Quantum Implicit Neural Representations

Jiaming Zhao [1] **Wenbo Qiao** [1] **Peng Zhang** [2] **Hui Gao** [2]

Abstract

Implicit neural representations have emerged as a powerful paradigm to represent signals such as images and sounds. This approach aims to utilize neural networks to parameterize the implicit function of the signal. However, when representing implicit functions, traditional neural networks such as ReLU-based multilayer perceptrons face challenges in accurately modeling high-frequency components of signals. Recent research has begun to explore the use of Fourier Neural Networks (FNNs) to overcome this limitation. In this paper, we propose **Q**uantum **I**mplicit **Re**presentation **N**etwork (QIREN), a novel quantum generalization of FNNs. Furthermore, through theoretical analysis, we demonstrate that QIREN possesses a quantum advantage over classical FNNs. Lastly, we conducted experiments in signal representation, image superresolution, and image generation tasks to show the superior performance of QIREN compared to state-of-the-art (SOTA) models. Our work not only incorporates quantum advantages into implicit neural representations but also uncovers a promising application direction for Quantum Neural Networks. Our code is available at https://github.com/GGorMM1/QIREN.

1. Introduction

Implicit neural representations (INRs) are a novel research field where traditional discrete grid-sampled signal representations (e.g., images as discrete grids of pixels) are replaced by continuous functions. These continuous functions are typically implicitly defined, and INRs approximate them with deep neural networks which take low-dimensional coordinates as inputs and output numerical values, such as color, amplitude, and density, at corresponding positions. INRs are not coupled to spatial resolution, thus offering improved memory efficiency and overcoming resolution limitations (Dupont et al., 2021a). Most of early research on INRs is built on ReLU-based multilayer perceptrons (MLPs) (Park et al., 2019; Sitzmann et al., 2019; Jiang et al., 2020; Genova et al., 2019). However, there exists a frequency principle (Xu et al., 2019; Rahaman et al., 2019) that the spectrum of ReLU-based MLPs rapidly decays with increasing frequency, which leads to the limited capability of ReLU-based MLPs in modeling high-frequency components of signals, as illustrated in Figure 1.

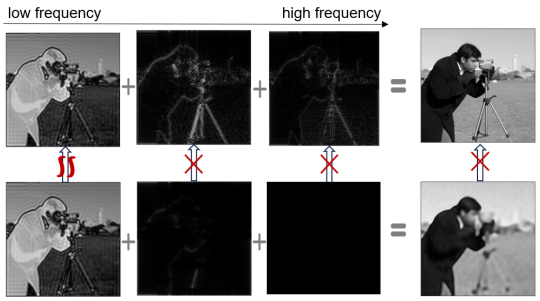


Figure 1. The different frequency components of a real image (top) and the image fitted by a ReLU-based MLP (bottom).

On one hand, recent studies (Mildenhall et al., 2021; Sitzmann et al., 2020; Tancik et al., 2020) have discovered that employing Fourier Neural Networks can partially address this issue. Specifically, by mapping the input using trigonometric functions or replacing the ReLU activation function with trigonometric functions, the capacity of the network to represent high-frequency components of the signal can be significantly enhanced. However, faced with increasingly complex fitting tasks in real-world applications, the classical Fourier Neural Network also requires a growing number of training parameters, which increases the demand for computing resources. It is crucial for the field of INRs and even machine learning to explore new solutions that can reduce these parameters and alleviate computational consumption.

On the other hand, quantum machine learning has recently been given heavy prominence in accelerating computing and saving parameters. Among them, Quantum Neural Networks (Farhi & Neven, 2018), also known as variational circuits (McClean et al., 2016), or quantum circuit learning (Mitarai et al., 2018), have demonstrated quantum advan-

[1]School of New Media and Communication, Tianjin University, China [2]College of Intelligence and Computing, Tianjin University, China. Correspondence to: Peng Zhang <pzhang@tju.edu.cn>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

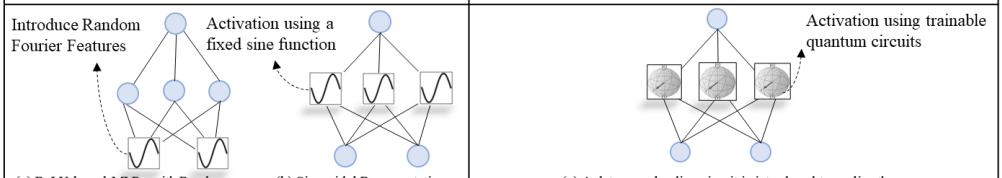
Domain	Implicit Neural Representations (INRs)	
Challenge	Modeling high-frequency components of signals	
Methodology	Fourier Neural Networks (FNNs)	
Model	Classical Implicit Representation Network	Quantum Implicit Representation Network (QIREN)
	 (a) ReLU-based MLPs with Random Fourier Features (RFF) (b) Sinusoidal Representation Network (SIREN)	(c) A data re-uploading circuit is introduced to realize the quantum generalization of FNN.
Signal representation capacity	Growing **linearly** with the model size	Growing **exponentially** with the size of quantum circuits under optimal conditions
Result	More parameters and less precise signal representation	**Fewer parameters and more precise signal representation**

Figure 2. Classical Fourier Neural Networks vs. Quantum Fourier Neural Networks.

tages over classical methods in various domains, including classification (Schuld et al., 2020), generative adversarial learning (Dallaire-Demers & Killoran, 2018), and deep reinforcement learning (Lockwood & Si, 2020). However, there is relatively little work demonstrating quantum advantages in real-world tasks that outperform classical methods. Identifying practical applications that reveal the benefits of quantum machine learning has been a goal of this community. Recent advances (Schuld et al., 2021; Yu et al., 2022) have found that data re-uploading quantum circuits, which involve uploading data repeatedly into quantum circuits, can be equivalently represented as Fourier series in the data. This inspires us to explore the application of Quantum Neural Networks with data re-uploading quantum circuits as core components in the field of INRs. We hope to achieve a more fine-grained representation of the signal with fewer parameters, as well as to find suitable task scenarios for quantum machine learning.

In this paper, we first demonstrate the exponential advantage of the data re-uploading quantum circuit over classical methods in representing Fourier series under optimal conditions. By combining these quantum circuits with classical layers, we have designed QIREN. Theoretically, QIREN has a stronger ability to fit the Fourier series than classical FNNs. This is reflected in the fact that QIREN can achieve a more refined signal representation with fewer parameters. Experimental results verify that QIREN indeed exhibits excellent performance, with a reduction in fitting error of up to 35% and a smaller number of parameters, compared to the SOTA model on the signal representation task. Additionally, it demonstrates improvements in performance for both image superresolution and image generation tasks. We exhibit the core ideas and main conclusions of this paper in Figure 2.

In summary, the following novel contributions are made in this paper:

- We propose QIREN, a novel quantum generalization of FNNs and highlight implicit neural representations as a field where the potential advantages of Quantum Neural Networks can truly be harnessed.
- We analyze the role of quantum circuits and classical layers in QIREN and demonstrate the quantum advantages of QIREN over classical FNNs.
- Through image representation and sound representation tasks, we show that QIREN outperforms SOTA models in signal representation, particularly in modeling high-frequency signals. Moreover, we have explored the application prospect of QIREN in tasks such as image superresolution and image generation.

The following sections are organized as follows. In Section 2, we review related work on implicit neural representations and data re-uploading quantum circuits. In Section 3, we introduce classical implicit representation networks. In Section 4, we propose quantum implicit representation networks and theoretically demonstrate the quantum advantage over classical implicit representation networks. In Section 5, we conduct experiments and show that QIREN has performance beyond the SOTA model in the field of INRs. We finally present our conclusions and outlook for future work in Section 6.

2. Related Work

2.1. Implicit Neural Representations

The original concept of INRs was first introduced in Compositional Pattern Producing Network (Stanley, 2007), which is a neuroevolution-based model that is trained to represent 2D images. Subsequently, INRs gained significant popularity in the field of 3D computer vision due to its provision of continuous, memory-efficient implicit representations for shape parts (Genova et al., 2019; Zhong et al., 2019), objects (Park et al., 2019), or scenes (Sitzmann et al., 2019; Jiang et al., 2020). However, early works mostly utilized ReLU-based MLPs, which lacked accuracy in modeling the high-frequency components of signals (Xu et al., 2019; Rahaman et al., 2019; Tancik et al., 2020). Mildenhall et al. (2021) significantly improved the capability of modeling high-frequency details in ReLU-based MLPs by applying Random Fourier Features (RFF) mapping to the inputs. Sitzmann et al. (2020) proposed Sinusoidal Representation Network (SIREN), which effectively represents high-frequency signals by using the sine activation function. The models proposed in these two works are considered SOTA models in the field of INRs and have inspired many subsequent works (Woo et al., 2023; Huang et al., 2022; Skorokhodov et al., 2021; Chan et al., 2021). Nevertheless, the capacity of these classical SOTA models to represent signals is constrained by their linear growth with model size. Our proposed model achieves a stronger signal representation capability, by leveraging the exponential advantage of the data re-uploading quantum circuit over classical methods in representing Fourier series.

2.2. Data Re-uploading Quantum Circuits

Pérez-Salinas et al. (2020) first introduced the concept of data re-uploading quantum circuits and demonstrated their capability in handling certain point set classification tasks. Wach et al. (2023) further explored their classification performance and conducted experiments on image classification. The aforementioned studies have only experimented on simple synthetic datasets and have not found a suitable practical application scenario for data re-uploading quantum circuits.

Some studies concentrate on investigating the mathematical properties of data re-uploading quantum circuits. Schuld et al. (2021) revealed the fundamental nature of data re-uploading quantum circuits as Fourier series. Yu et al. (2022) further investigated the universality of single-qubit data re-uploading circuits. These works only explore the multi-layer single-qubit circuits or single-layer multi-qubit circuits, and they do not compare with classical methods to demonstrate the advantages of data re-uploading quantum circuits. We extended our research to multi-layer multi-qubit quantum circuits. Furthermore, we demonstrated that data re-uploading circuits exhibit exponential advantages of fitting Fourier series over classical methods.

3. Classical Implicit Representation Network

Two SOTA models in the field of implicit neural representations, ReLU-based MLPs with Random Fourier Features and Sinusoidal Representation Network, have been shown to be FNNs by Benbarka et al. (2022). We review this in this section and further demonstrate that the signal representation capability of these classical FNNs grows linearly with model size.

A perceptron with RFF has the following form:

$$g(\mathbf{x}) = \mathbf{W} \cdot \begin{pmatrix} \cos(2\pi \mathbf{M} \cdot \mathbf{x}) \\ \sin(2\pi \mathbf{M} \cdot \mathbf{x}) \end{pmatrix} + b, \tag{1}$$

where $\mathbf{W} \in \mathbb{R}^{1\times 2m}$ is the parameter matrix, $\mathbf{M} \in \mathbb{R}^{m\times d_{in}}$ is the random Fourier mapping matrix (m determines the size of the matrix, serving as a measure of the model size), $\mathbf{x} \in \mathbb{R}^{d_{in}}$ is the input and $b \in \mathbb{R}$ is the bias. Next, we will relate Eq. (1) to the general equation of Fourier series.

A Fourier series is a weighted sum of sine and cosine functions. When the number of terms approaches infinity, the Fourier series can approximate any periodic function. Even though the function $g(\mathbf{x})$ may not be periodic, the inputs are generally bounded. Therefore, we can assume that $g(\mathbf{x})$ is periodic outside the boundaries of its inputs and its Fourier series takes the following form:

$$g(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{Z}^{d_{in}}} c_{\mathbf{n}} \mathrm{e}^{2\pi \mathrm{i} \frac{\mathbf{n}}{\mathbf{T}}\cdot\mathbf{x}}, \tag{2}$$

where $c_{\mathbf{n}}$ denotes the coefficient, and $\mathbf{T}$ denotes the period of $g(\mathbf{x})$. For real-valued functions, it holds that $c_{\mathbf{n}} = c^*_{-\mathbf{n}}$. Therefore Eq. (2) can be rewritten as

$$g(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{Z}^{d_{in}}} a_{\mathbf{n}} \cos(2\pi \frac{\mathbf{n}}{\mathbf{T}} \cdot \mathbf{x}) + b_{\mathbf{n}} \sin(2\pi \frac{\mathbf{n}}{\mathbf{T}} \cdot \mathbf{x}). \tag{3}$$

If we write Eq. (3) in vector form, we get

$$g(\mathbf{x}) = (a_{\mathbf{N}}, b_{\mathbf{N}}) \cdot \begin{pmatrix} \cos(2\pi \mathbf{N} \cdot \mathbf{x}) \\ \sin(2\pi \mathbf{N} \cdot \mathbf{x}) \end{pmatrix}, \tag{4}$$

where $\mathbf{N} = \{\frac{\mathbf{n}}{\mathbf{T}}\}_{\mathbf{n}\in\mathbb{Z}^{d_{in}}}$ denotes frequency spectrum. Now, when we compare Eq. (1) and (4), we find similarities. Obviously, a perceptron with RFF constructs a Fourier series, where $\mathbf{W}$ includes coefficients of $g(\mathbf{x})$, and $\mathbf{M}$ includes frequencies of $g(\mathbf{x})$.

SIREN is a MLP that utilizes the sine activation function. Considering SIREN with one hidden layer, we get

$$g(\mathbf{x}) = \mathbf{W_2} \cdot \sin(2\pi \mathbf{W_1} \cdot \mathbf{x} + \boldsymbol{\theta}) + b, \tag{5}$$

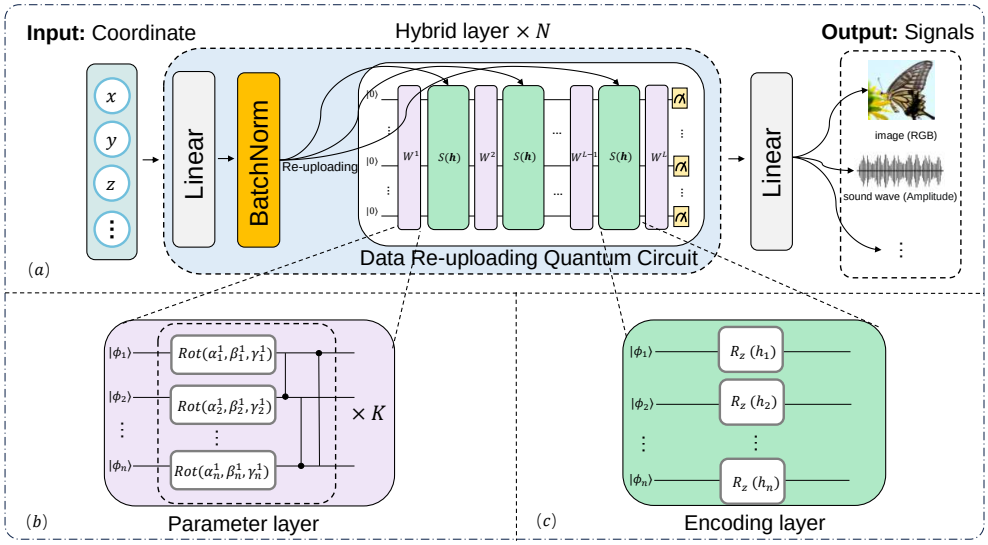


Figure 3. Architecture of QIREN. (a) presents the overall architecture of QIREN. (b) and (c) respectively illustrate the implementation details of the parameter layer and the encoding layer.

where $\mathbf{W_1} \in \mathbb{R}^{2m \times d_{in}}$ and $\mathbf{W_2} \in \mathbb{R}^{1 \times 2m}$ are the parameter matrices and $\boldsymbol{\theta} \in \mathbb{R}^{2m}$ is the bias vector. We can consider a specific case when $\boldsymbol{\theta} = (\pi/2, \ldots, \pi/2, 0, \ldots, 0)^T$. In this case, the SIREN and the perceptron with RFF have nearly identical forms, except that $\mathbf{W_2}$ is a trainable parameter matrix and $\mathbf{M}$ is an untrained random matrix. In fact, any situation can be transformed into this special case, as we can always express Eq. (5) as

$$g(\mathbf{x}) = \mathbf{W_2} \cdot \sin(2\pi \mathbf{W_1} \cdot \mathbf{x} + \hat{\boldsymbol{\theta}} + \begin{pmatrix} \frac{\pi}{2} \\ 0 \end{pmatrix}) + b, \quad (6)$$

where $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} - \begin{pmatrix} \frac{\pi}{2} \\ 0 \end{pmatrix}$.

At this point, we can conclude that both ReLU-based MLP with RFF and SIREN are FNNs as their fundamental blocks approximate functions in the form of Fourier series. Their performance is directly correlated with the size of the Fourier series they can represent. **Therefore, the signal representation capacity of classical FNNs increases linearly with the model size** m.

4. Quantum Implicit Representation Network

In Section 4.1 and 4.2, we introduce the architecture of QIREN and its nature as a Fourier Neural Network. In Section 4.3, we analyze the role of each component in QIREN.

4.1. Overall Architecture of QIREN

The overall architecture of QIREN is shown in Figure 3(a), which consists of N hybrid layers and a Linear layer at the end. The model takes coordinates as inputs and outputs signal values. Correspondingly, the dataset $D = \{(\mathbf{x}_i, \mathbf{y}_i)\}$ is a set of tuples of coordinates $\mathbf{x}_i \in \mathbb{R}^{d_{in}}$ along with signals $\mathbf{y}_i \in \mathbb{R}^{d_{out}}$. Next, we provide a detailed description of the model.

The data $\mathbf{x}_i$ initially enters the hybrid layer, starting with a Linear layer and a BatchNorm layer. We get

$$\mathbf{h}_i = \text{BatchNorm}(\mathbf{W}\mathbf{x}_i + \mathbf{b}), \quad (7)$$

where $\mathbf{W}$ and $\mathbf{b}$ are parameters of the Linear layer. Then $\mathbf{h}_i$ is input to the data re-uploading quantum circuit QC (Preliminaries of the quantum circuit is presented in the Appendix A). In Figure 3(b) and (c), we present the implementation structures for the parameter layer W and encoding layer $S(\mathbf{h}_i)$ of the quantum circuit. The parameter layer consists of K stacked blocks. Each block contains Rot gates applied to each qubit and $CNOT$ gates that connects qubits in a cyclic manner. The encoding layer applies R_Z gates on each qubit. Finally, we measure the expectation value of the quantum state with observables. The output of the quantum circuit is given by

$$\mathbf{f}_i = [\text{QC}(\mathbf{h}_i, O_1), \ldots, \text{QC}(\mathbf{h}_i, O_{d_f})], \quad (8)$$

where O denotes an arbitrary observable. The output $\mathbf{f}_i^{(n)}$ of the n-th hybrid layer will be used as the input of the $(n+1)$-th layer. At the end, we add a Linear layer to receive $\mathbf{f}_i^{(N)}$ and output $\mathbf{y'}_i$. We use Mean Squared Error (MSE) as the loss function to train the model:

$$\text{MSE} = \frac{1}{|D|} \sum_i (\mathbf{y}_i - \mathbf{y'}_i)^2. \quad (9)$$

4.2. QIREN as a Quantum Generalization of FNNs

Schuld et al. (2021) derives that the data re-uploading circuit is essentially the Fourier series and analyzes it on single-layer multi-qubit and multi-layer single-qubit circuits. Based on this, we generalize the analysis to multi-layer multi-qubit circuits and show that QIREN is essentially a Fourier Neural Network.

Each component of Eq. (8) can be defined as $f(\mathbf{h})$, given by the multi-layer multi-qubit data re-uploading quantum circuit, which is represented as follows:

$$f(\mathbf{h}) = \left\langle 0 \left| U^{\dagger}(\mathbf{h}) O U(\mathbf{h}) \right| 0 \right\rangle, \tag{10}$$

where

$$U(\mathbf{h}) = W^{(L)} S(\mathbf{h}) W^{(L-1)} \dots W^{(2)} S(\mathbf{h}) W^{(1)}. \tag{11}$$

In the data re-uploading quantum circuit, each encoding layer has the form $S(\mathbf{h}) := \mathrm{e}^{-\mathrm{i}h_1 H} \otimes \dots \otimes \mathrm{e}^{-\mathrm{i}h_{d_h} H}$, where H is a d-qubit Hamiltonian that depends on the actual quantum gates used, and there always exists an eigenvalue decomposition $H = V^{\dagger} \Sigma V$. Without loss of generality, we can assume that all Hamiltonians are diagonal as V and $V^{\dagger}$ can be absorbed into adjacent parameter layers. With this assumption, we note that the $S(\mathbf{h})$ is diagonal and l-th $S(\mathbf{h})$ has the form:

$$[S(\mathbf{h})]_{(j_1^{(l)},\dots,j_{d_h}^{(l)}),(j_1^{(l)},\dots,j_{d_h}^{(l)})} = \mathrm{e}^{-\mathrm{i}\left(\lambda_{j_1^{(l)}},\dots,\lambda_{j_{d_h}^{(l)}}\right)\cdot \mathbf{h}}, \tag{12}$$

where λ is the eigenvalue of H. We can rewrite Eq. (12) by introducing the multi-index $\mathbf{j}^{(l)} = \left\{ j_1^{(l)}, \dots, j_{d_h}^{(l)} \right\} \in \{1, \dots, 2^d\}^{d_h}$ as follows:

$$[S(\mathbf{h})]_{\mathbf{j}^{(l)},\mathbf{j}^{(l)}} = \mathrm{e}^{-\mathrm{i}\boldsymbol{\lambda}_{\mathbf{j}^{(l)}}\cdot \mathbf{h}}, \tag{13}$$

where $\boldsymbol{\lambda}_{\mathbf{j}^{(l)}} = \left(\lambda_{j_1^{(l)}}, \dots, \lambda_{j_{d_h}^{(l)}} \right)$. In this way, we can derive the mathematical expression of the quantum state $|\psi\rangle$:

$$|\psi\rangle_i = [U(\mathbf{h})|0\rangle]_i = \sum_{\mathbf{j}^{(1)}\dots\mathbf{j}^{(L-1)}} \mathrm{e}^{-\mathrm{i}\left(\boldsymbol{\lambda}_{\mathbf{j}^{(1)}} + \dots + \boldsymbol{\lambda}_{\mathbf{j}^{(L-1)}}\right)\cdot \mathbf{h}} \times W^{(L)}_{i\mathbf{j}^{(L-1)}} \dots W^{(2)}_{\mathbf{j}^{(2)}\mathbf{j}^{(1)}} W^{(1)}_{\mathbf{j}^{(1)}1}, \tag{14}$$

where the subscript of W indicates the row and column. We introduce another multi-index $\mathbf{J} = \left\{ \mathbf{j}^{(1)}, \dots, \mathbf{j}^{(L-1)} \right\} \in \{1, \dots, 2^d\}^{d_h \times (L-1)}$ to simplify Eq. (14) as follows:

$$|\psi\rangle_i = \sum_{\mathbf{J}} \mathrm{e}^{-\mathrm{i}\boldsymbol{\Lambda}_{\mathbf{J}}\cdot \mathbf{h}} W^{(L)}_{i\mathbf{j}^{(L-1)}} \dots W^{(2)}_{\mathbf{j}^{(2)}\mathbf{j}^{(1)}} W^{(1)}_{\mathbf{j}^{(1)}1}, \tag{15}$$

where $\boldsymbol{\Lambda}_{\mathbf{J}} = \boldsymbol{\lambda}_{\mathbf{j}^{(1)}} + \dots + \boldsymbol{\lambda}_{\mathbf{j}^{(L-1)}}$. Next, according to Eq. (10), we calculate the expectation value of $|\psi\rangle$ and its conjugate transpose $\langle\psi|$ with respect to O, and get

$$f(\mathbf{h}) = \langle\psi|O|\psi\rangle = \sum_{\mathbf{K},\mathbf{J}} \mathrm{e}^{\mathrm{i}(\boldsymbol{\Lambda}_{\mathbf{K}} - \boldsymbol{\Lambda}_{\mathbf{J}})\cdot \mathbf{h}} \sum_{i,i'} W^{\dagger(1)}_{\mathbf{k}^{(1)}1} W^{\dagger(2)}_{\mathbf{k}^{(2)}\mathbf{k}^{(1)}} \dots W^{\dagger(L)}_{i\mathbf{k}^{(L-1)}} O_{i,i'} W^{(L)}_{i'\mathbf{j}^{(L-1)}} \dots W^{(2)}_{\mathbf{j}^{(2)}\mathbf{j}^{(1)}} W^{(1)}_{\mathbf{j}^{(1)}1}. \tag{16}$$

We are currently not concerned with the specific form of the observable O and the parameter layers. So we substitute $a_{\mathbf{K},\mathbf{J}}$ for the second summation term in Eq. (16), resulting in a more intuitive form:

$$f(\mathbf{h}) = \sum_{\mathbf{K},\mathbf{J}} a_{\mathbf{K},\mathbf{J}} \mathrm{e}^{\mathrm{i}(\boldsymbol{\Lambda}_{\mathbf{K}} - \boldsymbol{\Lambda}_{\mathbf{J}})\cdot \mathbf{h}}, \tag{17}$$

where $a_{\mathbf{K},\mathbf{J}}$ denotes the coefficient, $\boldsymbol{\Lambda}_{\mathbf{K}} - \boldsymbol{\Lambda}_{\mathbf{J}}$ denotes the frequency and $\{\boldsymbol{\Lambda}_{\mathbf{K}} - \boldsymbol{\Lambda}_{\mathbf{J}}\}_{\mathbf{K},\mathbf{J}}$ denotes the frequency spectrum. Eq. (17) is now evidently equivalent to the Fourier series. Utilizing this quantum circuit enables QIREN to approximate functions in the form of the Fourier series. Thus, we have demonstrated that QIREN is a quantum generalization of FNNs.

4.3. Theoretical Analysis of QIREN

In this subsection, we analyze the roles of each component in QIREN and demonstrate the quantum advantages of QIREN over classical FNNs. We summarize them into three claims.

Claim 1 *The capability of the data re-uploading quantum circuit to represent Fourier series grows exponentially with the size of the circuit under optimal conditions.*

We consider a specific quantum circuit with $d \times d_h$ qubits, where the data is repeatedly uploaded L times, and $H = \sum_{q=1}^{d} Z^{(q)}/2$ ($Z^{(q)}$ denotes Pauli Z acting on the q-th qubit). Regarding H, it has 2^d eigenvalues with the $d+1$ unique entries:

$$\mathbf{P} = \{p - \frac{d}{2}, \mid p \in \{0, \dots, d\}\}. \tag{18}$$

In this case, the frequency spectrum of $f(\mathbf{h})$ can be represented as

$$\begin{aligned} \{\boldsymbol{\Lambda}_{\mathbf{K}} - \boldsymbol{\Lambda}_{\mathbf{J}}\}_{\mathbf{K},\mathbf{J}} &= \left\{ \boldsymbol{\lambda}_{\mathbf{k}^{(1)}} + \dots + \boldsymbol{\lambda}_{\mathbf{k}^{(L)}} - (\boldsymbol{\lambda}_{\mathbf{j}^{(1)}} + \dots + \boldsymbol{\lambda}_{\mathbf{j}^{(L)}}) \right\} \\ &= \left\{ \left(\lambda_{k_1^{(1)}} + \dots + \lambda_{k_1^{(L)}} \right) - \left(\lambda_{j_1^{(1)}} + \dots + \lambda_{j_1^{(L)}} \right) \mid \right. \\ &\qquad \left. \lambda_{k_1^{(1)}}, \dots, \lambda_{k_1^{(L)}}, \lambda_{j_1^{(1)}}, \dots, \lambda_{j_1^{(L)}} \in \mathbf{P} \right\}^{d_h} \\ &= \{-dL, \dots, dL\}^{d_h}. \end{aligned} \tag{19}$$

It can be seen that the spectrum grows exponentially with d_h. **With the help of the Linear layer, the spectrum can**

Method	Sound Representation			Image Representation					Image Superresolution		
	Cello	#params	#mem(%)	Astronaut	Camera	Coffee	#params	#mem(%)	Astronaut	Camera	Coffee
Nearest	-	-	-	-	-	-	-	-	26.6	10.4	13.6
Bilinear	-	-	-	-	-	-	-	-	25.2	9.2	12.2
ReLU	6.8	831	16.9	9.9	2.7	4.2	841	17.9	33.8	11.3	13.5
Tanh	14.0	831	16.9	20.7	5.8	14.8	841	17.9	47.8	15.2	26.7
ReLU+RFF*	6.0	791	20.9	5.1	1.9	4.9	791	22.8	39.9	13.3	23.9
SIREN*	**5.5**	691	30.9	9.0	1.5	2.3	701	31.5	77.0	26.3	15.7
QIREN (ours)	**5.5**	**649**	**35.1**	**4.0**	**1.1**	**1.5**	**657**	**35.8**	**24.3**	**7.9**	**9.4**

Table 1. MSE ($\times 10^{-3}$) of different models on signal representation and image superresolution tasks. The best results are highlighted in **bold**. The models widely regarded as SOTA are marked with $\star$. #params denote the number of model parameters, and #mem denotes the memory saved by the model compared to the discrete grid-sampled representation.

grow exponentially with the size of the circuit $d \times d_h$ at most, as will be explained in Claim 2. However, one unsatisfactory point is that we cannot guarantee that the coefficients $a_{\mathbf{K},\mathbf{J}}$ are arbitrary because this would require us to have the ability to construct arbitrary unitary matrices to some extent. According to the Solovay-Kitaev theorem (Nielsen & Chuang, 2010) which follows that any arbitrary unitary matrix on n qubits can be approximated within an error ϵ using $O\left(n^2 4^n \log^c\left(n^2 4^n/\epsilon\right)\right)$ two-qubit gates (c is a constant), constructing arbitrary unitary matrices would consume exponential resources.

Fortunately, the coefficients of most functions we need to fit in practice are not arbitrary but rather follow certain prior distributions (e.g., the amplitudes of high-frequency components are often small in natural signals). This can draw an analogy to quantum many-body physics, where the majority of natural systems have Hamiltonians that can be expressed as a sum of local Hamiltonians (Gao et al., 2022). Theorem 1 (Poulin et al., 2011) states that approximating the evolution process of such a system only requires a polynomial number of parameters.

Theorem 1 *The time evolution operator corresponding to a Hamiltonian composed of L k-body terms, with a total evolution time of T, can be simulated by a quantum circuit of polynomial size. The total number of standard universal 2-qubit quantum gates needed to approximate the complete time evolution with an error ϵ is upper bounded by*

$$G_{\text{tot}} = d_{SK} \frac{c_{\max}^2 T^2 L^3}{\epsilon} \log^{cSK}\left(\frac{c_{\max}^2 T^2 L^3}{\epsilon^2}\right) \tag{20}$$

where cSK and dSK are constants, $c_{\max}$ is the maximum norm of the k-body terms in the Hamiltonian.

Hence, we can assume that the Hamiltonian corresponding to the function we aim to fit can also be decomposed into a sum of local Hamiltonians. When the ansatz of the parameter layer is well-adapted to approximate this Hamiltonian, a polynomial number of parameters in the parameter layer is sufficient to effectively fit the function. At this point, we can conclude Claim 1. Furthermore, utilizing this quantum circuit can empower QIREN with quantum advantages over classical FNNs.

Claim 2 *The role of the Linear layer is to further expand the spectrum and adjust the frequency, leading to improved fitting performance.*

By observing Eq. (13), we can understand that applying a Linear layer before uploading the data to the quantum circuit is equivalent to adjusting the eigenvalues in the Hamiltonian, ultimately affecting the frequency spectrum. Using the Linear layer brings two advantages. First, it can make the frequency spectrum larger. From Eq. (18) and (19), it can be deduced that certain sums of eigenvalues generate the same term in the spectrum. This redundancy can be reduced by using the Linear layer, and the size of the spectrum can be extended from $(2dL+1)^{d_h}$ to up to $((3^d-1)L+1)^{d_h}$. Second, it enables the coverage range of the frequency spectrum to be adjustable, aiming to cover frequencies with larger coefficients, which are more important. Therefore, incorporating Linear layers can further enhance the fitting performance of QIREN. A more detailed explanation of Claim 2 can be found in Appendix B.

Claim 3 *The role of the Batchnorm layer is to accelerate the convergence of our quantum model.*

In feedforward neural networks, the data often passes through a BatchNorm layer before the activation function, which effectively prevents the vanishing gradient problem (Ioffe & Szegedy, 2015). Similarly, in QIREN, the quantum circuits replace activation functions and play a role in providing non-linearity (The quantum circuit itself is linear, but the process of uploading classical data to the quantum circuit is non-linear). Therefore, we add the BatchNorm layer here, with the aim of stabilizing and accelerating the convergence of the model. We have further validated its

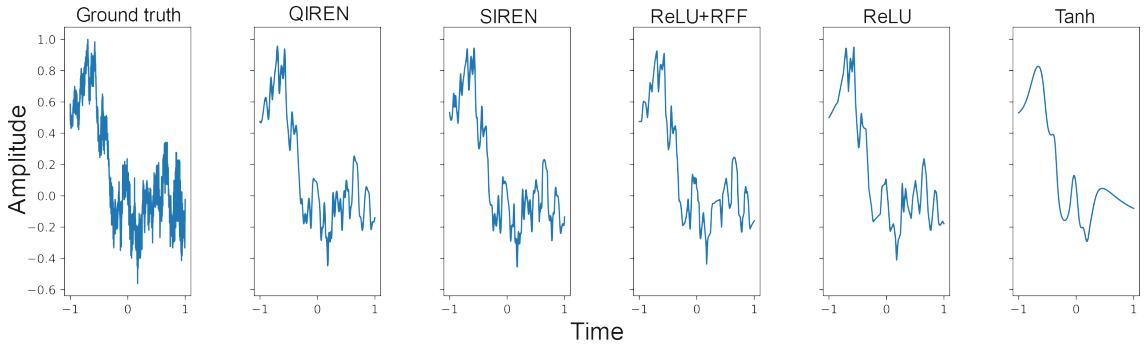


Figure 4. Results of sound representation. The Y-axis represents the normalized amplitude of the sound wave. The X-axis represents the time.

Figure 5. Results of image representation.

effectiveness through ablation experiments in Appendix E.

5. Experiments

In Section 5.1, we validate the superior performance of QIREN in representing signals, especially high-frequency signals, through image representation and sound representation tasks. In Section 5.2 and 5.3, we show the expanded applications of QIREN in tasks such as image superresolution and image generation. To ensure a fair comparison, we adjust the hyperparameters to approximately match the number of parameters between the QIREN and the baseline models. In each experiment, each model is trained five times, and the best performance is selected as the comparison result. Further details regarding the model implementation can be found in Appendix C. All our experiments are conducted on a simulation platform, utilizing the Pennylane (Bergholm et al., 2018) and TorchQauntum (Wang et al., 2022).

5.1. Signal Representation

Task definition. The goal is to use neural networks to represent signals. Specifically, we aim to train a model that takes one-dimensional or two-dimensional coordinates as input and outputs the corresponding amplitude or grayscale value at that position.

Datasets and Evaluation. In the sound representation task, we take a small section of one movement from Bach's Cello Suites, sample 1000 points at equal intervals, and normalize the amplitudes. We set the corresponding timestamps to be evenly spaced within $[-1, 1]$. In the image representation task, we utilize three popular images: Astronaut, Camera, and Coffee (Van der Walt et al., 2014). These images are cropped and down-sampled to dimensions of 32×32 pixels, resulting in three datasets, each consisting of 1024 pixels. We select the following models as our baselines: ReLU-based MLP, Tanh-based MLP, ReLU-based MLP with RFF (Mildenhall et al., 2021) and SIREN (Sitzmann et al., 2020). MSE is used as a metric to evaluate the difference between the images or sound fitted by the model and the ground truth.

Results. The experimental results are presented in Figure 4, Figure 5 and Table 1. QIREN and SIREN exhibit similar performance on the sound representation task. Although the performance of these two models seems comparable,

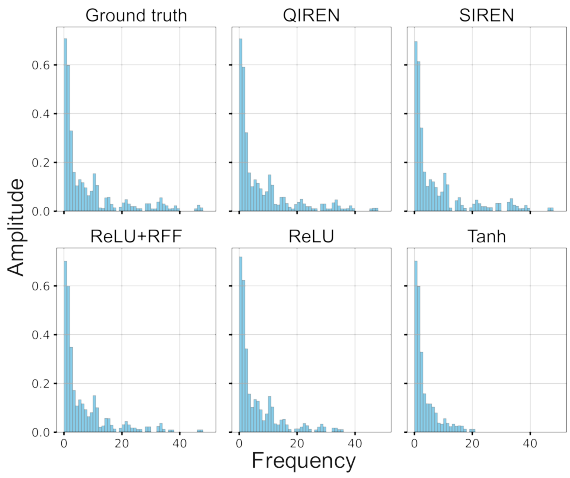


Figure 6. The frequency spectrum of the output of models on the sound representation task.

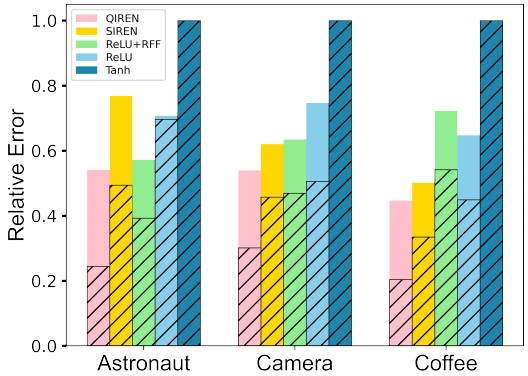


Figure 7. The relative error of each model compared to the Tanh-based MLP. The shaded area represents the low-frequency error, while the non-shaded area represents the high-frequency error.

it is worth emphasizing that our model achieves a significant memory saving of 35.1% with the minimum number of parameters and the convergence of SIREN requires proper setting of hyperparameters, whereas our model does not encounter such limitations. Then we explored the output of models from a frequency perspective. We visualize the frequency spectrum of the output of models in Figure 6. It is evident that the low-frequency distribution of the outputs from different models all resembles the ground truth. However, when it comes to the high-frequency distribution, QIREN and SIREN exhibit superior performance, followed by the ReLU-based MLP with RFF. ReLU-based and Tanh-based MLPs even lack the high-frequency parts.

QIREN achieved optimal performance with the fewest parameters on the image representation task, resulting in a maximum reduction of 34.8% in error compared to the SOTA model. To further explore the signal representation capabilities of the models, we used filters to separate the high-frequency and low-frequency components of their outputs and compared the fitting errors of these two components individually. The results are presented in the Figure 7. QIREN consistently achieved the lowest error in fitting both high-frequency and low-frequency components.

5.2. Image Superresolution

Task definition. The goal is to increase the resolution of an image while maintaining its content and details as much as possible. In this task, we partition the coordinates into a grid of size 64×64 and use them as input for the model trained on 32×32 pixel images to reconstruct 64×64 pixel images.

Datasets and Evaluation. This task directly utilizes the models trained on image representation tasks. The images Astronaut, Camera, and Coffee are down-sampled to create 64×64 pixel images, which serve as the ground truth for evaluation. For a more comprehensive comparison, in addition to the baseline models used in the signal representation task, we also include bilinear interpolation and nearest-neighbor interpolation. The evaluation metric remains MSE.

Results. The experimental results are presented in Figure 8 and Table 1. As observed in Table 1, the performance of classical Implicit Representation Networks failed to surpass that of interpolation methods, possibly due to the limited number of parameters in the network. However, even in such circumstances, QIREN outperforms interpolation methods by achieving a maximum error reduction of 24.0%. This stems from the superior signal representation capability of QIREN, as compared to other models, the parameterized implicit function of QIREN exhibits a distribution that is more similar to the target function.

5.3. Image Generation

Task definition. The goal of this task is to generate facial images that are indistinguishable from real images with the INR-based generator. Recent advancements (Skorokhodov et al., 2021; Shaham et al., 2021; Anokhin et al., 2021; Dupont et al., 2021b; Chan et al., 2021) have introduced a groundbreaking framework and extended INRs to image generation. More specifically, this framework utilizes a hypernetwork (Ha et al., 2016) that takes random distributions as input to generate the parameters of the Implicit Representation Network. Subsequently, these generated parameters are assigned to the Implicit Representation Network. Finally, the Implicit Representation Network takes coordinate-based inputs and generates images as output. An adversarial approach (Goodfellow et al., 2020) is employed to ensure the generated images align with our desired outcomes. In this task, we adopt such a framework and build

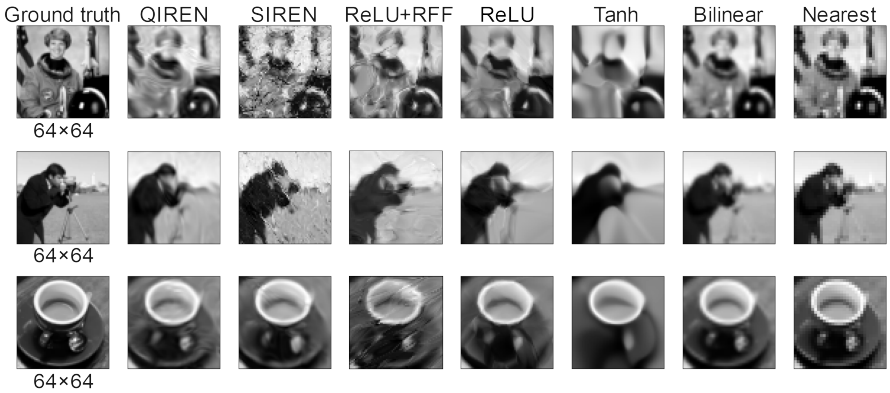


Figure 8. Results of image superresolution.

upon the foundation of StyleGAN2 (Karras et al., 2020).

Datasets and Evaluation. FFHQ is a high-resolution dataset of 70k human faces (Karras et al., 2019) and CelebA-HQ is a high-quality version of CelebA that consists of 30k images (Karras et al., 2017). We downsample the images in these two datasets to obtain 32×32 pixel images for experimentation. In this task, we utilize the same baseline models as those used in the signal representation task. We evaluate the model using Frechet Inception Distance (FID) (Heusel et al., 2017) metric using 50k images to compute the statistics.

Results. The experimental results are presented in Table 2. We observed that the ReLU-based MLP is completely inadequate for image generation tasks, whereas RFF can significantly enhance its performance. Tanh-based MLP and SIREN show similar performance. Compared to these classical implicit representation networks, QIREN achieves superior performance with the fewest number of parameters. In Appendix D, we will show the images generated by QIREN and further explore its exciting properties, such as out-of-the-box superresolution and meaningful image-space interpolation.

Method	FFHQ	CelebA-HQ	#params
Tanh	26.98	25.17	1.16M
ReLU	84.94	110.81	1.16M
ReLU+RFF*	15.01	13.91	1.14M
SIREN*	22.31	20.97	1.16M
QIREN (ours)	**11.53**	**11.78**	**1.13M**

Table 2. FID scores of different models on FFHQ and CelebA-HQ datasets.

6. Conclusion and Future Work

In this paper, we first derive that QIREN is a quantum generalization of FNNs. Subsequently, we perform a theoretical analysis of the data re-uploading quantum circuits, Linear layers, and BatchNorm layers, demonstrating the ability of QIREN to represent the Fourier series is exponentially stronger over classical FNNs. Lastly, we conduct experiments on signal representation, image superresolution, and image generation. The results demonstrate that QIREN significantly outperforms SOTA models.

It is worth emphasizing that INRs have many other potential applications such as representing scenes or 3D objects, time series forecasting, and solving the differential equation. For a large class of tasks that model continuous signals, we can consider introducing INRs as the basic component. Based on the theoretical and experimental foundations in this paper, we can extend QIREN to these applications in future work and QIREN is anticipated to yield superior results with fewer parameters in these domains. Simultaneously, we have found a suitable application scenario for quantum machine learning. This will stimulate further practical and innovative research within the quantum machine learning community.

Acknowledgements

This work is supported in part by the Natural Science Foundation of China (grant No. 62276188), TJU-Wenge joint laboratory funding.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Anokhin, I., Demochkin, K., Khakhulin, T., Sterkin, G., Lempitsky, V., and Korzhenkov, D. Image generators with

conditionally-independent pixel synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14278–14287, 2021.

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

Bergholm, V., Izaac, J., Schuld, M., Gogolin, C., Ahmed, S., Ajith, V., Alam, M. S., Alonso-Linaje, G., Akash-Narayanan, B., Asadi, A., et al. Pennylane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.

Chan, E. R., Monteiro, M., Kellnhofer, P., Wu, J., and Wetzstein, G. pi-gan: Periodic implicit generative adversarial networks for 3d-aware image synthesis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5799–5809, 2021.

Dallaire-Demers, P.-L. and Killoran, N. Quantum generative adversarial networks. *Physical Review A*, 98(1):012324, 2018.

Dupont, E., Goliński, A., Alizadeh, M., Teh, Y. W., and Doucet, A. Coin: Compression with implicit neural representations. *arXiv preprint arXiv:2103.03123*, 2021a.

Dupont, E., Teh, Y. W., and Doucet, A. Generative models as distributions of functions. *arXiv preprint arXiv:2102.04776*, 2021b.

Farhi, E. and Neven, H. Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*, 2018.

Gao, X., Anschuetz, E. R., Wang, S.-T., Cirac, J. I., and Lukin, M. D. Enhancing generative models via quantum correlations. *Physical Review X*, 12(2):021037, 2022.

Genova, K., Cole, F., Vlasic, D., Sarna, A., Freeman, W. T., and Funkhouser, T. Learning shape templates with structured implicit functions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 7154–7164, 2019.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

Ha, D., Dai, A. M., and Le, Q. V. Hypernetworks. *CoRR*, abs/1609.09106, 2016. URL http://arxiv.org/abs/1609.09106.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

Huang, X., Liu, H., Shi, B., Wang, Z., Yang, K., Li, Y., Wang, M., Chu, H., Zhou, J., Yu, F., et al. A universal pinns method for solving partial differential equations with a point source. In *IJCAI*, pp. 3839–3846, 2022.

Ioffe, S. and Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pp. 448–456. pmlr, 2015.

Jiang, C., Sud, A., Makadia, A., Huang, J., Nießner, M., Funkhouser, T., et al. Local implicit grid representations for 3d scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6001–6010, 2020.

Karras, T., Aila, T., Laine, S., and Lehtinen, J. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.

Karras, T., Laine, S., and Aila, T. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401–4410, 2019.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8110–8119, 2020.

Lockwood, O. and Si, M. Reinforcement learning with quantum variational circuit. In *Proceedings of the AAAI conference on artificial intelligence and interactive digital entertainment*, volume 16, pp. 245–251, 2020.

McClean, J. R., Romero, J., Babbush, R., and Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

Mitarai, K., Negoro, M., Kitagawa, M., and Fujii, K. Quantum circuit learning. *Physical Review A*, 98(3):032309, 2018.

Nielsen, M. A. and Chuang, I. L. *Quantum computation and quantum information*. Cambridge university press, 2010.

Park, J. J., Florence, P., Straub, J., Newcombe, R., and Lovegrove, S. Deepsdf: Learning continuous signed distance

functions for shape representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 165–174, 2019.

Pérez-Salinas, A., Cervera-Lierta, A., Gil-Fuster, E., and Latorre, J. I. Data re-uploading for a universal quantum classifier. *Quantum*, 4:226, 2020.

Poulin, D., Qarry, A., Somma, R., and Verstraete, F. Quantum simulation of time-dependent hamiltonians and the convenient illusion of hilbert space. *Physical review letters*, 106(17):170501, 2011.

Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y., and Courville, A. On the spectral bias of neural networks. In *International Conference on Machine Learning*, pp. 5301–5310. PMLR, 2019.

Schuld, M., Bocharov, A., Svore, K. M., and Wiebe, N. Circuit-centric quantum classifiers. *Physical Review A*, 101(3):032308, 2020.

Schuld, M., Sweke, R., and Meyer, J. J. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3): 032430, 2021.

Shaham, T. R., Gharbi, M., Zhang, R., Shechtman, E., and Michaeli, T. Spatially-adaptive pixelwise networks for fast image translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14882–14891, 2021.

Shin, S., Teo, Y., and Jeong, H. Exponential data encoding for quantum supervised learning. *Physical Review A*, 107 (1):012422, 2023.

Sitzmann, V., Zollhöfer, M., and Wetzstein, G. Scene representation networks: Continuous 3d-structure-aware neural scene representations. *Advances in Neural Information Processing Systems*, 32, 2019.

Sitzmann, V., Martel, J., Bergman, A., Lindell, D., and Wetzstein, G. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems*, 33:7462–7473, 2020.

Skorokhodov, I., Ignatyev, S., and Elhoseiny, M. Adversarial generation of continuous images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10753–10764, 2021.

Stanley, K. O. Compositional pattern producing networks: A novel abstraction of development. *Genetic programming and evolvable machines*, 8:131–162, 2007.

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

Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.

Wach, N. L., Rudolph, M. S., Jendrzejewski, F., and Schmitt, S. Data re-uploading with a single qudit. *arXiv preprint arXiv:2302.13932*, 2023.

Wang, H., Ding, Y., Gu, J., Lin, Y., Pan, D. Z., Chong, F. T., and Han, S. Quantumnas: Noise-adaptive search for robust quantum circuits. In *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, pp. 692–708. IEEE, 2022.

Woo, G., Liu, C., Sahoo, D., Kumar, A., and Hoi, S. Learning deep time-index models for time series forecasting. In *International Conference on Machine Learning*, pp. 37217–37237. PMLR, 2023.

Xu, Z.-Q. J., Zhang, Y., Luo, T., Xiao, Y., and Ma, Z. Frequency principle: Fourier analysis sheds light on deep neural networks. *arXiv preprint arXiv:1901.06523*, 2019.

Yu, Z., Yao, H., Li, M., and Wang, X. Power and limitations of single-qubit native quantum neural networks. *Advances in Neural Information Processing Systems*, 35: 27810–27823, 2022.

Zhong, E. D., Bepler, T., Davis, J. H., and Berger, B. Reconstructing continuous distributions of 3d protein structure from cryo-em images. *arXiv preprint arXiv:1909.05215*, 2019.

A. Preliminaries of Quantum Circuit

In quantum computing, information is often carried by qubits over Hilbert space. A pure quantum state consists of one or more qubits and is usually represented by Dirac's notation, which denotes a unit vector $\mathbf{v}$ as a ket $|v\rangle$ and its conjugate transpose $\mathbf{v}^\dagger$ as a bra $\langle v|$. The inner product between $|v\rangle$ and $|u\rangle$ is denoted as $\langle u|v\rangle$, and the outer product is $|u\rangle\langle v|$. The evolution of a quantum state $|v\rangle$ is accomplished by sequentially applying quantum gates on it, i.e. $|v'\rangle = U_K ... U_2 U_1 |v\rangle$, where U_k is the unitary matrix representing the quantum gate and $|v'\rangle$ is the quantum state after evolution. Common single-qubit gates are as follows:

$$\begin{aligned}
&H := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
&X := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, R_X(\theta) := \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \\
&Y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, R_Y(\theta) := \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \\
&Z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, R_Z(\theta) := \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix},
\end{aligned} \tag{A1}$$

where H denotes the Hadamard gate, X, Y, Z denote the Pauli gates, $R_X(\theta), R_Y(\theta), R_Z(\theta)$ denote the rotation gates. A common multi-qubit gate is $CNOT$ gate:

$$CNOT := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{A2}$$

In a quantum circuit, the initial quantum state is generally $|0\rangle^{\otimes N}$, and after applying a sequence of quantum gates, the measurement will be used to convert quantum information into classical information. For instance, we can design quantum measurements to obtain the expectation $\langle v|O|v\rangle$ of the quantum state $|v\rangle$ about an observable O.

B. Additional Explanation on Claim 2

Firstly, we introduce an exponential encoding scheme (Shin et al., 2023), where the specific form of H is as follows:

$$H = \sum_{q=1}^{d} \beta_q Z^{(q)}/2. \tag{A3}$$

In the absence of degeneracy, there can be 2^d distinct eigenvalues. Thus, the spectrum can contain at most $4^d - 2^d + 1$ distinct frequencies (including zero). But if we naively encode all $\beta_q = 1$, then there will be only $d + 1$ distinct eigenvalues and the size of the spectrum $\Omega = \{\lambda_k - \lambda_j\}$ will be limited to $2d + 1$. Therefore, we will now derive how to choose the appropriate β values to maximize the size of the spectrum.

With $H = \beta_1 Z/2$ and upon setting $\beta_1 = 1$, we have $\Omega^{(1)} = \{-1, 0, 1\}$ as the frequency spectrum. If we append one more encoding gate to H,we have $\Omega^{(2)} = \{-\beta_2 - 1, -\beta_2, -\beta_2 + 1, -1, 0, 1, \beta_2 - 1, \beta_2, \beta_2 + 1\}$. More generally, let $\Omega^{(k)}$ as the k-th frequency spectrum as a result of using k encoding gates. Then $\Omega^{(k)}$ has all elements from $\Omega^{(k-1)}$, along with the new ones generated by adding $\pm\beta_k$ to all elements of $\Omega^{(k-1)}$:

$$\Omega^{(k)} = \left\{\Omega^{(k-1)} - \beta_k, \Omega^{(k-1)}, \Omega^{(k-1)} + \beta_k\right\}. \tag{A4}$$

As $\Omega^{(k)}$ is always symmetric about zero, to generate a maximally non-degenerate integer-valued frequency spectrum, the inequality

$$\max\left\{\alpha \in \Omega^{(k-1)}\right\} < \beta_k - \max\left\{\alpha \in \Omega^{(k-1)}\right\} \tag{A5}$$

is to be satisfied, where we can easily deduce that $\max\left\{\alpha \in \Omega^{(k-1)}\right\} = \sum_{j=1}^{k-1} \beta_j$. To construct a dense integer-valued spectrum starting from zero, the following recursive equation

$$2\sum_{j=1}^{k-1} \beta_j + 1 = \beta_k \tag{A6}$$

must be satisfied. When $\beta_1 = 1$, we have $\beta_k = 3^{k-1}$, which is the exponential encoding scheme. We can also construct a non-dense spectrum by choosing large spacings. By applying this scheme, the size of the spectrum is extended from $2d + 1$ to 3^d.

Next, we analyze the Linear layer before data re-uploading circuit. If we ignore the bias term, the Linear layer essentially maps the input x to $(w_1 x, \ldots, w_d x)$. Considering the form of the encoding gate as $S(x) = \otimes_{q=1}^{d} \mathrm{e}^{-\mathrm{i}(w_q x)Z/2} = \mathrm{e}^{-\mathrm{i}x(w_q Z/2)}$, it can be concluded that the Linear layer can be equivalent to an encoding scheme. This encoding scheme is learned through training, and the upper bound is to extend the size of the spectrum to 3^d just as the exponential encoding scheme does. In addition, the Linear layer encoding scheme has the advantage of not being restricted to integer spectrum and the sparsity of the spectrum is adaptive, which makes QIREN more flexible in fitting frequencies. The analyses we have discussed so far have been based on a single variable, but they can be naturally extended to multiple variables, which will not be further elaborated here.

C. Additional Implementation Details

C.1. Signal Representation

We use MSE as the loss function and use Adam optimizers with the parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 1e-8$. The models are trained for 600 epochs for Astronaut, 300 epochs for Camera, Coffee or sound.

ReLU-based MLP includes one input linear layer, one output linear layer, and six hidden linear layers with a hidden dimension of 10. Except for the output layer, the output of each linear layer will pass through a BatchNorm layer and then the activation function.

Tanh-based MLP is set up the same way as the ReLU-based MLP.

ReLU-based MLP with Random Fourier Features is set up the same way as the ReLU-based MLP, except that the input layer is replaced with an RFF layer which consists of a non-trainable random parameter matrix followed by sine and cosine activation functions.

SIREN is set up the same way as the ReLU-based MLP, but instead of using BatchNorm layers, it employs a special initialization method and hyperparameter tuning to regulate the data distribution.

QIREN includes three Hybrid layers and one output linear layer. Each Hybrid Layer consists of a linear layer with a hidden dimension of 8, a BatchNorm layer and a quantum circuit with 8 qubits.

C.2. Image Generation

We build upon the StyleGAN2 framework (Karras et al., 2020), replacing a convolutional generator with our INR-based one as illustrated in Figure A1. For several classical INR-based generators, we directly generated all their parameters using a hypernetwork. However, for QIREN, based on our experiments, we found that the hypernetwork is unable to comprehend the quantum components, leading to poor performance when directly generating parameters. Therefore, we added a ReLU-based MLP after QIREN, and the hypernetwork only generates parameters for this ReLU-based MLP. In this setup, QIREN plays a role similar to that of RFF. All generators include a hypernetwork with eight hidden layers with a hidden dimension of 256. We use non-saturating logistic loss for training:

$$\begin{aligned} Loss_G &= \tfrac{1}{m}\textstyle\sum_{i=1}^{m} -\log\left(D\left(G\left(z^{(i)}\right)\right)\right), \\ Loss_D &= \tfrac{1}{m}\textstyle\sum_{i=1}^{m} \left[\log D\left(\mathbf{x}_i\right) + \log\left(1 - D\left(G\left(\mathbf{z}_i\right)\right)\right)\right], \end{aligned} \tag{A7}$$

and use Adam optimizers with the parameters $\beta_1 = 0.0$, $\beta_2 = 0.98$ and $\epsilon = 1e-8$. All models are trained for 6048 kimg. Additionally, we utilize a trick named factorized multiplicative modulation layer, which was proposed in previous works (Skorokhodov et al., 2021), to alleviate the parameter burden of the hypernetworks. For more specific details about the architecture of the INR-based generators, please refer to the code.

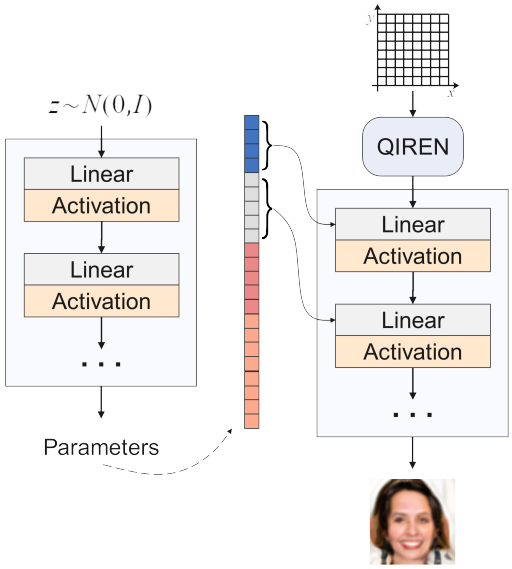


Figure A1. QIREN-based generator.

Figure A2. Results of image generation.

D. Results on Image Generation

In Figure A2, we present the 32×32 pixel images generated by the QIREN-based generator. Next, we will demonstrate two exciting properties that set the QIREN-based generator apart from traditional convolutional generators: out-of-the-box superresolution and meaningful image-space interpolation.

Our QIREN-based generator is able to produce images of higher resolution than it was trained on. For this, we evaluate our model on a denser coordinates grid. Specifically, we input coordinate grids of size 64×64 and 128×128 separately into the QIREN-based generator, and the results are shown in Figure A3.

As is well known, image-space interpolation often exhibits poor performance. However, in Figure A4, we demonstrate that the images generated by the QIREN-based Generator can be interpolated reasonably.

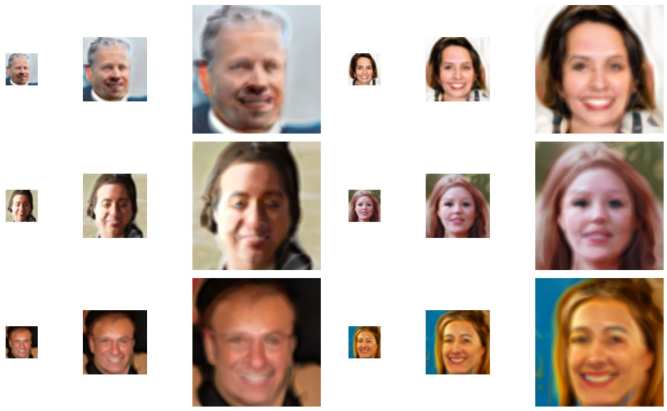


Figure A3. Superresolution of generated images.

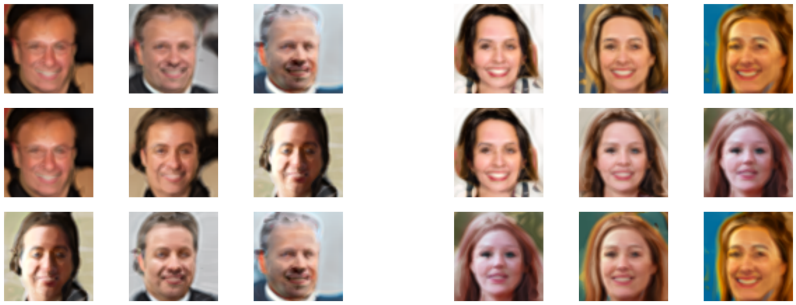


Figure A4. The intermediate image is the result of interpolating between two images on the left and right. To interpolate between $G(\theta_1)$ and $G(\theta_2)$, we compute interpolation parameters $\theta = \alpha\theta_1 + (1-\alpha)\theta_2$ and evaluate $G(\theta)$ for the provided coordinates grid. The interpolated images can also achieve superresolution.

E. Ablation Analysis

E.1. Signal Representation

We conducted four ablation experiments on the sound representation task and presented the results in Figure A5. First, we ablate the importance of the BatchNorm layer. The loss function value of QIREN with BatchNorm consistently remained lower than the one without BatchNorm. This finding demonstrates that the role of the activation function has been taken over by the quantum circuit. As a result, the BatchNorm layer, which is typically placed before the activation function in classical models to deal with vanishing gradient problems, can also be applied to our quantum model to effectively promote fast convergence.

Next, we investigated the influence of the number of data re-uploading L, also known as circuit depth, on the representation

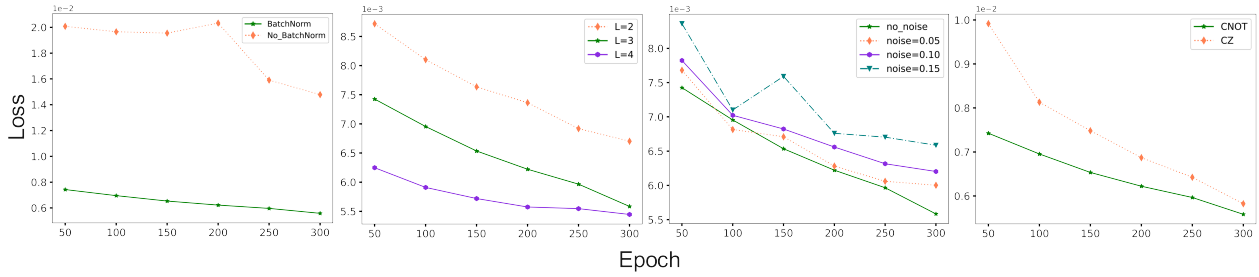


Figure A5. Loss functions of ablation experiments on the sound representation task.

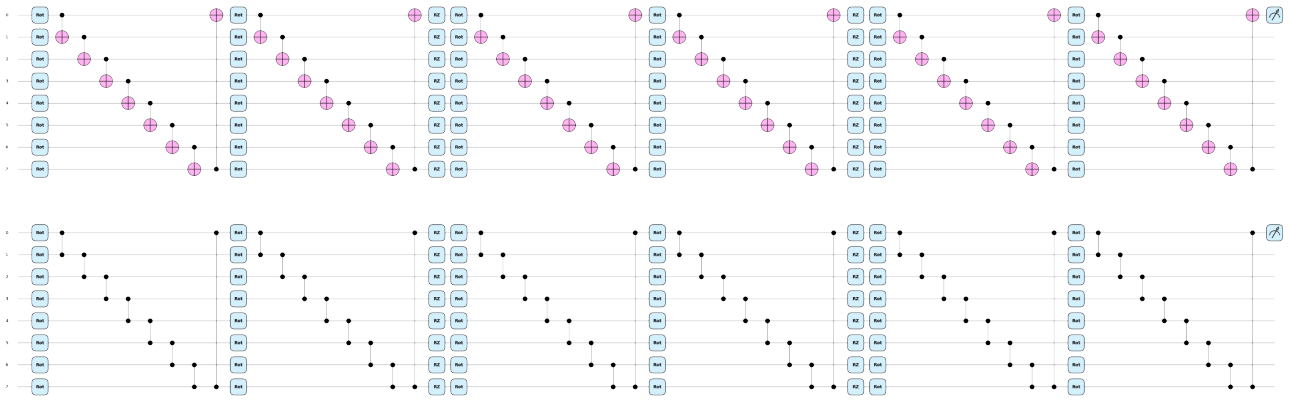


Figure A6. Circuit diagrams for the two ansatz.

capacity. In the original version of QIREN, L was set to 3. It can be observed that within our experimental range, as L increased, the representation capability of QIREN was strengthened.

Considering the real-world scenario of running on a quantum computer, we investigated the impact of noise on QIREN. We introduced noise by applying an $R_X(\theta)$ on each qubit before the measurement, where $\theta \sim \mathrm{U}(0, \text{noise})$ and noise $= 0.05, 0.10, 0.15$. It can be observed that noise does impact the performance of QIREN, but even with the presence of noise, QIREN demonstrates remarkable signal representation capability comparable to SIREN and ReLU-based MLP with RFF.

Finally, we investigated the impact of the quantum circuit ansatz on performance. In the vanilla QIREN, we used $CNOT$ gates to provide quantum entanglement. We attempted to replace them with CZ gates, and both ansatz showed similar performance. The circuit diagrams for both ansatz are shown in Figure A6.

We performed an ablation analysis of the Linear layer on the image representation task, and as can be seen in Table A1, the model fits the image poorly after removing the Linear layer. This is because the Linear layer serves to extend and adjust the spectrum, allowing QIREN to capture most of the important frequency bands of the signal with a small amount of quantum resources, as analyzed in Claim 2 and Appendix B. However, a limitation of this ablation analysis is that after removing the Linear layer, QIREN converts to a pure quantum circuit. The current limitations of hardware devices restrict us to training only a lightweight quantum circuit, with the number of parameters involved not being on the same order of magnitude as in QIREN. This is also the reason why we adopted a hybrid quantum neural network: to obtain quantum advantages in real-world tasks with limited quantum resources. Pure quantum circuits are often limited to simple synthetic tasks.

Method	Astronaut	Camera	Coffee	#params
Prue quantum circuit	71.1	51.9	73.2	72
QIREN	4.0	1.1	1.5	657

Table A1. MSE ($\times 10^{-3}$) of different models on the image representation task.

E.2. Image Generation

We conducted ablation experiments on the FFHQ dataset, and the results are presented in Table A2. First, we employed the factorized multiplicative modulation technique (Skorokhodov et al., 2021), which significantly reduced the number of model parameters while ensuring high-quality generated images. Next, we expanded the encoding dimension from 8 to 32 by increasing the number of quantum circuits. This expansion resulted in a substantial improvement in generation quality. We also experimented with increasing the number of data re-uploading and found that it enhanced the performance of the model. Finally, we tested the performance of the mainstream generator model, StyleGAN2 (Karras et al., 2020). In order to make a fair comparison, we adjusted the hyperparameters of StyleGAN2 to ensure that the total number of parameters in both models was approximately the same. While StyleGAN2 outperforms our QIREN-based generator in terms of image

generation quality, it is important to note that superior image generation performance is not the primary advantage of the INR-based generator model. As shown in several studies (Skorokhodov et al., 2021; Dupont et al., 2021b) and Appendix D, their strength lies in unique properties that mainstream generator models do not possess.

Method	FID	#params
QIREN	11.77	1.45M
+ Factorized Multiplicative Modulation	11.53	1.13M
+ Encoding Dimension	7.67	1.19M
+ Number of Data Re-uploading	6.88	1.19M
StyleGAN2	5.14	1.38M

Table A2. Results of ablation experiments on the FFHQ dataset.