

Exploring Trainability of Quantum Fourier Models via Data-Tailored Encoding Schemes [4]

Felix Paul*, Björn Minneker, Peter Jung

Introduction

In Quantum Machine Learning (QML), a lot of research focus has been directed towards Parameterized Quantum Circuits (PQCs) - quantum circuits containing trainable parameters that can be optimized using classical or quantum-mechanical optimization techniques. These circuits encode classical data into quantum states through unitary operations, similar to how classical neural networks process information.

A remarkable discovery from recent years revealed that quantum models employing data re-uploading schemes can be mathematically represented as truncated Fourier series [1]. This representation shows that

- The model's frequency spectrum depends exclusively on the eigenvalues λ of the data-encoding Hamiltonians
- The Fourier coefficients are determined by the trainable parameters and the observable
- The model's expressivity is governed by both spectrum size and coefficient flexibility

Understanding the properties of such Quantum Fourier Models (QFMs) is essential for insights into their capabilities and limitations. Here, we look into the trainability of QFMs for different encoding strategies, focusing on the redundancy of certain frequencies and how it affects the training process.

Framework

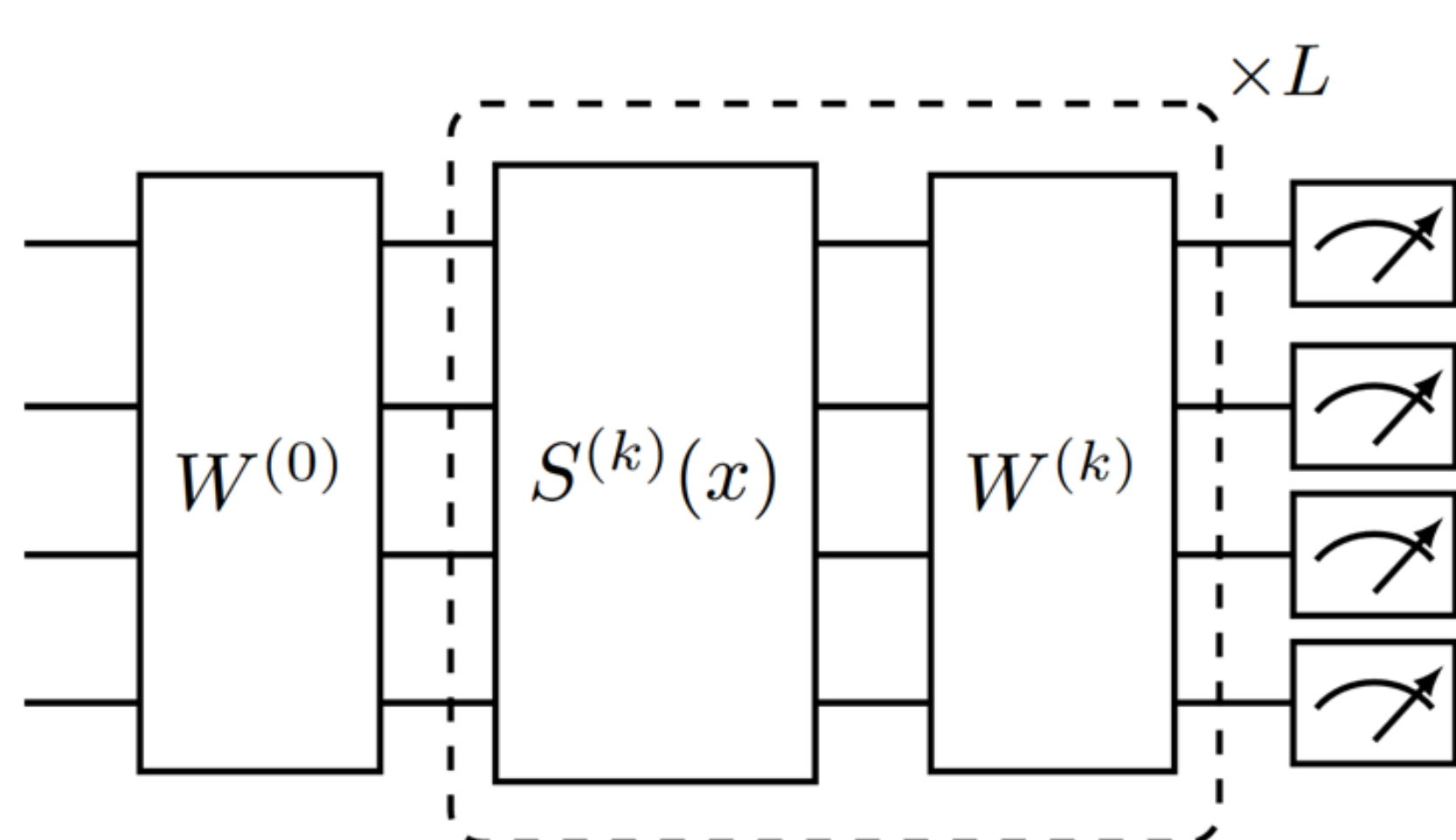
Objective functions of QML models is expectation w.r.t. some observable

$$f_{\theta}(x) = \langle 0 | U^{\dagger}(x; \theta) \mathcal{M} U(x; \theta) | 0 \rangle$$

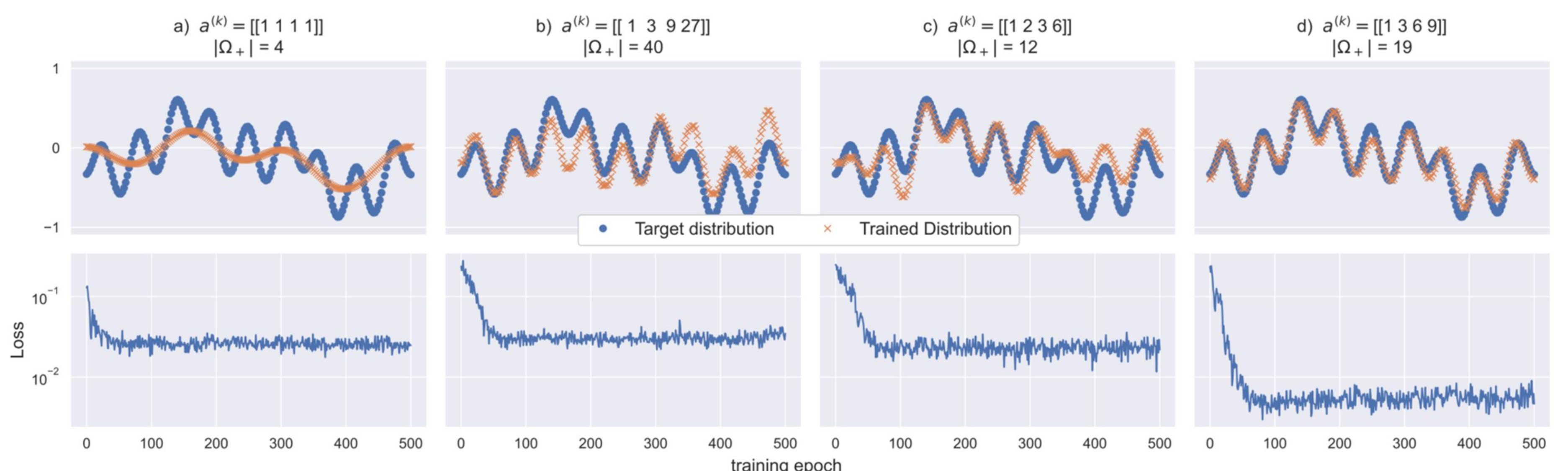
Constructing the unitary evolution $U(x; \theta)$ as depicted below, the objective function can be expressed as **truncated Fourier series** [1]

$$f_{\theta}(x) = \sum_{\omega \in \Omega} c_{\omega}(\theta) e^{-i\omega x}$$

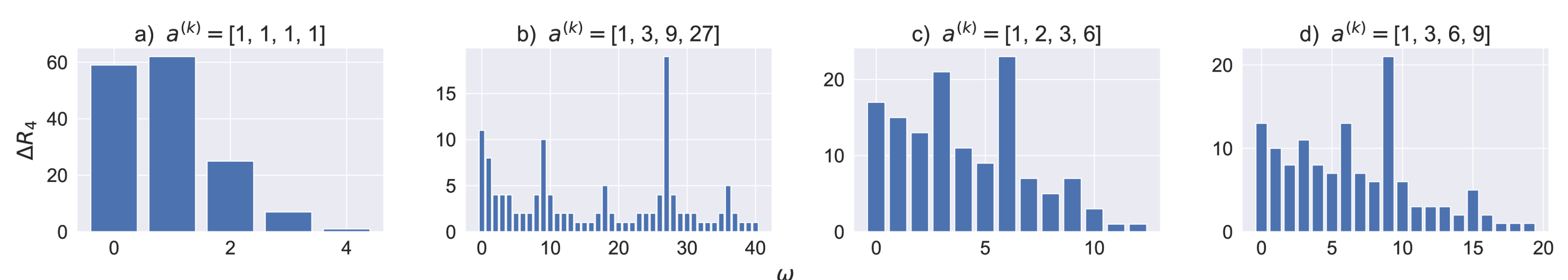
with the spectrum depending on the number of data re-uploads L and qubits n .



Circuit diagram of re-uploading scheme. $W^{(k)}$ are trainable blocks, $S^{(k)}$ are (layer-dependent) data-encoding blocks



Comparison of representative training performances and trained distributions for different data re-uploading schemes, namely a) constant, b) exponential, c) alternative and d) data-tailored encoding. The scale for the loss is logarithmic. For each setting, the (positive) spectrum size $|\Omega_+|$ is noted in the top.



Partial redundancy difference distributions $\Delta R_4(\omega)$ for the considered data re-uploading schemes, namely a) constant, b) exponential, c) alternative d) data-tailored encoding

Introducing a **layer-dependent encoding** affects the model's available spectrum

$$S(x) \mapsto S^{(k)}(x) = \exp(-ixa^{(k)}Z/2)$$

Different choices of the scaling coefficients $a^{(k)}$ changes the **spectrum size** ($n = 1$ examples):

Constant encoding:

$$a^{(k)} = 1 \quad \forall k \Rightarrow |\Omega_+| = L$$

Exponential encoding [2]:

$$a^{(k)} = 3^{k-1} \Rightarrow |\Omega_+| = (3^L - 1)/2$$

Frequency Redundancy $|R(\omega)|$

How often a frequency comes up during the calculation of the spectrum

$$\Omega = \left\{ \sum_{i=1}^L \lambda_{j_i} - \sum_{i=1}^L \lambda_{k_i} \mid \mathbf{j}, \mathbf{k} \in [d]^L \right\}$$

Example: constant encoding, $n = 1, L = 2$

$$\Omega = \{-2, -1, 0, 1, 2\}$$

With the following redundancies

ω	0	± 1	± 2
$ R(\omega) $	6	4	1

The **variance** of the Fourier coefficients w.r.t. the optimization parameters can be expressed via the (adapted) **frequency redundancy** (R_k^L refers to redundancy obtained by a circuit from the k -th layer onward) [3]:

$$\text{Var}_{\theta}[c_{\omega}(\theta)] = C \left[\underbrace{|R_1^L(\omega)| - |R_2^L(\omega)|}_{\Delta R_L(\omega)} + \sum_{j=3}^L 3^{j-2} |R_j^L(\omega)| \right]$$

Setup & Results

Method: Simulation in PennyLane:

- $n = 1$ qubit
- $L = 4$ layers
- Variational blocks consist of 3 independent U_3 -gates + Entangler

Goal: Approximate real-valued Fourier series with frequencies $\{0, 3, 6, 9\}$ using different encodings.

Question: How does a **higher variance** for important coefficients **influence** the **training**?

Observation:

- **Higher variance** of important frequency-coefficients $\Rightarrow$ **Better Training** performance
- Too **small spectrum** fails to catch important components
- Too **large spectrum** results in **bad training**

Next step: Understand **structure of variance** in order to **adapt encoding for better training prerequisites**.

References

- [1] M. Schuld et al. "Effect of data encoding on the expressive power of variational quantum-machine-learning models" Phys. Rev. A, vol. 103, p. 032430, 2021
- [2] S. Shin et al. "Exponential data encoding for quantum supervised learning" Phys. Rev. A, vol. 107, no. 1, p. 012422, 2023
- [3] H. Mhiri et al. "Constrained and vanishing expressivity of quantum Fourier models", Quantum, vol. 9, p. 1847, 2025
- [4] F. Paul et al. "Exploring Trainability of Quantum Fourier Models via Data-Tailored Encoding Schemes" to appear in Proc. of the IEEE ICASSP 2026 Satellite Workshop on QML

