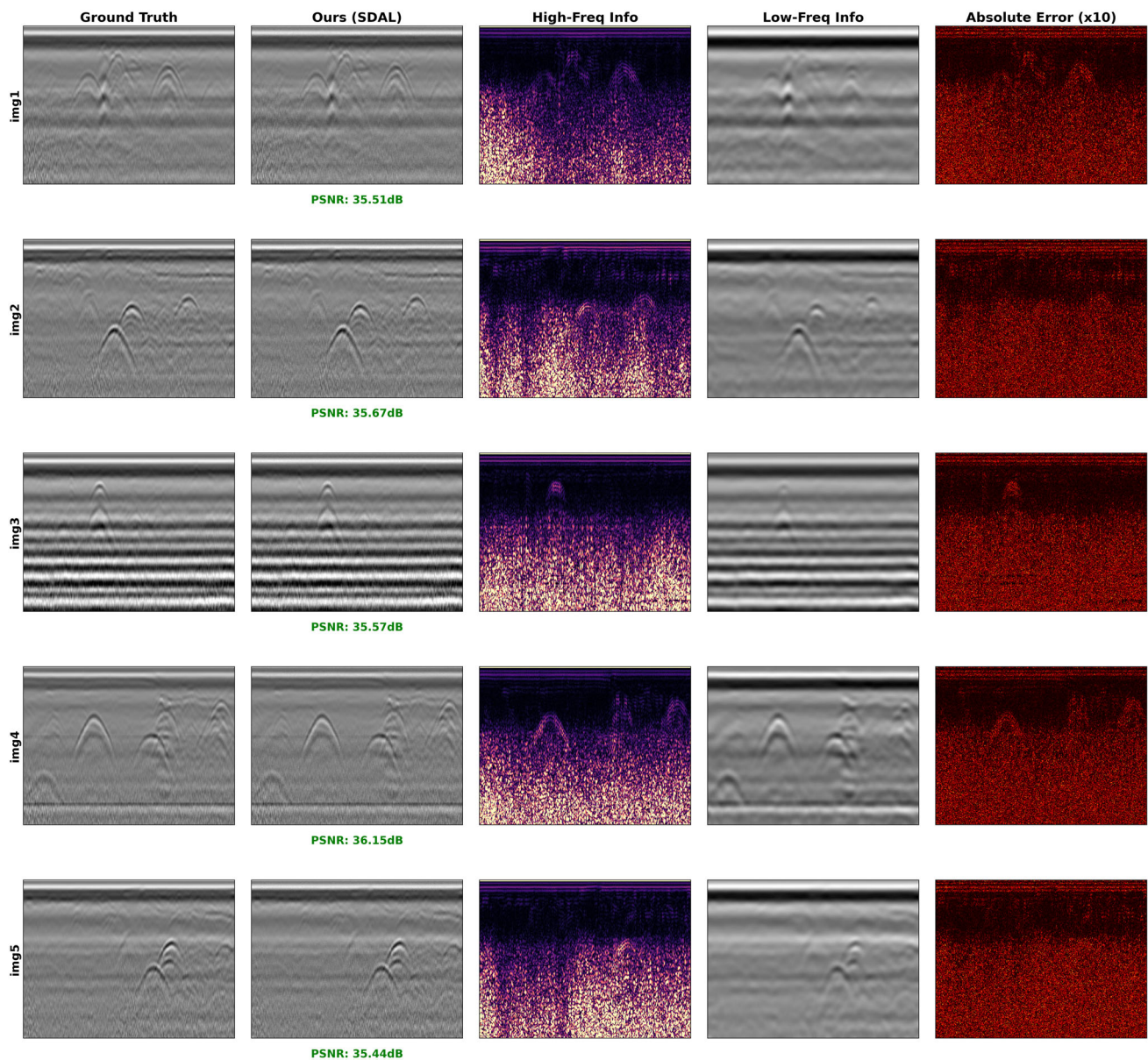


Fig. 7 Qualitative and quantitative evaluation on real-world severely degraded topologies (GPR B-scans, img1–img5). The SDAL framework achieves high-fidelity reconstruction (PSNR > 35 dB) via its analytical cold-start solver. Columns 3 and 4 visually demonstrate the framework’s unique Spectral Disentanglement capability, effectively

isolating high-frequency hyperbolic diffraction patterns from intense low-frequency background clutter. The residual maps (Absolute Error $\times 10$) highlight the precision of our method in resolving sharp oscillatory visual features with minimal error in real-world extremes

set conservatively, as underestimating the maximum frequency causes high-frequency loss while overestimating wastes anchor capacity. Finally, regarding the scope of our scale-robustness claims, these are validated empirically on Urban100 and theoretically under limiting conditions. We do not claim strict scale-invariance under all finite-sample, noisy, non-band-limited scenarios. Rather, SDAL exhibits improved scale robustness compared to gradient-based methods.

6 Conclusion

In this work, we presented the SDAL framework, introducing a spectral-domain parameterization approach for continuous neural fields. By transitioning from localized spatial-domain gradient descent to explicit spectral-domain parameterization, SDAL addresses the spectral bias inherent to standard coordinate-based visual networks.

We established that our Multi-Ring Spectral Constraint—conceptually defined as the Spider-Web Topology—acts as a geometrically-aware, adaptive filter bank. This structural regularization mathematically enables improved scale robustness and significantly reduced aliasing across diverse topological domains. Furthermore, our proposed Alternating Optimization Strategy addresses the slow-convergence bottleneck that has limited implicit neural representations. By decoupling the highly non-convex spatial learning problem and resolving the linear projection weights via a closed-form analytical solver, our model achieves a zero-iteration cold-start capability. It achieves high visual fidelity at network initialization, while simultaneously reducing the trainable parameter space by over an order of magnitude compared to conventional SIREN architectures.

Rigorous empirical evaluations across highly oscillatory continuous spatial fields, real-world severely degraded topologies (GPR), and complex natural visual domains corroborated our theoretical assertions. Notably, on extreme cross-domain GPR benchmarks (Dou et al., 2016) characterized by heterogeneous continuous environments and structural targets hidden at diverse scales (0.1m–0.7m), SDAL demonstrated a superior capacity to disentangle localized high-frequency hyperbolic signatures from intense low-frequency background clutter and severe signal degradation, achieving a high-fidelity reconstruction (PSNR > 35 dB) without the need for iterative fine-tuning. The framework consistently demonstrated optimal spectral disentanglement—simultaneously resolving rigid periodic boundaries and chaotic stochastic textures without succumbing to the high-frequency speckle oscillations or topological blurring that critically compromise baseline models.

Ultimately, this research elevates implicit neural representations from mere spatial interpolators to robust, geometrically regularized solvers. By bridging the gap between deep representation learning and rigorous signal processing in continuous visual sensing, the SDAL framework lays a highly scalable foundation for next-generation AI-driven spatial representation. Looking forward, we envision extending this paradigm to higher-dimensional spatio-temporal continuums and non-stationary dynamic visual fields. We aim to explore the capture of complex transient dynamics in high-dimensional manifolds, further consolidating SDAL as a computationally efficient and mathematically interpretable engine for large-scale computer vision and continuous spatial learning.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under Grant 42104078 and Grant 62421002, and in part by the National Key Research and Development Program of China under Grant 2023YFA1011704 and Grant 2021YFB0300101.

Data Availability The datasets analyzed and generated during the current study are detailed as follows: (1) The natural image benchmarks (Kodak24) are publicly available in the Kodak Lossless True Color Image Suite repository (Company, 2026). (2) The Urban100 dataset for high-resolution urban image benchmarking is publicly available in the repository of Huang et al. (2015). (3) The real-world Ground Penetrating Radar (GPR) subsurface B-scan dataset is derived from the published work of Dou et al. (2016) and is available from their original repository.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aghaee, A., & Khan, M. O. (2024). Performance of Fourier-based activation function in physics-informed neural networks for patient-specific cardiovascular flows. *Computer Methods and Programs in Biomedicine*, 247, 108081.
- Ahn, D., Chandran, S., Leibovici, D., Kovachki, N., Papalexakis, E.E., & Kossai, J. (2025). Lightweight fourier neural operator for time-dependent partial differential equations. In *Machine Learning and the Physical Sciences Workshop*, NeurIPS 2025.
- Beurling, A., Carleson, L.: *The Collected Works of Arne Beurling, Volume 2: Harmonic Analysis. Contemporary Mathematicians*. Birkhäuser, Boston (1989)
- Bizzi, A., Moreira, L., Marques, M., Mendonça, L., Oliveira, C., Balestro, V., Santos Fernandez, L., Yukimura, D., Petrov, P., Pereira, J., et al. (2026). Neuro-spectral architectures for causal physics-informed networks. *Advances in Neural Information Processing Systems*, 38, 54419–54449.
- Cai, Z., Zhu, H., Shen, Q., Wang, X., & Cao, X. (2024). Batch normalization alleviates the spectral bias in coordinate networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 25160–25171).
- Candes, E., & Recht, B. (2012). Exact matrix completion via convex optimization. *Communications of the ACM*, 55(6), 111–119.
- Chai, X., Cao, W., Li, J., Long, H., & Sun, X. (2024). Overcoming the spectral bias problem of physics-informed neural networks in solving the frequency-domain acoustic wave equation. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–20.
- Company, E.K. (2026). Kodak Lossless True Color Image Suite. <http://r0k.us/graphics/kodak/>. Accessed: 2026–03–27
- Dou, Q., Wei, L., Magee, D. R., & Cohn, A. G. (2016). Real-time hyperbola recognition and fitting in GPR data. *IEEE Transactions on Geoscience and Remote Sensing*, 55(1), 51–62.
- Huang, J.-B., Singh, A., & Ahuja, N. (2015). Single image super-resolution from transformed self-exemplars. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5197–5206).
- Jiao, Y., Liu, Y., Yang, J. Z., & Yuan, C. (2024). A stabilized physics informed neural networks method for wave equations. *arXiv preprint arXiv:2403.19090*.
- Kim, M., Jun-Seong, K., Yun, S.-Y., & Kim, J.-H. (2024). Synergistic integration of coordinate network and tensorial feature for improving neural radiance fields from sparse inputs. In *Forty-first International Conference on Machine Learning*.
- Kania, A., Mihajlovic, M., Prokudin, S., Tabor, J., & Spurek, P. (2025). Fresh: Frequency shifting for accelerated neural representation learning. In *International Conference on Learning Representations* (Vol. 2025, pp. 76383–76406).

- Kim, N., & Fridovich-Keil, S. (2026). Grids often outperform implicit neural representation at compressing dense signals. *Advances in Neural Information Processing Systems*, 38, 51926–51980.
- Li, D., Pan, L., Yu, Z., & Yuan, L. (2026). A continuous implicit neural representation framework with gradient regularization for sea surface height reconstruction from satellite altimetry. *EGUsphere*, 2026, 1–21.
- Liu, Z., Zhu, H., Zhang, Q., Fu, J., Deng, W., Ma, Z., Guo, Y., & Cao, X. (2024). Finer: Flexible spectral-bias tuning in implicit neural representation by variable-periodic activation functions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2713–2722).
- Liu, X., Gao, S., Zeng, B., Zhang, L., Wang, T., Liu, J., & Zhang, B. (2025). Implicit diffusion models for continuous super-resolution. *International Journal of Computer Vision*, 133(9), 6535–6557.
- Long, D., Xu, Z., Yuan, Q., Yang, Y., & Zhe, S. (2024). Invertible fourier neural operators for tackling both forward and inverse problems. *arXiv preprint arXiv:2402.11722*.
- Mello Rella, E., Chhatkuli, A., Konukoglu, E., & Van Gool, L. (2025). Neural vector fields for implicit surface representation and inference. *International Journal of Computer Vision*, 133(4), 1855–1878.
- Mohseni, P., & Duffield, N. (2026). Spectral convolutional conditional neural processes. *Advances in Neural Information Processing Systems*, 38, 40416–40446.
- Qi, S., & Sarris, C. D. (2024). Hybrid physics-informed neural network for the wave equation with unconditionally stable time-stepping. *IEEE Antennas and Wireless Propagation Letters*, 23(4), 1356–1360.
- Qi, X., Zhang, D., Wang, F., Bi, W., & Lu, M. (2025). NDAWL-PINN: A new non-dimensionalization and multi-task learning approach for efficient training of physics-informed neural networks to solve the shallow water equations. *Engineering Applications of Computational Fluid Mechanics*, 19(1), 2535015.
- Qi, Z., Peng, Y., & Earls, C. (2026). Fourier neural operators for time-periodic quantum systems: Learning Floquet hamiltonians, observable dynamics, and operator growth. *PRX Quantum*, 7(2), 020340.
- Rahimi, A., & Recht, B. (2007). Random features for large-scale kernel machines. *Advances in Neural Information Processing Systems*, 20.
- Saragadam, V., LeJeune, D., Tan, J., Balakrishnan, G., Veeraraghavan, A., & Baraniuk, R. G. (2023). Wire: Wavelet implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 18507–18516).
- Schoder, S. (2025). Physics-informed neural networks for modal wave field predictions in 3D room acoustics. *Applied Sciences*, 15(2), 939.
- Shi, K., Zhou, X., & Gu, S. (2024). Improved implicit neural representation with fourier reparameterized training. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 25985–25994).
- Staub, D. (2024). *Solving the elastic wave equation with physics-informed neural networks: A robust and critical assessment*. Switzerland: ETH Zurich. <https://doi.org/10.3929/ethz-b-000668359>
- Tie, M., Wei, J., Wu, K., Wang, Z., Yuan, S., Zhang, K., Jia, J., Zhao, J., Gan, Z., & Ding, W. (2024). O 2 v-mapping: Online open-vocabulary mapping with neural implicit representation. *European Conference on Computer Vision* (pp. 318–333). Cham: Springer.
- Xu, S., Wang, Y., Luo, X., Yoon, B.-J., & Qian, X. (2025). Scale-invariant implicit neural representations for object counting. In *Proceedings of the Computer Vision and Pattern Recognition Conference* (pp. 2308–2318).
- Ye, H., Xu, C., Zhou, Y., & Meng, X. (2025). Spatial-temporal, parallel, physics-informed neural networks for solving forward and inverse PDE problems via overlapping domain decomposition. *Aerospace Research Communications*, 3, 14842.
- Zhang, Z., Yang, R., Suo, J., Cheng, Y., & Dai, Q. (2025). Lightweight high-speed photography built on coded exposure and implicit neural representation of videos. *International Journal of Computer Vision*, 133(3), 991–1011.
- Zhao, H., Gao, Z., Wang, Y., Xiong, R., & Zhang, Y. (2025). Adaptive wavelet-positional encoding for high-frequency information learning in implicit neural representation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 39, pp. 10430–10438).
- Zheng, X., Xia, H., Wang, H., Zhang, C., Duan, J., Liu, J., & Tang, S. (2025). A Fourier neural operator-enhanced parabolic equation framework for highly efficient underwater acoustic field prediction. *Frontiers in Marine Science*, 12, 1692899.
- Zhu, H., Xie, S., Liu, Z., Liu, F., Zhang, Q., Zhou, Y., Lin, Y., Ma, Z., & Cao, X. (2024). Disorder-invariant implicit neural representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(8), 5463–5478.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.