Towards a Sampling Theory for Implicit Neural Representations

Mahrokh Najaf   Gregory Ongie
Department of Mathematical and Statistical Sciences
Marquette University
Milwaukee, WI

Abstract—**Implicit neural representations (INRs) have emerged as a powerful tool for solving inverse problems in computer vision and computational imaging. INRs represent images as continuous domain functions realized by a neural network taking spatial coordinates as inputs. However, unlike traditional pixel representations, little is known about the sample complexity of estimating images using INRs in the context of linear inverse problems. Towards this end, we study the sampling requirements for recovery of a continuous domain image from its low-pass Fourier coefficients by fitting a single hidden-layer INR with ReLU activation and a Fourier features layer using a generalized form of weight decay regularization. Our key insight is to relate minimizers of this non-convex parameter space optimization problem to minimizers of a convex penalty defined over a space of measures. We identify a sufficient number of samples for which an image realized by a width-1 INR is exactly recoverable by solving the INR training problem, and give a conjecture for the general width-W case. To validate our theory, we empirically assess the probability of achieving exact recovery of images realized by low-width single hidden-layer INRs, and illustrate the performance of INR on super-resolution recovery of more realistic continuous domain phantom images.**

I. INTRODUCTION

Many inverse problems in signal and image processing are naturally posed in continuous domain. For example, in magnetic resonance imaging, measurements are modeled as samples of the Fourier transform of a function $f(x)$ defined over continuous spatial coordinates $x \in \mathbb{R}^3$ [1]. Often these continuous domain inverse problems are approximated using discrete image or signal representations, along with an appropriate discretization of the forward model, such as the discrete Fourier transform in place of the continuous Fourier transform. However, this can lead to undesirable discretization artifacts, such as Gibb's ringing [2], [3], [4], [5]. Such artifacts may be reduced by increasing the resolution of the discretization grid, but at the cost of additional computational complexity and potentially making the problem ill-posed without additional regularization.

Recently, continuous domain image representations using neural networks have gained traction in a variety of inverse problems arising in computer vision and computational imaging (see, e.g., the recent review [6]). These so-called *implicit neural representations* (INRs), also variously known as *coordinate-based neural networks* or *neural radiance fields*, parameterize a signal or image as a neural network defined over continuous spatial coordinates [7], [8], [9]. The main

benefits of INRs over discrete representations are that (1) they offer a compact (non-linear) representation of signals/images with far fewer parameters than a discrete representation, (2) they allow for more accurate continuous domain forward modeling, which can reduce the impact of discretization artifacts, and (3) standard neural network training can be used to fit an INR to measurements, sidestepping the need for custom iterative solvers.

However, the fundamental limits of image/signal estimation using INRs have not been explored in detail. One basic unanswered question is: How many linear measurements are necessary and sufficient for accurate reconstruction of a signal/image by fitting an INR? Additionally, how does the architecture of the INR influence the recovered signal/image? As a step in the direction of answering these questions, we consider the recovery of a signal/image from its low-pass Fourier coefficients using INRs. In particular, we focus on a popular INR architecture originally introduced in [9] that combines a fully-connected neural network with a Fourier features layer.

Recent work has shown that training a single hidden-layer ReLU network with ℓ^2-regularization on its parameters (also known as "weight decay" regularization) imparts a sparsity effect, such that the network learned has low effective width [10], [11], [12], [13]. Bringing this perspective to INRs, we study the conditions under which a signal/image realizable by a single-hidden layer, low-width INR is exactly recoverable from low-pass Fourier samples by minimizing a generalized weight decay objective. Our main theoretical insight is to relate minimizers of this non-convex parameter space optimization problem to the minimizers of a convex optimization problem posed over a space of measures. Using this framework, we identify a sufficient number of samples such that an image realizable by a width-1 INR is the unique minimizer of the INR training problem, and state a conjecture for widths greater than one. We validate this conjecture with numerical simulations. Finally, we illustrate the practical performance of INRs on super-resolution recovery of continuous domain phantom images, highlighting the impact of parameter-space regularization and network depth on recovery performance.

II. PROBLEM FORMULATION

We consider the recovery of a d-dimensional real-valued continuous domain signal/image $f : [0, 1]^d \to \mathbb{R}$ from a

sampling of its Fourier coefficients $\hat{f}$ defined by

$$\hat{f}[\boldsymbol{k}] = \int_{[0,1]^d} f(\boldsymbol{x}) e^{-2\pi i \boldsymbol{k}^\top \boldsymbol{x}} d\boldsymbol{x},$$

where $\boldsymbol{k} \in \mathbb{Z}^d$ is any d-tuple of integer spatial frequencies. In particular, suppose we have access to all low-pass Fourier coefficients belonging to a uniformly sampled square region in frequency domain $\Omega = \{\boldsymbol{k} \in \mathbb{Z}^d : \|\boldsymbol{k}\|_\infty \leq K\}$ where K is a positive integer. Let $\mathcal{F}_\Omega f = (\hat{f}[\boldsymbol{k}])_{\boldsymbol{k} \in \Omega}$ denote the vector of Fourier coefficients of f restricted to Ω. We study the inverse problem of recovering the function f given $\boldsymbol{y} = \mathcal{F}_\Omega f$.

Clearly, this is an ill-posed inverse problem if no further assumptions are made on f. The INR approach attempts to resolve this ill-posedness by assuming f is well-approximated by a function realized by a "simple" neural network f_θ with trainable parameters θ. Under this assumption, one can attempt to recover f from measurements $\boldsymbol{y}$ by solving the (non-linear) least squares problem

$$\min_\theta \tfrac{1}{2} \|\mathcal{F}_\Omega f_\theta - \boldsymbol{y}\|_2^2. \tag{1}$$

However, this problem may still be ill-posed (i.e., (1) may have multiple solutions) unless the number of trainable parameters used to define the INR f_θ is heavily constrained. Yet, useful solutions of (1) may be found by relying on implicit regularization induced by practical gradient-based algorithms [14].

As an alternative to using an overly constrained network architecture or relying on implicit regularization to resolve ill-posedness, we propose incorporating an explicit parameter space regularizer $R(\theta)$ to the least squares objective:

$$\min_\theta \tfrac{1}{2} \|\mathcal{F}_\Omega f_\theta - \boldsymbol{y}\|_2^2 + \lambda R(\theta), \tag{2}$$

where $\lambda > 0$ is a regularization parameter. In particular, we focus on a class of regularizers $R(\theta)$ that generalize weight decay regularization, i.e., the squared ℓ^2-norm of all trainable parameters; this class is described in more detail below in Section II-B.

To tightly enforce data consistency constraints, the regularization parameter λ in (2) should be very small. As a model for this situation, we will focus on the equality constrained problem

$$\min_\theta R(\theta) \quad s.t. \quad \mathcal{F}_\Omega f_\theta = \boldsymbol{y}, \tag{3}$$

which can be thought of as the limiting case of (2) as $\lambda \to 0$. Our main goal is to characterize global minimizers of (3). In particular, we are interested in its *function space minimizers*, i.e., functions f_{θ^*} where θ^* is a global minimizer of (3). In particular, given $\boldsymbol{y} = \mathcal{F}_\Omega f$ where f is realizable as an INR, we ask: when is f the unique function space minimizer of (3)? Next, we describe the architectural assumptions we put on the INR f_θ and the associated regularizer $R(\theta)$.

A. INR Architecture

We focus on an INR architecture first proposed in [9] that combines a trainable fully connected network with a Fourier features layer. For simplicity, we primarily consider a
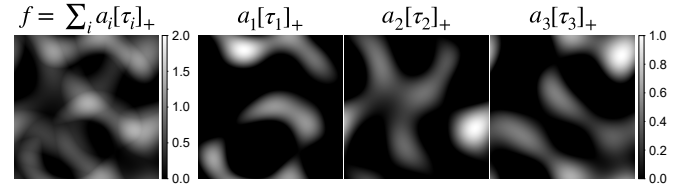

$f = \sum_i a_i [\tau_i]_+ \qquad a_1[\tau_1]_+ \qquad a_2[\tau_2]_+ \qquad a_3[\tau_3]_+$

Fig. 1. The single-hidden layer INR architecture considered in this work represents an image f as a linear combination of rectified trigonometric polynomials $[\tau_i]_+ = [\boldsymbol{w}_i^\top \boldsymbol{\gamma}(\boldsymbol{x})]_+$. We study the sample complexity of estimating an image of this type from its low-pass Fourier coefficients.

"shallow" INR architecture, where the fully connected network has a single-hidden layer with ReLU activation.

Specifically, consider a single-hidden layer ReLU network $g_\theta : \mathbb{R}^D \to \mathbb{R}$ of width-W as given by

$$g_\theta(\boldsymbol{v}) = \sum_{i=1}^W a_i [\boldsymbol{w}_i^\top \boldsymbol{v}]_+, \tag{4}$$

where $[\cdot]_+$ denotes the ReLU activation function, and $\theta = \{(a_i, \boldsymbol{w}_i)\}_{i=1}^W$ denotes the vector of all trainable parameters, which belongs to the space $\Theta_W \simeq \mathbb{R}^{(1+d)DW}$. By choosing $D = d$, this architecture could be used as an INR directly. However, as shown in [9], it is challenging to recover high-frequency content in images using such an architecture.

To enhance recovery of high-frequency content, [9] proposed adding an initial Fourier features layer to the architecture. In this work, we focus on a Fourier features layer $\boldsymbol{\gamma} : \mathbb{R}^d \to \mathbb{R}^D$ of the form

$$\boldsymbol{\gamma}(\boldsymbol{x}) = [1, \sqrt{2}\cos(2\pi \boldsymbol{k}_1^\top \boldsymbol{x}), \cdots, \sqrt{2}\cos(2\pi \boldsymbol{k}_p^\top \boldsymbol{x}),$$
$$\sqrt{2}\sin(2\pi \boldsymbol{k}_1^\top \boldsymbol{x}), \cdots, \sqrt{2}\sin(2\pi \boldsymbol{k}_p^\top \boldsymbol{x})]^\top, \tag{5}$$

where $\Omega_0 = \{\boldsymbol{k}_1, ..., \boldsymbol{k}_p\} \subset \mathbb{R}^d$ is a collection of frequency vectors. For a single hidden-layer fully connected network, this gives the INR architecture:

$$f_\theta(\boldsymbol{x}) = g_\theta(\boldsymbol{\gamma}(\boldsymbol{x})) = \sum_{i=1}^W a_i [\boldsymbol{w}_i^\top \boldsymbol{\gamma}(\boldsymbol{x})]_+. \tag{6}$$

To better understand the impact of the Fourier features layer, for any weight vector $\boldsymbol{w} \in \mathbb{R}^D$, let us write $\boldsymbol{w} = [w^{(0)}, \boldsymbol{w}^{(1)}, \boldsymbol{w}^{(2)}]$ where $w^{(0)} \in \mathbb{R}$ and $\boldsymbol{w}^{(1)}, \boldsymbol{w}^{(2)} \in \mathbb{R}^p$. Then defining $\tau(\boldsymbol{x}) := \boldsymbol{w}^\top \boldsymbol{\gamma}(\boldsymbol{x})$, we have

$$\tau(\boldsymbol{x}) = w^{(0)} + \sqrt{2} \sum_{j=1}^p \left(w_j^{(1)} \cos(2\pi \boldsymbol{k}_j^\top \boldsymbol{x}) + w_j^{(2)} \sin(2\pi \boldsymbol{k}_j^\top \boldsymbol{x}) \right).$$

Note that τ is a *trigonometric polynomial*, i.e., τ is a bandlimited function with frequency support contained in the discrete set Ω_0. Therefore, a single-hidden layer INR with Fourier features layer is a weighted linear combination of functions of the form $\boldsymbol{x} \mapsto [\tau(\boldsymbol{x})]_+$. We call such a function a *rectified trigonometric polynomial*; see Figure 1 for an illustration.

Originally, [9] proposed defining the frequency set Ω_0 by sampling normally distributed random vectors $\boldsymbol{k} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$, where the variance σ is treated as a tunable parameter. To

simplify the mathematical theory, in this work, we focus on the case where Ω_0 consists of uniform sampled integer frequencies with maximum frequency K_0, i.e., $\Omega_0 = \{\boldsymbol{k} \in \mathbb{Z}^d : \|\boldsymbol{k}\|_\infty \leq K_0\}$. This implies that the INR is a periodic function on $[0,1]^d$. Using integer frequencies in place of random frequencies for an INR Fourier features layer was also investigated in [15].

B. Generalized Weight Decay Regularization

We focus on a family of parameter space regularizers $R(\theta)$ that strictly generalize "weight decay" regularization [16], i.e., the squared ℓ^2-norm of all trainable parameters. Consider any non-negative weighting function $\eta : \mathbb{R}^D \to \mathbb{R}$ defined over inner-layer weight vectors $\boldsymbol{w} \in \mathbb{R}^D$. We say η is an *admissible* weighting function if it is continuous, positively homogeneous (i.e., $\eta(\alpha\boldsymbol{w}) = \alpha\eta(\boldsymbol{w})$ for all $\alpha \geq 0$), and $\eta(\boldsymbol{w}) > 0$ whenever $[\boldsymbol{w}^\top \boldsymbol{\gamma}(\cdot)]_+ \neq 0$. Given any admissible weight function, for any parameter vector $\theta = \{(a_i, \boldsymbol{w}_i)\}_{i=1}^W$ we define the regularizer $R(\theta)$ by

$$R(\theta) = \frac{1}{2}\sum_{i=1}^W \left(|a_i|^2 + \eta(\boldsymbol{w}_i)^2\right). \tag{7}$$

For example, if $\eta(\boldsymbol{w}) = \|\boldsymbol{w}\|_2$ we recover standard weight decay regularization. For the sampling theory given in Section III-C, we focus on the alternative weighting function

$$\eta(\boldsymbol{w}) = \|\mathcal{F}_\Omega[\boldsymbol{w}^\top\boldsymbol{\gamma}(\cdot)]_+\|_2, \tag{8}$$

and call the resulting $R(\theta)$ *modified weight decay regularization*. Assuming Ω is a sufficiently large set of frequencies, by Parseval's Theorem we have $\|\mathcal{F}_\Omega[\boldsymbol{w}^\top\boldsymbol{\gamma}(\cdot)]_+\|_2 \approx \|[\boldsymbol{w}^\top\boldsymbol{\gamma}(\cdot)]_+\|_{L^2([0,1]^d)}$. This shows that modified weight decay regularization can be seen as approximately penalizing the L^2-norm contribution made by each unit $[\boldsymbol{w}^\top\boldsymbol{\gamma}(\cdot)]_+$ to the overall image.

III. THEORY

Our main theoretical insight is to show that minimizers of the non-convex INR training problem (3) coincide with minimizers of a *convex* optimization problem defined over an infinite-dimensional space of measures. We use this equivalence to prove our main theorem, which identifies a sufficient number of Fourier samples needed to recover a function realizable as a width-1 INR (i.e., a truncated trigonometric polynomial). Proofs of all results in this section are omitted for space reasons.

A. Generalized Weight Decay Induces Unit Sparsity

Let $\Theta_W^* \subset \Theta_W$ denote the subset of parameter vectors $\theta = \{(a_i, \boldsymbol{w}_i)\}_{i=1}^W$ that satisfy the normalization constraint $\|\boldsymbol{w}\|_2 = 1$ for all $i \in [W]$. Then we have the following equivalence:

Lemma 1. *The function space minimizers of* (3) *coincide with the function space minimizers of*

$$\min_{\theta \in \Theta_W^*} \sum_{i=1}^W |a_i|\eta(\boldsymbol{w}) \ s.t. \ \mathcal{F}_\Omega f_\theta = \boldsymbol{y}. \tag{9}$$

The proof follows by an argument similar to the proof of [10, Lemma 1], and relies on the positive homogeneity of

the ReLU activation and of the weighting function η. This result shows that by minimizing weight decay we implicitly minimize a weighted ℓ^1-norm on the outer-layer weights a_i of the network subject to a normalization constraint on the inner-layer weight vectors $\boldsymbol{w}_i$. Due to the sparsity-promoting property of the ℓ^1-norm, intuitively this result shows that by minimizing $R(\theta)$ we should promote functions that are realizable by a "sparse" INR, i.e., one with few active ReLU units.

B. Infinite-width Convexification

One technical hurdle to characterizing minimizers of (9) is that set of parameters $\theta \in \Theta_W^*$ satisfying the constraint $\mathcal{F}_\Omega f_\theta = \boldsymbol{y}$ is non-convex. To circumvent this issue, we show how to reformulate (9) as a convex optimization problem posed over a space of measures that represents the limit as the hidden-layer width $W \to \infty$. We note that similar convex formulations for single hidden-layer neural networks have been explored in [17], [18], [19].

First, observe that for any $\theta = (\{a_i, \boldsymbol{w}_i\}_{i=1}^W \in \Theta_W^*$, by linearity we have

$$\mathcal{F}_\Omega f_\theta = \sum_{i=1}^W a_i \mathcal{F}_\Omega[\boldsymbol{w}_i^\top\boldsymbol{\gamma}(\cdot)]_+.$$

We now show the sum on the right can be expressed as the integration of a function against a signed measure. In particular, define $\mathcal{U}_\eta = \{\boldsymbol{w} \in \mathbb{R}^D : \|\boldsymbol{w}\|_2 = 1, \eta(\boldsymbol{w}) > 0\}$. For any finite signed measure μ defined over $\mathcal{U}_\eta$, define the linear operator

$$\mathcal{K}_\Omega\mu := \int_{\mathcal{U}_\eta} \mathcal{F}_\Omega[\boldsymbol{w}^\top\boldsymbol{\gamma}(\cdot)]_+ d\mu(\boldsymbol{w})$$

Then for the sparse measure $\mu^* = \sum_{i=1}^W a_i\delta_{\boldsymbol{w}_i}$ where $\delta_{\boldsymbol{w}}$ denotes a Dirac delta measure centered at $\boldsymbol{w}$, we have $\mathcal{F}_\Omega f_\theta = \mathcal{K}_\Omega\mu^*$ and so the constraint $\mathcal{F}_\Omega f_\theta = \boldsymbol{y}$ in (9) is equivalent to $\mathcal{K}_\Omega\mu^* = \boldsymbol{y}$.

Next, define the weighted total variation norm $\|\mu\|_{TV,\eta}$ for any signed measure μ over $\mathcal{U}_\eta$ by

$$\|\mu\|_{TV,\eta} = \int_{\mathcal{U}_\eta} \eta(\boldsymbol{w})d|\mu|(\boldsymbol{w}),$$

where $|\mu|$ is the total variation measure associated with μ. Then we have $\|\mu^*\|_{TV,\eta} = \sum_{i=1}^K |a_i|\eta(\boldsymbol{w}_i)$ which coincides with the objective in (9).

Therefore, letting $\mathcal{M}_W(\mathcal{U}_\eta)$ denote the set of measures defined over $\mathcal{U}_\eta$ expressible as a weighted linear combination of at most W Diracs, we have shown that (9) is equivalent to the optimization problem

$$P_W^* = \min_{\mu \in \mathcal{M}_W(\mathcal{U}_\eta)} \|\mu\|_{TV,\eta} \ s.t. \ \mathcal{K}_\Omega\mu = \boldsymbol{y}. \tag{10}$$

In particular, if the measure $\mu^* = \sum_{i=1}^W a_i\delta_{\boldsymbol{w}_i}$ is a minimizer of (10), then the parameter vector $\theta^* = \{(a_i, \boldsymbol{w}_i)\}_{i=1}^W \in \Theta_W^*$ is a minimizer of (9), and hence by Lemma 1, the INR f_{θ^*} is also a function space minimizer of the original INR training

problem (3). This shows that we can describe all function space minimizers (3) by identifying the minimizers of (10).

However, analyzing (10) directly is challenging since domain $\mathcal{M}_W(\mathcal{U}_\eta)$ is not a vector space. To circumvent this issue, we pass to the larger space $\mathcal{M}(\mathcal{U}_\eta)$ of signed measures defined over $\mathcal{U}_\eta$ with finite total variation norm, and consider the optimization problem

$$P^* = \min_{\mu \in \mathcal{M}(\mathcal{U}_\eta)} \|\mu\|_{TV,\eta} \quad s.t. \quad \mathcal{K}_\Omega \mu = \boldsymbol{y}. \tag{11}$$

This can be thought of as the "infinite-width" analogue of (9). Note (11) is now a convex optimization problem, though one posed over an infinite dimensional space of measures.

Since $\mathcal{M}_W(\mathcal{U}_\eta) \subset \mathcal{M}(\mathcal{U}_\eta)$ we have $P^* \leq P_W^*$. Additionally, if a minimizer μ^* to (11) is a finite linear combination of at most W Diracs, i.e., $\mu^* \in \mathcal{M}_W(\mathcal{U}_\eta)$, then μ^* must also be a minimizer of (10), and so $P^* = P_W^*$. This is the key property we exploit in proving the sampling guarantees given below.

C. Provable Recovery of Width-1 INRs

Now we give our main theorem, which identifies a sufficient number of samples for which an image realizable as a width-1 INR is exactly recoverable as the unique global minimizer of (3). For technical reasons, here we focus on the modified weight decay regularizer $R(\theta)$ defined using the specific weighting function $\eta(\boldsymbol{w})$ given in (8). Also, we assume the Fourier feature frequency set $\Omega_0 = \{\boldsymbol{k} \in \mathbb{Z}^d : \|\boldsymbol{k}\|_\infty \leq K_0\}$, and for any positive integer n, we let $n\Omega_0$ denote the dilated frequency set $\{\boldsymbol{k} \in \mathbb{Z}^d : \|\boldsymbol{k}\|_\infty \leq nK_0\}$. Then we have the following result:

Theorem 1. *Let* $f(\boldsymbol{x}) = a[\tau]_+$ *where* τ *is a real-valued trigonometric polynomial with frequency support in* Ω_0. *Suppose* $\Omega \supseteq 3\Omega_0$, *and* $\boldsymbol{y} = \mathcal{F}_\Omega f$. *Then* f *is the unique function space minimizer of the INR training problem* (3) *using modified weight decay regularization.*

The main idea of the proof is to pass to the convex dual of (11), which reduces to a semi-infinite program, i.e., a finite dimensional linear program with infinitely many constraints. Similar to previous works that have focused on super-resolution of sparse spikes [20], [21], [22], we prove unique recovery is equivalent to the construction of a particular dual optimal solution known as a *dual certificate*. In the case of a width-1 INR, we explicitly construct the dual certificate.

More generally, we conjecture that Theorem 1 can be extended to the case of standard weight decay regularization (i.e., weighting function $\eta(\boldsymbol{w}) = \|\boldsymbol{w}\|_2$). This is verified empirically in the next section. Additionally, when f is a width W INR, we conjecture that sampling over the frequency set $3W\Omega_0$ is sufficient to ensure f is the unique minimizer of (11), both with standard and modified weight decay regularization.

IV. EXPERIMENTS

A. Exact Recovery Experiments

To validate the proved and conjectured sampling guarantees, we generate random continuous domain images realizable as a single hidden-layer "teacher network" INR of varying widths W; Figure 1 shows an example for $W = 3$. We investigate the maximum sampling frequency K needed to recover the image by solving the optimization problem (3) using a "student network" INR with the same architecture but larger width. As an attempt to solve the equality constrained problem (3) we use the Augmented Lagrangian (AL) method (see, e.g., [23, Ch 7]). This results in a sequence of "outer loop" INR training problems, each of which approximately optimized with $5,000$ iterations of the Adam optimizer with a learning rate of 0.001, for a maximum of 60 outer loop iterations.

For both the teacher and student networks, we use a Fourier features layer with maximum frequency $K_0 = 2$. We sweep over K values from 2 to 30 in increments of 2, while varying the teacher network width, W, from 1 to 5. For each regularization approach and each (K, W) pair, we run 10 random trials. The width of student networks is 100. We say the recovery is "exact" if the image domain mean-squared-error (MSE) is less than 1×10^{-9} at any point during training.
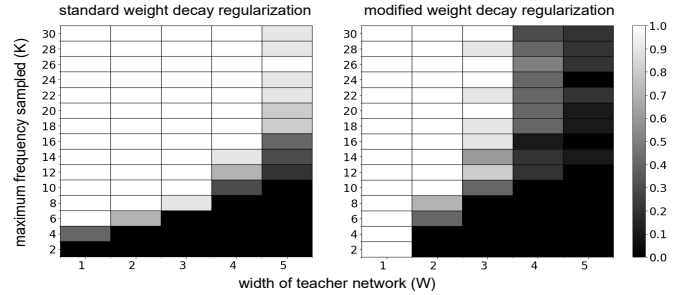

Fig. 2. Empirical probability of exact recovery of images representable by a single hidden-layer INR of width-W from its low-pass Fourier coefficients with maximum sampling frequency K by solving (3) using standard and modified weight decay regularization.

Figure 2 shows the results of these experiments as probability tables. In the table corresponding to standard weight decay regularization, we see that the conjectured linear trend between the teacher network width W and the minimal sampling cutoff K needed to achieve exact recovery holds with high probability. For modified weight decay regularization, although the linear trend is less obvious, exact recovery is achieved in all random trials when $W = 1$ and remains attainable in many cases when $W > 1$. For both regularization methods, the parameter space optimization problem is non-convex, and so convergence of the AL method to a global optimum is not guaranteed. Therefore, it is unclear if the relatively worse empirical performance of modified weight decay regularization at achieving exact recovery is due to a difficulty in reaching a global minimizer with the proposed optimization method, or a more fundamental lack of identifiability. Despite this deficiency, we note there may be practical benefits to using modified weight decay over standard weight decay regularization in achieving *approximate* recovery, as we show in the next set of experiments with more realistic phantom images.
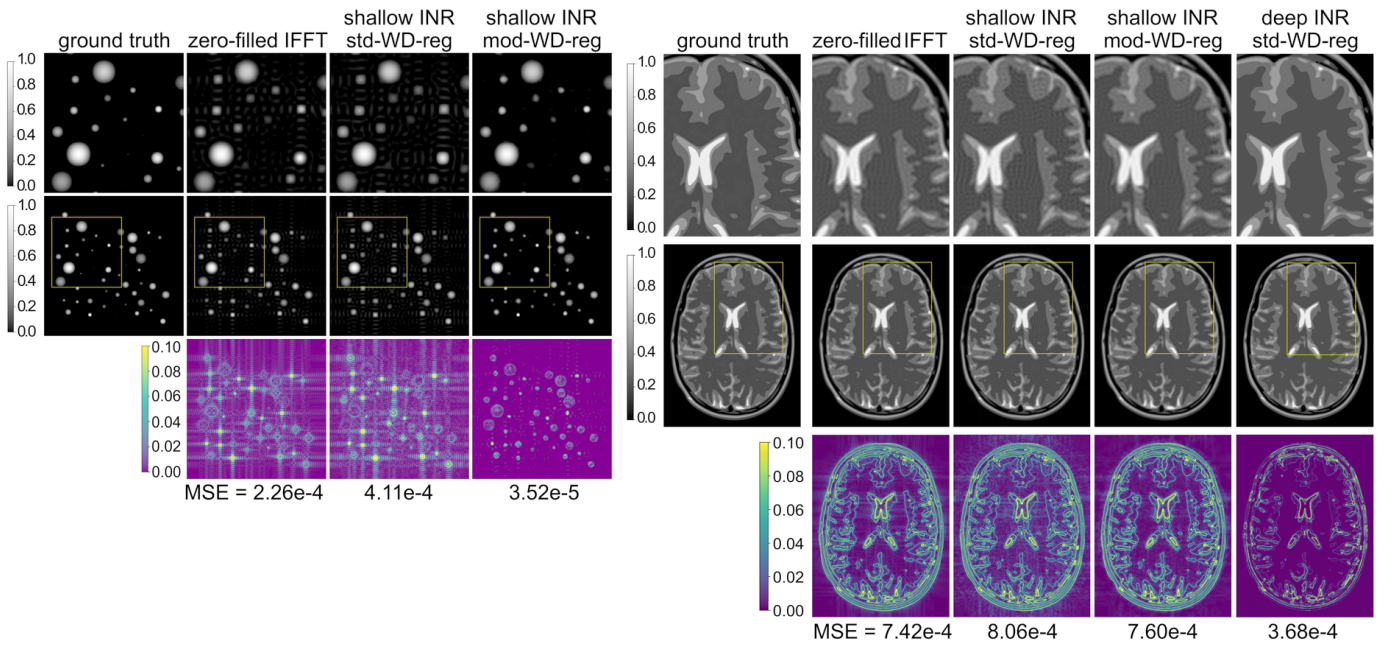
Fig. 3. **Recovery of continuous domain phantoms from low-pass Fourier samples.** Left panel: "dot phantom" image created for this study features randomly generated circle-shaped features. Right panel: piecewise constant brain MRI phantom introduced in [24]. We compare the zero-filled IFFT reconstruction with a single-hidden layer INR trained with standard weight decay regularization (shallow INR, std-WD-reg), and with the proposed modified weight decay regularization (shallow INR, mod-WD-reg). For the brain phantom, we also compare with a depth 10 INR architecture trained with standard weight decay regularization (deep INR, std-WD-reg). The bottom row shows the absolute value of the difference with the ground truth.

B. Recovery of Continuous Domain Phantom Images

In Figure 3, we illustrate the use of INRs for the super-resolution recovery of two continuous domain phantoms whose Fourier coefficients are known exactly: a "dot phantom" image created for this study that features randomly generated circle-shaped features, and a piecewise constant MRI "brain phantom" introduced in [24]. The dot phantom is continuous, piecewise smooth, and exactly realizable by a shallow INR architecture of width $W \geq 50$ and $K_0 \geq 8$. Since the brain phantom is piecewise constant, and hence discontinuous, it is not exactly realizable with the INR considered in this work. We also note both phantoms are not bandlimited, such that a simple reconstruction obtained by zero-filling in the Fourier domain and applying the IFFT results in significant ringing artifacts (see the second column in each panel of Figure 3).

For the dot phantom, the maximum sampling frequency is set to $K = 32$, and we fit a shallow INR with maximum frequency $K_0 = 10$ for the Fourier features layer and width 100. For the brain phantom, the maximum sampling frequency is set to $K = 64$, and we fit a shallow INR of width 100 with maximum frequency $K_0 = 10$ for the Fourier features layer. To illustrate the role of INR depth, in the case of the brain phantom, we also fit a deep INR with depth 10 and width 100 in each layer. We train the INR by minimizing the regularized least squares formulation (2) using the Adam optimizer for $40,000$ iterations with a learning rate of 0.001 followed by an additional $10,000$ iterations with a learning rate of 0.0001. In each case, the regularization parameter λ was selected via grid search to minimize image domain MSE.

Unlike the exact recovery experiments, for both phantoms, we find that using a shallow INR trained with modified weight decay regularization results in lower image MSE and fewer artifacts in the reconstructions compared to using a shallow INR trained with standard weight decay regularization. In particular, the MSE is a magnitude lower on the dot phantom with modified weight decay and the reconstruction has far fewer ringing artifacts. A more modest improvement is shown on the brain phantom, where training with modified weight decay yields a slight improvement in MSE, but some ringing artifacts are still visible. However, we find that using a deep INR leads to further improvements in the brain phantom's recovery, eliminating nearly all residual ringing artifacts.

V. Conclusion

In this work, we introduce a mathematical framework for studying exact recovery of continuous domain images from low-pass Fourier coefficients with INRs. We show that fitting a single-hidden layer INR architecture using a generalized weight decay regularization term is equivalent to minimizing a convex loss over a space of measures. Using this perspective, we identify a sufficient number of samples to uniquely recover an image realizable as a width-1 INR, which we conjecture generalizes larger widths, as well.

While this work focused on a specific INR architecture introduced in [9], several others have been proposed [8], [25], [26], [27]. Extending the present results to these alternative architectures is an interesting direction for future work. Also, for mathematical tractability, we focused on a single hidden-layer INR architecture, but many studies use INR architectures with two or more hidden layers. Extensions of the theory to deeper architectures will also be a focus of future work.

Authorized licensed use limited to: Marquette University. Downloaded on July 01,2025 at 14:16:32 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] J. A. Fessler, "Model-based image reconstruction for MRI," *IEEE signal processing magazine*, vol. 27, no. 4, pp. 81–89, 2010.

[2] R. Archibald and A. Gelb, "A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity," *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 305–319, 2002.

[3] J. Veraart, E. Fieremans, I. O. Jelescu, F. Knoll, and D. S. Novikov, "Gibbs ringing in diffusion MRI," *Magnetic resonance in medicine*, vol. 76, no. 1, pp. 301–314, 2016.

[4] G. Ongie and M. Jacob, "Recovery of piecewise smooth images from few fourier samples," pp. 543–547, 2015.

[5] ——, "Off-the-grid recovery of piecewise constant images from few fourier samples," *SIAM Journal on Imaging Sciences*, vol. 9, no. 3, pp. 1004–1041, 2016.

[6] Y. Xie, T. Takikawa, S. Saito, O. Litany, S. Yan, N. Khan, F. Tombari, J. Tompkin, V. Sitzmann, and S. Sridhar, "Neural fields in visual computing and beyond," in *Computer Graphics Forum*, vol. 41, no. 2. Wiley Online Library, 2022, pp. 641–676.

[7] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[8] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 7462–7473, 2020.

[9] M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 7537–7547, 2020.

[10] P. Savarese, I. Evron, D. Soudry, and N. Srebro, "How do infinite width bounded norm networks look in function space?" in *Conference on Learning Theory*. PMLR, 2019, pp. 2667–2690.

[11] G. Ongie, R. Willett, D. Soudry, and N. Srebro, "A function space view of bounded norm infinite width relu nets: The multivariate case," in *International Conference on Learning Representations*, 2019.

[12] R. Parhi and R. D. Nowak, "Banach space representer theorems for neural networks and ridge splines," *Journal of Machine Learning Research*, vol. 22, no. 43, pp. 1–40, 2021.

[13] ——, "Deep learning meets sparse regularization: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 40, no. 6, pp. 63–74, 2023.

[14] S. Gunasekar, J. Lee, D. Soudry, and N. Srebro, "Characterizing implicit bias in terms of optimization geometry," in *International Conference on Machine Learning*. PMLR, 2018, pp. 1832–1841.

[15] N. Benbarka, T. Höfer, A. Zell *et al.*, "Seeing implicit neural representations as fourier series," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 2041–2050.

[16] A. Krogh and J. Hertz, "A simple weight decay can improve generalization," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 4, 1991.

[17] Y. Bengio, N. Roux, P. Vincent, O. Delalleau, and P. Marcotte, "Convex neural networks," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 18, 2005.

[18] F. Bach, "Breaking the curse of dimensionality with convex neural networks," *Journal of Machine Learning Research*, vol. 18, no. 19, pp. 1–53, 2017.

[19] T. Ergen and M. Pilanci, "Convex geometry and duality of over-parameterized neural networks," *Journal of Machine Learning Research*, vol. 22, no. 212, pp. 1–63, 2021.

[20] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[21] C. Poon and G. Peyré, "Multidimensional sparse super-resolution," *SIAM Journal on Mathematical Analysis*, vol. 51, no. 1, pp. 1–44, 2019.

[22] A. Eftekhari, T. Bendory, and G. Tang, "Stable super-resolution of images: theoretical study," *Information and Inference: A Journal of the IMA*, vol. 10, no. 1, pp. 161–193, 2021.

[23] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer.

[24] M. Guerquin-Kern, L. Lejeune, K. P. Pruessmann, and M. Unser, "Realistic analytical phantoms for parallel magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 626–636, 2011.

[25] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM transactions on graphics (TOG)*, vol. 41, no. 4, pp. 1–15, 2022.

[26] V. Saragadam, D. LeJeune, J. Tan, G. Balakrishnan, A. Veeraraghavan, and R. G. Baraniuk, "Wire: Wavelet implicit neural representations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 18 507–18 516.

[27] J. Shenouda, Y. Zhou, and R. D. Nowak, "ReLUs are sufficient for learning implicit neural representations," *arXiv preprint arXiv:2406.02529*, 2024.