Out of Tune: Demystifying Noise-Effects on Quantum Fourier Models

Maja Franz[*]
Technical University of
Applied Sciences Regensburg
Regensburg, Germany
maja.franz@othr.de

Melvin Strobl[*]
Karlsruhe Institute of Technology
Karlsruhe, Germany
melvin.strobl@kit.edu

Leonid Chaichenets
Karlsruhe Institute of Technology
Karlsruhe, Germany
leonid@chaichenets.me

Eileen Kuehn
Karlsruhe Institute of Technology
Karlsruhe, Germany
eileen.kuehn@kit.edu

Achim Streit
Karlsruhe Institute of Technology
Karlsruhe, Germany
achim.streit@kit.edu

Wolfgang Mauerer
Technical University of
Applied Sciences Regensburg
Siemens AG, Technology
Regensburg/Munich, Germany
wolfgang.mauerer@othr.de

Abstract—The field of variational quantum algorithms, in particular Quantum Machine Learning (QML), produced numerous theoretical and empirical insights in recent years. As Variational Quantum Circuits (VQCs) can be represented by Fourier series that contain an exponentially large spectrum in the number of input features, hope for quantum advantage remains. Nevertheless, properties of Quantum Fourier Models (QFMs) are not yet fully understood, in particular how they could potentially outperform classical alternatives. Viewing VQCs with Fourier lenses opens up possibilities to analyse which classes of functions can be tackled by variational algorithms such as QML, while also illuminating and quantifying remaining constraints and challenges.

Considering that noise and imperfections remain dominant factors in the development trajectory from noisy intermediate-scale to fault-tolerant quantum computers, the aim of this work is to shed light on key properties of QFMs when exposed to noise. In particular, we systematically analyse the effect of noise on the Fourier spectrum, expressibility and entangling capability of QFMs by conducting large-scale numerical simulations of quantum systems. This may help to better utilise hardware resources, and guide the construction of tailored error correction schemes.

We find that decoherent noise exerts a uniform deleterious effect on all the tested ansätze, manifesting in the vanishing of Fourier coefficients, expressibility and entangling capability. We note however, that the detrimental influence of noise is less pronounced in some ansätze than in others, suggesting that these might possess greater resilience to noise.

Index Terms—Quantum Machine Learning, Fourier Analysis, Expressibility, Entanglement

I. INTRODUCTION

Given the remarkable advancements in machine learning, significant optimism has been directed towards Quantum Machine Learning (QML). However, such enthusiasm has frequently been tempered by the prevailing limitations of contemporary quantum hardware [1], together with algorithmic shortcomings. On the other hand, exact capabilities of QML are not yet fully understood, yet considerable opportunities to achieve advantage over classical approaches remain. As the concept of QML (see

the introductory review by Schuld *et al.* [2]) has received considerable attention in the literature during the last years, crucial limitations and trade-offs have been identified [3], especially regarding trainability challenges [4] and model complexity constraints [5]. Even despite technology progresses from Noisy Intermediate Scale Quantum (NISQ) devices [6] towards Fault-Tolerant Quantum Computing (FTQC), noise and imperfections will continue to influence Variational Quantum Circuit (VQC) performance and properties in the foreseeable future, particularly when considering the best possible use of available hardware resources [7].

In this work, we seek to shed light on key properties of VQCs in QML, with a particular focus on providing numerical insights into the impact of noise. Our contributions are illustrated in Fig. 1. We emphasise that our results are hardware agnostic, that is, they are not only applicable to error-corrected architectures, but also to intermediate Quantum Processing Units (QPUs) and algorithms (*e.g.* as in Refs.[8]–[14]), until full FTQC are available.

Common metrics used to assess VQCs include expressibility and entangling capability [16], which generally determine how effectively a VQC can explore the Hilbert space. This exploration is crucial because it it indicates the capacity of a VQC to learn various functions, with the solutions to these functions residing in this Hilbert space. However, a highly expressive VQC can face challenges such as the Barren-Plateau (BP) problem [4], where the optimisation landscape becomes difficult to navigate. Therefore, a successful VQC must strike a balance: it needs to be sufficiently expressive so that the portion of the Hilbert space it accesses contains potential solutions, while maintaining a manageable optimisation landscapes. Another way to gauge this property is by examining the *expressiveness*, that is the number and values of the VQC's Fourier coefficients [15], which dictate the types of functions the VQC can learn. Conversely, if an ansatz lacks sufficient expressiveness, the "trainable Hilbert space" it accesses is highly unlikely to contain viable solutions.

---

[*]Equal contribution
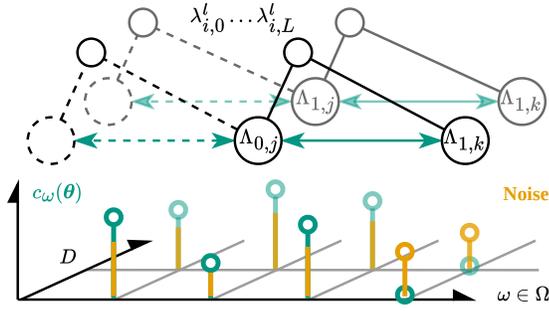
(a) Quantum spectrum tree [15] where Λ is constituted by the sum of the individual eigenvalues λ of each generator in the D-dimensional input encoding. The spectrum Ω is determined by the differences between each Λ where each coefficient $c_\omega(\theta)$ depends on a parameter vector θ. Noisy coefficients generally have a reduced magnitude, which can lead to non-trainable frequency components.
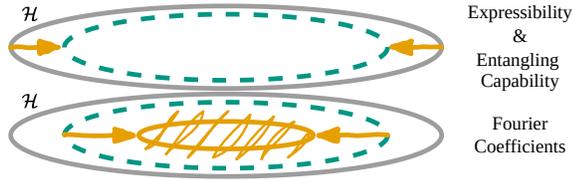
(b) Implications of the above considerations on Hilbert space level (solid grey), denoted by $\mathcal{H}$: Expressibility and entangling capability limits depend on the level of noise (arrows) and the accessible space (dashed) of the model. Additionally, noise further reduces this space when viewed through the Fourier lens, as it reduces the ability to change frequency components while adding untrainable frequency artifacts (hatched outside of solid orange).

Fig. 1: Overview of factors (and their interplay) that influence various properties of QML as studied in this work.

II. QUANTUM FOURIER MODELS

We consider an n-qubit quantum circuit whose expectation value, measured on each qubit (n-local) using the observable $\mathcal{M}$, is given by

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = \langle 0|^{\otimes n} U^\dagger(\boldsymbol{x}, \boldsymbol{\theta}) \mathcal{M} U(\boldsymbol{x}, \boldsymbol{\theta})|0\rangle^{\otimes n}, \quad (1)$$

which is parametrised by the D-dimensional input $\boldsymbol{x} = (x_1, \ldots, x_D) \in \mathbb{R}^D$ and the trainable parameter vector $\boldsymbol{\theta} =$

$(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{L+1}) \in [\mathbb{R}^K]^{L+1}$ with L layers of K parameters per layer[1].

We frame the unitaries $U(\boldsymbol{x}, \boldsymbol{\theta})$ as multiple layers ($\ell \in [1, L]$) of consecutive encoding $S^{(\ell)}(\boldsymbol{x})$ and trainable unitaries $W^{(\ell)} := W^{(\ell)}(\boldsymbol{\theta}_\ell)$ as follows:

$$U(\boldsymbol{x}) = W^{(L+1)} S^{(L)}(\boldsymbol{x}) W^{(L)} \cdots W^{(2)} S^{(1)}(\boldsymbol{x}) W^{(1)} \quad (2)$$

where an additional last trainable layer $W^{(L+1)}$ is added. To allow for an arbitrary circuit structure and placement of unitary and encoding gates, we construct $S^{(\ell)}(\boldsymbol{x})$ to comprise combinations of D unitaries $S_i^{(\ell)}(\boldsymbol{x})$, each of which encodes one single input feature x_i:

$$S_i^{(\ell)}(\boldsymbol{x}) = \exp\left(-\imath \boldsymbol{e}_i^T \boldsymbol{x} G_i^{(\ell)}\right) = \exp\left(-\imath x_i G_i^{(\ell)}\right) \quad (3)$$

with Hermitian generator $G_i^{(\ell)}$ and standard basis vector $\boldsymbol{e}_i$.

As shown in the seminal works of [17], [18], such an architecture allows rewriting Eq. 1 as a partial Fourier series

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{\theta}) &= \sum_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} c_{\boldsymbol{\omega}}(\boldsymbol{\theta}) e^{\imath \boldsymbol{\omega}^T \boldsymbol{x}} \\ &= \sum_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} |c_{\boldsymbol{\omega}}(\boldsymbol{\theta})| \left(\cos(\varphi(\boldsymbol{\theta})) \cos(\boldsymbol{\omega}^T \boldsymbol{x}) \right. \\ &\quad \left. - \sin(\varphi(\boldsymbol{\theta})) \sin(\boldsymbol{\omega}^T \boldsymbol{x})\right), \end{aligned} \quad (4)$$

where $\boldsymbol{\Omega}$ contains the frequencies resulting from the eigenvalues $\{\lambda_{i,j}^{(\ell)}\}$ of each generating encoding Hamiltonian $G_i^{(\ell)}$, and $\{c_{\boldsymbol{\omega}}(\boldsymbol{\theta})\} = \{|c_{\boldsymbol{\omega}}(\boldsymbol{\theta})|e^{\imath\varphi(\boldsymbol{\theta})}\}$ are the corresponding complex Fourier coefficients.

The resulting set of frequencies is

$$\boldsymbol{\Omega} = \left\{ \times_{i=1}^D \{\Lambda_{i,\boldsymbol{j}} - \Lambda_{i,\boldsymbol{k}}\} \mid \boldsymbol{j}, \boldsymbol{k} \in \prod_{\ell=1}^L [\![1, d_\ell]\!] \right\}, \quad (5)$$

where $\Lambda_{\boldsymbol{j}}$ present the sum of all d_ℓ eigenvalues of the ℓ-th generator grouped across all layers, given by

$$\Lambda_{i,\boldsymbol{j}} = \lambda_{i,1}^{(\ell)} + \cdots + \lambda_{i,L}^{(\ell)}. \quad (6)$$

Here, Eq. 5 shows a finding from Ref. [19] where it was shown that the set of frequencies is constituted by the Cartesian product over all the individual sets that stem from each input encoding. Note that while there are no theoretical constraints on the encoding strategy of each input, the practically useable set of frequencies only increases if each data encoding is orthogonal to the others.

Another descriptive metric of the Fourier spectrum that also provides an upper bound for the expressibility was introduced in Ref. [15] in form of the frequency redundancy which is defined by the size of the redundancy generator

$$R(\boldsymbol{\omega}) =$$
$$\left\{ (\boldsymbol{j}, \boldsymbol{k}) \mid \times_{i=1}^D \{\Lambda_{i,\boldsymbol{j}} - \Lambda_{i,\boldsymbol{k}}\} = \boldsymbol{\omega} \wedge \boldsymbol{j}, \boldsymbol{k} \in \prod_{\ell=1}^L [\![1, d_\ell]\!] \right\}. \quad (7)$$

In addition to that, the entangling capability is a metric to judge the amount of entanglement a VQC could create by adjusting the trainable parameters, while being restricted by its structure. This metric can be viewed as the "quantumness" a particular circuit can achieve and is therefore crucial to set it apart from classical surrogates.

Before we give an extensive overview of related work in Sec. III, we first introduce the concept behind Quantum Fourier Models (QFMs) in Sec. II. We then present our methods in Sec. IV where we introduce the noise models used in this work and how we measure the coefficients, entangling capability and expressibility. In Sec. V we present our numerical results, provide a discussion in Sec. VI and finally conclude in Sec. VII.

---

[1] Without loss of generality, we choose the same number of parametrised gates for each layer.

By increasing the number of layers, the number of frequencies $|\mathbf{\Omega}|$ increases linearly as the number of redundancies $\sum_{\boldsymbol{\omega} \in \mathbf{\Omega}} |R(\boldsymbol{\omega})|$ increases exponentially considering single qubit Pauli-encodings. Notably, it was also shown in Ref. [17], that the exact same result can be achieved by increasing the number of qubits instead. It is well known that the increasing number of frequencies stem from the gaps between the eigenvalues of the encoding generating Hamiltonians, which can visually be represented by a tree structure [15], as shown in Fig. 1a.

The coefficients $c_{\boldsymbol{\omega}}(\boldsymbol{\theta})$, parametrised by the trainable parameters in the VQC, are solely determined by the circuit architecture, although they are partially affected by the encoding as shown in Ref. [20]. The way parameters act on the coefficients is not trivial with an analytical description of this relation as published recently in Refs. [21], [22].

III. RELATED WORK

Various aspects can be considered, when working with QFMs, which range from the *Fourier analysis*, over the relation to *trainability*, *dequantisation*, *expressibility* and *entanglement* to the influence of *noise*. In this section, we review the main related work focusing on these topics.

A. Fourier Analysis

The seminal work [18] and following study in Ref. [17] sparked a whole subfield of research within QML. With the former paper introducing the "Data-Reuploading" technique that enables a non-linear transformation between input and output of a VQC with only unitary transformations, the latter extended this work by deriving first formulations on how the spectrum of a QFM changes with the input encoding. Ref. [19] then further extended this to a formulation for multidimensional input features where they found that for some ansätze, the spectrum size grows faster than the available degrees in Hilbert space. While aforementioned work considered a fixed encoding strategy, which results in an evenly spaced spectrum, Jaderberg *et al.* [23] found that by adding a trainable parameter to the input, the distances between the frequencies can be changed. The expressibility of QFM and the redundancies of eigenvalues that effectively determine the resulting spectrum was linked in [15], which also demonstrated that the variance of a frequencies coefficients is linearly dependent on the number of its redundancies. Nemkov *et al.* [21] introduced an analytical description between Fourier coefficients and trainable parameters. Wiedmann *et al.* [22] use this analytical description to argue that certain coefficients, and therefore also frequencies, can vanish dependent on the trainable parameters, which leads to a reduction of the hypothetical spectrum.

B. Trainability

While there has been a significant effort to build mathematical formulations and find analytical relations around QFM, the trainability of such models has to be regarded as well. One of the fundamental problems is called BP phenomenon, introduced in Ref. [24]. Generally it describes the exponential decay of gradients in a VQC caused by (1) the circuit expressibility,

(2) degree of entanglement, (3) locality of the observable [4], [25] and (4) noise [4], [26].

One could argue, that reducing mentioned sources would help, but it ultimately boils down to the fact that if a circuit does not exhibit a BP, it is classically simulable, ruining any potential advantage [27]. However, while this holds true for a majority of use cases, it is not generally applicable. Also, Ref. [20] showed that the encoding strategy plays a significant role in the trainability of QML models.

C. Dequantisation

To assess whether QML offers advantages over classical Machine Learning (ML), one approach is dequantisation, where classical methods replicate or surpass QML's performance. Schreiber *et al.* [28] demonstrated that such a classical surrogate model can outperform QFMs for small problem instances, but this becomes intractable as the number of Fourier coefficients grows exponentially with input features. To mitigate this, Fontana *et al.* [29] and Landmann *et al.* [30] proposed approximating VQC outcomes by respectively trimming frequencies or using Random Fourier Features (RFFs). Sweke *et al.* [31] established that efficient dequantisation for regression is possible with RFFs if the spectrum of the QFM is polynomially concentrated, requiring a polynomial number of frequencies.

All insights in the above mentioned references, result from looking at the VQC, optimisation algorithm and training data as a whole, which is reasonable, given that this is the most likely way that QML approaches problems. However, this holistic way of considering QML makes it hard to validate the actual VQC structure that are at the core.

D. Expressibility and Entanglement

Characterising the properties of VQCs to identify suitable ansätze for variational quantum algorithm is a common approach, while the implications are often left as open questions in the field. In this context, Sim *et al.* [16] conducted an extensive study to measure the expressibility (cf. Sec. IV-C) and entangling capability (cf. Sec. IV-D) of different ansätze. expressibility, which indicates how effectively an ansatz can explore the Hilbert space, has been shown to be related to the redundancies in the Fourier spectrum [15].

The role of entanglement in QML and other variational algorithms remains far from fully understood. When dealing with input data that is inherently quantum in nature, entanglement is often considered as the key resource for successful learning [32]. Wang *et al.* [33] demonstrated that increasing entangling capability up to a certain threshold can lead to improved model performance. However, when the input data is classical, empirical studies, such as those in Refs. [34], [35], have shown that a low-entanglement circuits can perform similarly well, or even better than highly entangled ones on specific learning tasks. Furthermore, Joch *et al.* [36] highlighted the potential downside of excessive entanglement in learning scenarios, suggesting that too high an entangling capability can lead to the BP phenomenon. This indicates a critical threshold beyond

which entanglement may no longer be beneficial for learning tasks. To our knowledge, there is currently no established connection in previous work between the entangling capability of a circuit and its Fourier spectrum, to which we contribute in this work.

E. Noise

Given that fault-tolerance is not yet fully in reach, the influence of noise on QFM remains an open question. Fontana *et al.* [37] measure the effect of noise on the Fourier spectrum to make suggestions for noise mitigation and diagnostics. By cutting or filtering of certain noise-induced coefficients or frequencies, noiseless landscapes can be approximately reconstructed. Apart from Ref. [37], which focuses on the Quantum Approximate Optimisation Algorithm (QAOA) and Variational Quantum Eigensolver (VQE), we are not aware of other publications, investigating the direct influence of noise on the Fourier spectrum and other properties of VQCs, especially in the QML domain. With this article, we aim to contribute to filling this research gap.

IV. METHOD

A. Noise

In our work we investigate decoherent types of noise that are of stochastic nature and not necessarily follow unitary evolutions. As introduced in Ref. [6], we investigate three main decoherent noise channel categories, namely (1) damping noise (*i.e.* environmental effects, decoherence) encompassing Amplitude Damping (AD) and Phase Damping (PD), (2) decoherent gate errors, including Bit Flip (BF), Phase Flip (PF) and Depolarisation (DP) noise, and (3) State Preparation and Measurement (SPAM) errors. Additionally we consider a Coherent Gate Error (CGE).

The following paragraphs describe the individual noise models, inspired by Ref. [6]. For the decoherent noise types, we utilise the Kraus formalism with each Kraus operator K_i acting on the density matrix ρ as

$$\rho \to \sum_i K_i \rho K_i^\dagger. \tag{8}$$

1) Decoherent Gate Errors: The general idea of decoherent gate error channels is to apply some combination of Pauli-operations $\{I, X, Y, Z\}$ with a certain probability after each (noiseless) quantum gate. As a proof-of-concept, we only consider one-qubit channels in our experiments, although multi-qubit Pauli-channels are possible [6].

a) Bit Flip (BF): The BF error is modelled by the Kraus operators

$$K_0 = \sqrt{1 - p_{bf}}I, \qquad K_1 = \sqrt{p_{bf}}X, \tag{9}$$

with the probability of a BF error being p_{bf}.

b) Phase Flip (PF): Similarly, the PF error is modelled by the Kraus operators

$$K_0 = \sqrt{1 - p_{pf}}I, \qquad K_1 = \sqrt{p_{pf}}Z, \tag{10}$$

with the probability p_{pf} of a PF error.

c) Depolarisation (DP): The DP error creates a fully mixed state with probability p_{dp} and is modelled by

$$K_0 = \sqrt{1 - p_{dp}}I, \qquad K_1 = \sqrt{p_{dp}/3}X,$$
$$K_2 = \sqrt{p_{dp}/3}Y, \qquad K_3 = \sqrt{p_{dp}/3}Z. \tag{11}$$

2) Damping Errors: The group of errors to which we refer as *damping* errors corresponds to noise effects that occur over time when a quantum system interacts with the environment. Notable, the probabilities for damping depend on the relaxation time T_1 and dephasing time T_2 of a quantum system in question. As we are working with no particular QPU, in our experiments we assume fixed probabilities and apply the damping channels at the end of a circuit.

a) Amplitude Damping (AD): AD describes the natural decay of the excited state $|1\rangle$ to the ground state $|0\rangle$ due to energy exchange with the environment. The corresponding Kraus operators are

$$K_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1 - p_{ad}} \end{bmatrix}, \qquad K_1 = \begin{bmatrix} 0 & \sqrt{p_{ad}} \\ 0 & 0 \end{bmatrix}, \tag{12}$$

with AD probability p_{ad}.

b) Phase Damping (PD): Similarly, PD describes the transition of a quantum system towards classical behaviour, defined by the Kraus operators

$$K_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1 - p_{pd}} \end{bmatrix}, \qquad K_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{p_{pd}} \end{bmatrix}, \tag{13}$$

with PD probability p_{pd}.

3) State Preparation and Measurement (SPAM) Errors: State Preparation and Measurement (SPAM) on a real quantum device can be faulty, just like any other operation. Typically on current QPUs, a state is prepared in the all-$|0\rangle$ state, and a measurement returns either $|0\rangle$ or $|1\rangle$ with some probability, depending on the algorithm. However, there is a probability that State-Preparation (SP) fails, or that the measurement returns a $|1\rangle$ instead of a $|0\rangle$ and vice versa. Essentially, SPAM errors can be modelled with BF errors, applied to all qubits.

a) State-Preparation (SP): For SP these BF errors are applied at the beginning of a circuit with probability p_{sp}.

b) Measurement (ME): For the ME the additional BF errors with probability p_{me} are applied at the very end of a circuit.

4) Coherent Gate Error (CGE): Notable, since each operation is applied to the quantum system using an imperfect real QPU, the actual gate operation may deviate from the intended one. This coherent error can usually be mitigated in QML by accordingly adjusting the parameter that parametrises a gate. However, for the sake of completeness we provide this error in form of a Gaussian distribution, from which an error $\epsilon \sim \mathcal{N}(0, p_{cge}^2)$ is drawn randomly for each gate.

B. Coefficients

Based on the definition of a QFM from Eq. 4, we can investigate the coefficients of such a model after parametrisation using the Fast Fourier Transform (FFT). Given the correct number of frequencies, which is in the simplified scenario of Pauli-encoding and a single encoding layer, equal to the

number of qubits n, this transformation yields a numerical estimate of the coefficients. Note that this set of coefficients is a numerical approximation based on the expectation value of the model given a range of input samples $x \in \mathcal{X}$ and a fixed parameter vector $\boldsymbol{\theta}$. As stated in Eq. 4, these coefficients depend on the parameters. For each ansatz we assume that this set of coefficients is characterised by the mean and the standard deviation which is what we evaluate in the following numerical experiments. We also calculate the covariance matrices

$$\mathrm{Cov}_c = \begin{pmatrix} \mathrm{Cov}(\mathrm{Re}(c),\mathrm{Re}(c)) & \mathrm{Cov}(\mathrm{Re}(c),\mathrm{Im}(c)) \\ \mathrm{Cov}(\mathrm{Im}(c),\mathrm{Re}(c)) & \mathrm{Cov}(\mathrm{Im}(c),\mathrm{Im}(c)) \end{pmatrix}. \quad (14)$$

To estimate the model spectrum, we need to sample the parameter space Θ. We estimate the mean value of a coefficient contributing to the frequency $\boldsymbol{\omega}$,

$$\mu_c(\boldsymbol{\omega}) = \frac{1}{|\Theta|} \sum_{\boldsymbol{\theta} \in \Theta} |\mathrm{FFT}_{\mathcal{X}}(f(\cdot, \boldsymbol{\theta}))(\boldsymbol{\omega})|, \quad (15)$$

and the relative standard deviation

$$\sigma_c(\boldsymbol{\omega}) = \frac{1}{|\Theta|} \sum_{\boldsymbol{\theta} \in \Theta} \frac{\sqrt{|\mathrm{FFT}_{\mathcal{X}}(f(\cdot, \boldsymbol{\theta}))(\boldsymbol{\omega}) - \mu_c(\boldsymbol{\omega})|^2}}{\mu_c(\boldsymbol{\omega})}. \quad (16)$$

$\mathrm{FFT}_{\mathcal{X}}$ represents the discrete Fourier transform over $\mathcal{X}$. Analytical coefficients can be obtained by expanding the expectation value using trigonometric polynomials [21], [22]; we use these to cross-validate FFT result.

C. Expressibility

For the expressibility we utilise the Kullback-Leibler (KL) divergence [38] between the distributions obtained by sampling from the Haar integral $\int_{\mathrm{Haar}} (|\psi\rangle\langle\psi|)^{\otimes t} d\psi$ and the model $\int_{\boldsymbol{\theta}} (|\psi_{\boldsymbol{\theta}}\rangle\langle\psi_{\boldsymbol{\theta}}|)^{\otimes t} d\boldsymbol{\theta}$, as introduced in Ref. [16]:

$$D_{\mathrm{KL}}\left(\hat{P}_{\mathrm{Model}}(F; \boldsymbol{\theta}) \| P_{\mathrm{Haar}}(F)\right) \quad (17)$$

Here, the fidelity $F = |\langle\psi_{\boldsymbol{\varphi}} | \psi_{\boldsymbol{\phi}}\rangle|$ is the probability of state overlaps whereas the distributions of state overlaps is then $p(F = |\langle\psi_{\boldsymbol{\varphi}} | \psi_{\boldsymbol{\phi}}\rangle|)$.

This metric yields zero if $\hat{P}_{\mathrm{Model}}(F; \boldsymbol{\theta}) = P_{\mathrm{Haar}}(F)$, meaning the states sampled from the QFM are Haar distributed. For the least expressive case, *i.e.* the idle circuit, the KL divergence becomes $\ln(n_{\mathrm{bins}})$ where n_{bins} describes the number of bins that are used for discretising the probability distribution using a histogram. For the remainder of this work, we refer to the expressibility as the inverse of KL divergence.

D. Entangling Capability

There are different ways to calculate the entangling capability of a VQC, some of which we discuss in this section.

1) Meyer-Wallach (MW) Measure: The MW entangling capability [39], [40] is defined as the trace of the squared partial density matrix ρ_k where k indices the subsystem:

$$Q(|\psi\rangle) = 2\left(1 - 1/n \sum_{k=0}^{n-1} \mathrm{Tr}\left[\rho_k^2\right]\right). \quad (18)$$

This metric has the property that if $\mathrm{Tr}\left[\rho_j^2\right] = 1 \quad \forall j$, implying $Q = 0$, $|\psi\rangle$ is a product state whereas $Q = 1$ iff $\mathrm{Tr}\left[\rho_k^2\right] = 1/2 \quad \forall k$ meaning the state is maximally mixed.

Notable, this metric is only valid when working with pure states, but not for mixed states as they occur in decoherent noisy circuits.

2) Entanglement of Formation (EF): The EF [41] can be used as a metric on mixed states. In the extensive review on entanglement measures from Ref [42] it is defined as

$$Q(|\psi\rangle) := \inf\left\{\sum_i p_i E\left(|\psi_i\rangle\langle\psi_i|\right) : \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|\right\} \quad (19)$$

which is the average entanglement over all pure state decompositions of the density matrix ρ. The entanglement for pure states is then calculated using the MW measure as introduced in Eq. 18. Finding a pure state decomposition however is a non-trivial task. In this work we utilise an eigenvalue and eigenstate decomposition of the density matrix ρ. Based on the resulting decomposition, we proceed with calculating the entanglement of each eigenstate while weighting it by its eigenvalue as depicted in the left part of the right hand side of Eq. 19.

V. NUMERICAL RESULTS

This section describe the numerical results of our work. Generally, we consider a circuit with the structure as introduced in Eq. 1 with $L = 1$ layers, $D \in \{1, 2\}$ input features and $n \in [3 \ldots 6]$ qubits. Note that by convention, such a model would have two trainable layers, which is also used when calculating expressibility and entanglement, even if we discard the input gates in that case. For the trainable part, we investigate in the ansätze depicted in Fig. 2. We use the Strongly-Entangling Ansatz (SEA) as introduced in Ref. [43] and a circular structure for the Hardware-Efficient Ansatz (HEA). Circuit 15 and Circuit 19 stem from Ref. [16] and are chosen based on their different structural properties and distinguishable expressibility and entangling capability.

In the following experiments, we apply the types of noise, described in Sec. III-E to each ansatz, with seven linearly increasing noise levels from 0% to 3% yielding a resolution of 0.5%. The trainable parameters are randomly uniformly sampled from $[0 \ldots 2\pi]$. For each parameter sample, we perform n-local measurements on all qubits and then collect and average the expectation value over all samples. We use five different seeds for the parameter sampler to ensure statistical stability against initialisation. Throughout all our experiments, for each of the five seeds we use 250 samples per parameter value, which is scaled exponentially with the number of qubits if not stated otherwise (*e.g.* for a three qubit circuit with 10 parameters, we sample $10 \times 250 \times 2^3$ times for each seed).

All our results are fully reproducible [44], with the code available in Ref. [45], which is closely intertwined with the "QML-Essentials" package [46].

A. Coefficients

In this subsection we present the main result of our work, which is the effect that noise has on the Fourier coefficients of a

(a) SEA from Ref. [43]. (b) Circular HEA.

(c) Circuit 15 from Ref. [16]. (d) Circuit 19 from Ref. [16].
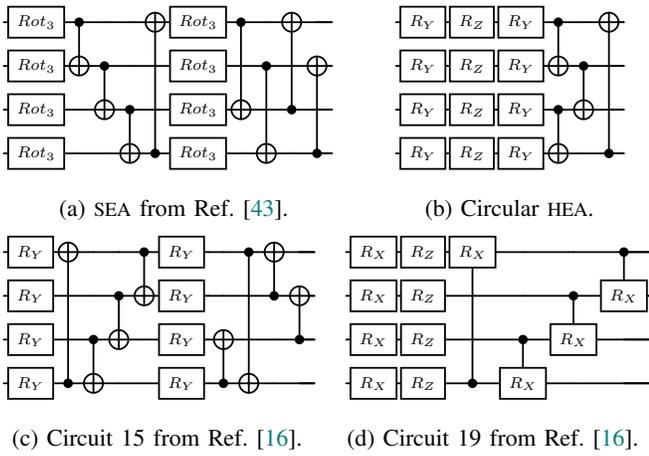
Fig. 2: ansätze investigated in this work, exemplarily for 4 qubits. Rot_3 represents an arbitrary single qubit rotational gate, which takes three parameters. The single qubit Pauli-rotation gates $R_{\{X,Y,Z\}}$ are parametrised by one parameter each.

Fig. 3: Absolute coefficient mean $\mu_c(\boldsymbol{\omega})$ for $[3\ldots6]$ qubits and $R_{\{X,Y,Z\}}$ encodings over frequencies.

Fig. 4: Coefficients, separated into real and imaginary parts for a circuit with six qubits and different single qubit Pauli-encodings. The individual frequency components are colour-coded.

QFM. As the standard deviation across seeds is small (*i.e.* $< 6\%$ for $\mu_c(\boldsymbol{\omega})$), figures show the mean of $\mu_c(\boldsymbol{\omega})$, $\sigma_c(\boldsymbol{\omega})$ or Cov_c over the five seeds. Note that if the samples for a coefficient result in a $\mu_c(\boldsymbol{\omega})$, smaller than 10^{-14}, it is assumed to be zero, indicating that the corresponding $\boldsymbol{\omega}$ is not the spectrum.

1) Input Encoding: As the coefficient values are greatly impacted by the encoding strategy [20], we first investigate in three different rotational Pauli-encodings around X, Y, and Z axis to encode one-dimensional data ($D = 1$) in a noiseless setting and observe how this not only changes the spectrum but also the real and imaginary part. In Fig. 3 we can observe that all circuits suffer from an exponential decay in the absolute coefficient value over the frequencies, varying across the structure. These findings are in line with the ones from [15]. Furthermore, the HEA with four to six qubits and Circuit 15 with five qubits resulted in $\mu_c(\boldsymbol{\omega})$ smaller than 10^{-14} for some frequencies, not exhibiting a full spectrum, as the other ansätze. However, we note that the absolute coefficient for individual parameter samples was above the threshold, indicating that a full spectrum is possible with these ansätze, but unlikely due to sampling resolution in parameter space.

While there is no obvious difference in $\mu_c(\boldsymbol{\omega})$ throughout the different encoding strategies, things change when we look at the real and imaginary parts of the individual coefficients $c_{\boldsymbol{\omega}}(\boldsymbol{\theta})$ separately, as depicted in Fig. 4. Here, Circuit 15 exhibits an imaginary part in the coefficients, only for the case of an R_Y encoding, missing a degree of freedom in the imaginary part for the R_X and R_Z encoding. Therefore, in the subsequent experiments, we use the R_Y encoding for Circuit 15, and R_X for the remaining ansätze.

2) Impact of Noise on the Real- and Imaginary Parts: Fig. 4 suggests that, the coefficients are evenly distributed in the real and imaginary part, with no clear correlation between the parts (apart from Circuit 15 with R_X and R_Y encoding). To corroborate this observation, we next compute the elements
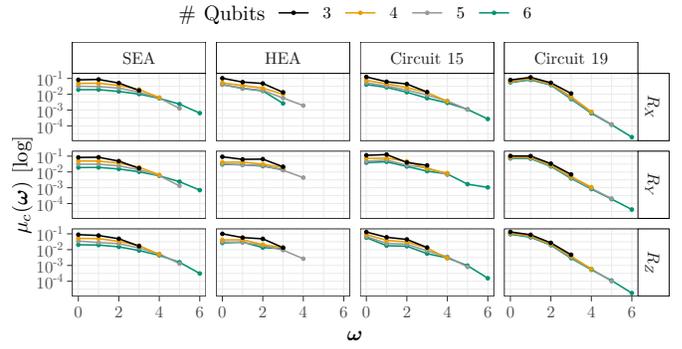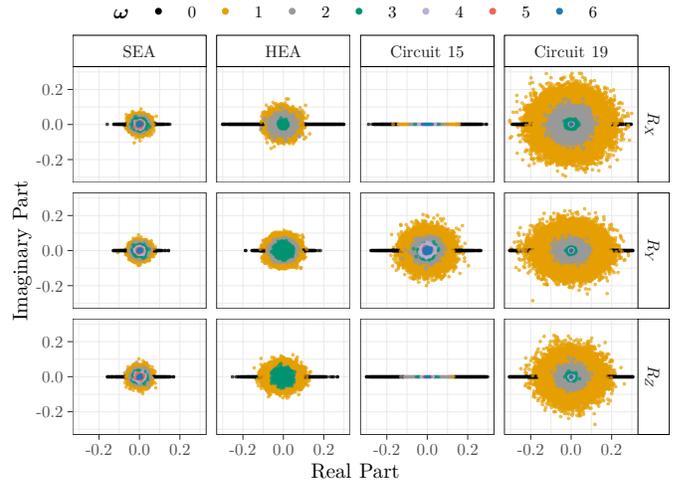
of the coefficients covariance matrix, as introduced in Eq. 14. In our experiment, we calculate these coefficients for noiseless samples and for samples subjected to each noise type at a probability of 3%. Since all circuit sizes yield consistent results, we focus on six-qubit circuits operating at the first and the last frequencies, as shown in Fig. 5.

These results align with those in Fig. 4, demonstrating that the variances of both the real and imaginary parts of the coefficients are approximately equally high across configurations. However, the covariance between these components is significantly smaller or even zero. This suggests that the real and imaginary parts of a coefficient are not strongly correlated, allowing us to consider only the absolute value of the coefficients in subsequent experiments. While additional variance experiments are detailed in Sec. V-B, Fig. 4 highlights that decoherent gate errors significantly reduce the variance of the coefficients, regardless of circuit structure. In contrast, other noise types appear to have no substantial impact on variance.

As shown in Fig. 4 coefficients are centred around zero for all configurations in a noiseless setting. Generally, this also
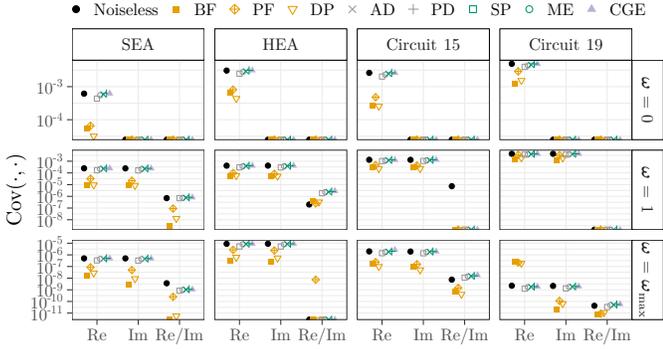
Fig. 5: Elements of the coefficients covariance matrix in noiseless and noisy (3% of each noise type) settings for six qubits. Points on the bottom border are zero. $\omega = \omega_{\max}$ denotes the maximum frequency in the noiseless spectrum, that is $\omega = 4$ for the HEA and $\omega = 6$ for the remaining ansätze.
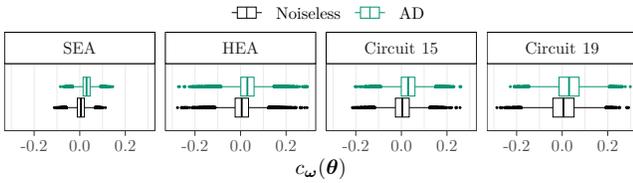
Fig. 6: Real coefficient for $\omega = 0$ in a noiseless setting vs. applying 3% of AD for circuits with six qubits.

(a) Absolute coefficient mean.

(b) Relative standard deviation of absolute coefficients.

Fig. 7: Absolute coefficient mean $\mu_c(\omega)$ and the corresponding relative standard deviation $\sigma_c(\omega)$ over noise levels for various types of noise and six qubit circuits. The individual frequency components are colour-coded.

applies to noisy scenarios with the exception of the AD channel, which shifts the coefficients of $\omega = 0$ towards positive values, as shown in Fig. 6. This indicates that the offset due to the Z-basis measurement is skewed when applying an AD channel.

B. Impact of Noise on the Absolute Value

Next, we examine how noise affects the coefficient mean and standard deviation across different ansätze configurations for one-dimensional inputs ($D = 1$). The results are shown in Fig. 7 for each type of noise, evaluated on six qubit circuits. Note that we conducted the same set of experiments for circuits with three to five qubits, which provided similar outcomes, but accordingly with fewer frequencies. A selection of the results is provided in Sec. V-C while the supplementary material in Ref. [45] provides full result sets.

Fig. 7a shows that decoherent gate errors generally lead to an exponential decay of coefficient mean; except for HEA and Circuit 19, where higher frequency components tend to increase with increasing noise level. This may seem counter-intuitive, but can be explained by a low coefficient mean for these frequencies without noise. Adding noise randomly distorts the expectation value, leading to a uniform increase (on average) of all frequencies in the spectral representation. The effect is less pronounced for lower frequency components as they typically have a higher mean value.

The relative standard deviation remains constant or decreases with noise for all frequencies and ansätze, while frequencies $\omega \neq 0$ generally lead to higher $\sigma_c(\omega)$. Since $\sigma_c(\omega)$ is proportional to $\mu_c(\omega)$, this implies that the absolute variance also decreases exponentially with increasing noise level under the influence of decoherent gate errors. Interestingly, in case of the SEA, the standard deviation for $\omega = 0$ increases with increasing noise level. Circuit 19 provides the highest standard deviation at 0% noise level which then decays rapidly when a small amount of decoherent noise is added. The remaining ansätze obtain $\sigma_c(\omega)$ values of a similar order of magnitude.

Generally SPAM, damping noise and CGE have minimal effect on both, coefficient mean relative standard deviation. However, we observe two exceptional high-frequency cases in Circuit 15 for CGE, where coefficients increase, with decreasing standard deviations decrease (see Sec. V-D for a detailed discussion).

C. Two-Dimensional Inputs

To assess the impact of noise on circuits with multiple input features, we conduct the same set of experiments detailed in Sec. V-B, also for $D = 2$, with the difference of only sampling a fixed number of times (*i.e.* 300) for all qubit numbers to save computational resources, resulting in a higher standard deviation in $\mu_c(\boldsymbol{\omega})$ of up to 29% between seeds for some six qubit cases. As different encoding strategies are presented in Ref. [19], this work focuses on a subsequent Pauli-encoding along different rotational axes. This is similar to the two-dimensional case, where it just satisfies the requirement of orthogonality. Given the substantial similarities across different noise types and frequencies, we present a curated selection of combined results for $D \in \{1, 2\}$ across all qubit numbers and ansätze in Fig. 8. The comprehensive results are available in our supplementary material in Ref. [45]. Our focus is on the effects of noise on two specific coefficient classes: the zero-frequency coefficient ($\boldsymbol{\omega} = \mathbf{0}$) and the coefficient corresponding to the maximum frequency in all input dimensions in the respective spectrum, determined at the 0% noise level ($\boldsymbol{\omega} = \boldsymbol{\omega}_{\max}$). For intermediate frequencies, as illustrated in Fig. 3, the coefficients decrease. Nevertheless, the effect that noise has on these intermediate coefficient, can also be observed either at $\boldsymbol{\omega} = \mathbf{0}$, or $\boldsymbol{\omega} = \boldsymbol{\omega}_{\max}$.

DP, as a representative of the class of decoherent gate errors, uniformly causes an exponential decline of $\mu_c(\boldsymbol{\omega})$ across all coefficients and frequencies, with the exception of higher frequencies in Circuit 19 (cf. Sec. V-B).
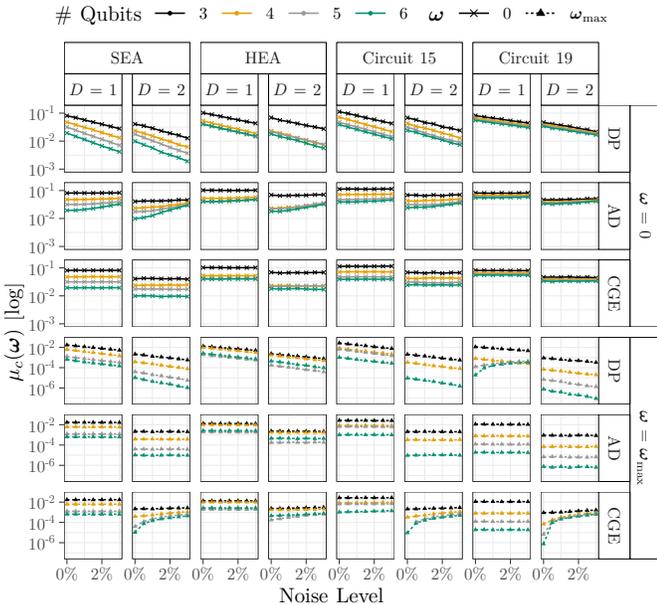
Fig. 8: Absolute coefficient mean $\mu_c(\boldsymbol{\omega})$ for the lowest frequency $\boldsymbol{\omega} = \mathbf{0}$ and highest frequency $\boldsymbol{\omega} = \boldsymbol{\omega}_{\max}$ in the respective spectrum under the influence of varying noise levels. We considered one-dimensional ($D = 1$), and two-dimensional inputs ($D = 2$). The y-axis for each facet row are equal throughout the respective $\boldsymbol{\omega}$, but differs in between.

For the zero-coefficient, we observe that applying an AD leads to an increase in $\mu_c(\boldsymbol{\omega})$. This effect is more pronounced in circuits with a higher number of qubits and input features. While this shift is more evident in the SEA and HEA, it is less noticeable for Circuit 15 and Circuit 19, suggesting these ansätze potentially exhibit greater resilience against AD. For the remaining frequencies (specifically $\boldsymbol{\omega}_{\max}$), AD has an imperceptible effect. Similarly, SPAM errors, and PD, not detailed in Fig. 8, exhibit no noticeable, or as in Fig. 7a only a slight decrease in $\mu_c(\boldsymbol{\omega})$ for all $\boldsymbol{\omega}$. Overall, these observations align with the proofs on decoherent noise effects in Ref. [37], showing (uniform) contractions for (DP) noise channels.

The relative standard deviation $\sigma_c(\boldsymbol{\omega})$ for $D = 2$, which is not explicitly shown here, demonstrates similar behaviour, as observed in Fig. 7b, with constant or decreasing values throughout the noise levels, indicating an overall decrease of variance with $\mu_c(\boldsymbol{\omega})$. Regarding CGEs, their influence is minimal, when $\boldsymbol{\omega}$ is small. However, for high frequencies, there is a significant increase in $\mu_c(\boldsymbol{\omega})$ with increasing noise levels in configurations with more than three qubits and $D = 2$. These findings are presented in depth in the following subsection.

D. Effect of Coherent Noise

In certain configurations shown in Fig. 7 and Fig. 8, a noticeable increase in higher-frequency coefficients can be observed. This effect is attributed to the CGE, which are modelled by adding $\boldsymbol{\epsilon}_{\boldsymbol{x}} \sim \mathcal{N}(0, p_{\text{cge}}^2)$ to the inputs The resulting Fourier series becomes

$$
\begin{aligned}
f_{\text{cge}}(\boldsymbol{x} + \boldsymbol{\epsilon}_{\boldsymbol{x}}, \boldsymbol{\theta}) &= \sum_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} c_{\boldsymbol{\omega}}(\boldsymbol{\theta}) e^{\imath \boldsymbol{\omega}^T (\boldsymbol{x} + \boldsymbol{\epsilon}_{\boldsymbol{x}})} \\
&= \sum_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} c_{\boldsymbol{\omega}}(\boldsymbol{\theta}) e^{\imath (\boldsymbol{\omega} + \boldsymbol{\epsilon}_{\boldsymbol{x}})^T \boldsymbol{x}},
\end{aligned}
\tag{20}
$$

where we neglegt that similarly a $\boldsymbol{\epsilon}_{\boldsymbol{\theta}} \sim \mathcal{N}(0, p_{\text{cge}}^2)$ is added to the trainable parameters.

This demonstrates that CGE not only alter the coefficients but also induce shifts in the frequency spectrum. Notably this effect is very similar to what was observed in Ref. [23] in the context of additional trainable parameters added to the encoding. The increase in higher-frequency coefficients, as shown in Fig. 7, is likely due to these spectral shifts, which lead to a more uniform distribution of coefficients.

The impact of CGE on the frequency spectrum becomes evident when comparing the number of frequencies with and without coherent noise, as illustrated in Fig. 9. In particular, coefficients that are computed to be zero in the FFT under noiseless conditions become non-zero in the presence of CGE.

Notably, HEA and in case of R_Y encoding also Circuit 15 do not result in a complete frequency spectrum. This is likely due to limitations in the ansatz structure and input encodings, which restrict the set of achievable eigenvalue differences (cf. Eq. 5). Such incomplete spectra are indicative of reduced expressiveness in these configurations. While CGE seems to solve this problem, we would like to point out that although the circuit under CGE contains more frequencies, this does not mean that they are individually tunable. This can result in a
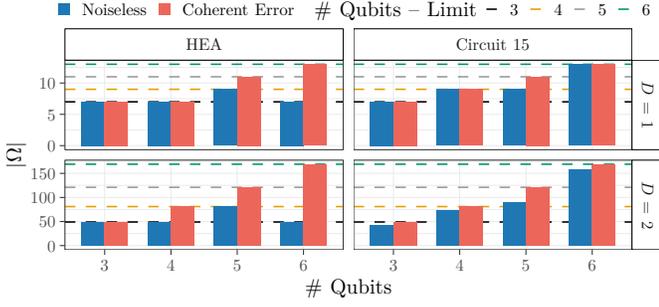
Fig. 9: Number of frequencies in the spectrum with and without applying a CGE for D-dimensional inputs and only those circuits where the maximum possible number of frequencies is not achieved in the noiseless case.

spectrum that is significantly correlated, a phenomenon that is undesirable in the context of most learning problems.

E. Expressibility

The expressibility is calculated as introduced in Sec. IV-C and quantified using the KL-divergence to the Haar distribution (cf. Eq. 17). The experimental results for all ansätze and qubit counts are presented in Fig. 10.
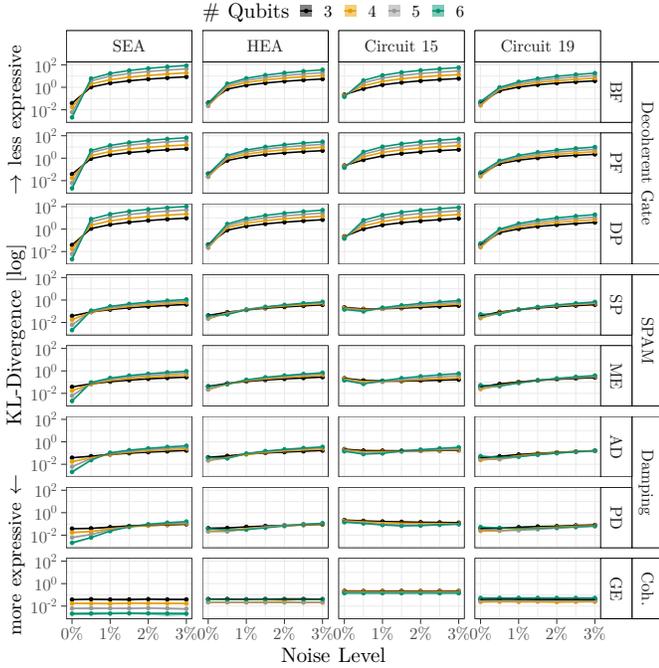
Fig. 10: Expressibility (*i.e.* inverse of the KL divergence) under the influence of increasing noise levels. The points represent the mean and the small shaded areas around it refer to the minimum and maximum across all five seeds.

We observe that all decoherent gate errors lead to an increase in the KL-divergence, thereby reducing expressibility. This effect is also consistent across SPAM and damping noise, though the impact is less pronounced. The observed decrease in expressibility aligns with the reduced variance of coefficients

noted in Sec. V-A and the connection to frequency redundancies discussed in Ref. [15].

Coherent noise, on the other hand, has no measurable effect on expressibility. This is not surprising since we omitted the input gates for the CGE to act upon, leaving the system in a coherent state. Effects of shifts in the trainable parameters appear to cancel out over the parameter samples.

In a noiseless environment, the SEA achieves a higher expressibility compared to the other ansätze. However, when coherent noise is applied, the expressiveness is quickly equalised. While the effect is consistent across all circuits tested, it can be observed that it becomes more pronounced as the number of qubits increases. This suggests that all forms of decoherent noise degrade the quantum nature of the circuits.

F. Entanglement

We first utilise the MW measure as introduced in Sec. IV-D1 to compute the entanglement of the circuits without noise. Subsequently, the EF (cf. Sec. IV-D2) is used to measure the entangling capability for the mixed states with increasing noise levels. The results for are shown in Fig. 11. Notable for the noiseless circuits with the MW measure, an increasing qubit count leads to a higher overall value for the entangling capability, while the specific values for the SEA, HEA and Circuit 15 are similar. Circuit 19, which has controlled rotation gates as entangling gates, results in a lower MW entangling capability for more than 4 qubits.
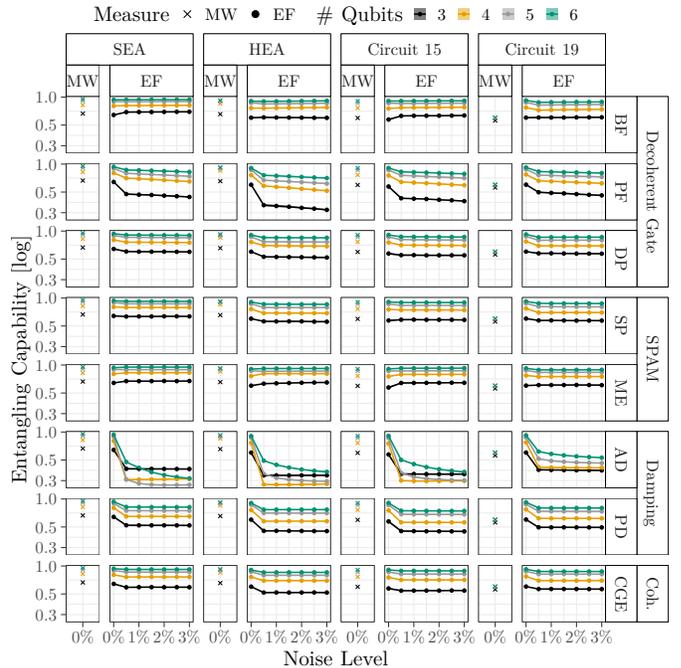
Fig. 11: MW entangling capability and EF under the influence of increasing noise levels. Points represent the mean of five seeds and lines are a linear interpolation to guide the eye. Shaded areas present the minimum/maximum entanglement across the five seeds. As the difference between seeds is small, it is not visible in the plot.

Notably, EF results in different values than the MW measure, also for noiseless circuits, especially observable for Circuit 19. This observation may be a consequence of the properties of the EF measure: At zero noise-probability for the EF, an eigendecomposition results in only one pure state, over which the MW measure is computed. However, this decomposition is not unique (cf. Sec. IV-D), and only one of many possibilities, potentially resulting in different values for the entangling capability. Therefore, we interpret the values in Fig. 11 as upper bounds of the entangling capability, which drop below the MW entangling capability with applied noise in most cases.

Within the decoherent gate-, and SPAM-error groups a similar behaviour can be observed for each ansatz. The values here slightly decrease or remain constant for both decoherent gate- and SPAM-errors, with some exceptions for three and four qubits with BF- and ME-noise, where the entanglement slightly increases from 0%-level noise to 0.5%. Similarly, for PD, the entangling capability also decreases with noise for all ansätze. For AD-noise entangling capability decreases more drastically for all circuits, and number of qubits with increasing noise level Although the state remains pure, when applying CGEs, a decrease in EF can also be observed in this case. Our results suggest that noise, even coherent noise, is overall detrimental for the entangling capability, especially for AD-noise.

VI. Discussion

In this work we explore the effects that noise has on the coefficients, expressibility and entangling capability of a QFMs. Empirically, we show an overall detrimental effect of noise on all of these properties, with the main general findings for QML outlined as follows: (1) We confirm the the statement of uniform contraction of coefficients with noise made in Ref. [29], and demonstrated that the decrease in the coefficient value is even exponential for decoherent gate errors, which may have a crucial impact on the kinds of functions that can be handled by the QFM. (2) We show that the expressibility, measured by the coefficients variance and the distance to the Haar distribution, also vanishes (potentially exponentially) with the coefficients value with increasing noise level. (3) The entangling capability, which is a general indicator of the "quantumness" of a VQC is affected by noise in a similar manner.

For all of the above, the ansätze that we investigated in showed a consistent, but not uniform pattern. While our findings are based on limited circuit and input dimension sizes (up to six qubits and two features), we observe consistent patterns as the number of qubits increases from three to six, suggesting potential for broader applicability in larger circuits. Additionally, similar results across one-dimensional and two-dimensional inputs offer a good indication for generalisation despite the exponential scaling of frequencies in the spectrum.

As each learning problem in QML has different requirements on the ansatz that is employed, a general "all-fits-one" ansatz is unlikely to exist. Nevertheless, based on our experiments, we can make some statements on the measures of quality for the tested ansätze:

(1) The SEA is, as the name suggests the ansatz with the highest entangling capability, although the difference to the HEA and Circuit 15 in that regard is small. Among the tested ansätze, the SEA also utilises most trainable parameters per qubit, resulting in a comparatively higher expressibility in a noiseless setting (cf. Fig. 10), which, however quickly vanishes, when applying only a small amount of noise. The corresponding Fourier spectrum is full (at least for up to two input dimensions), and comparatively uniformly distributed (cf. Fig. 3), also indicating expressiveness. (2) In the HEA, the amount by which the depth of the entangling sequence grows, is a constant, compared to the other ansätze, where it scales linearly with the number of qubits. This desirable property comes with the cost of lacking the full Fourier spectrum on average, even at one input dimension. (3) **Circuit 15** demonstrates similar behaviour to the SEA and HEA in the scaling of entanglement and expressibility under noise. It further has a unique characteristic, where the imaginary part in the coefficients appears exclusively for a R_Y encoding. As the HEA, Circuit 15 does not consistently achieve a full spectrum. Additionally, it seems to be more prone to CGE as other ansätze, yet less to AD. (4) **Circuit 19** achieves a full spectrum for up to two input dimensions, as the SEA. However, the coefficients vanish more quickly with increasing frequency, even in the noiseless case (cf. Fig. 3), indicating that not all parts of the spectrum may be effectively utilised. Despite having the lowest MW entangling capability, Circuit 19 features tunable parameters in the controlled-R_X gates, which could potentially allow for adjusting this property. Like Ciruit 15, it seems to be more susceptible to CGE, and less to AD.

VII. Conclusion and Outlook

Given the potential for conducting an extensive array of numerical experiments on this subject, the presented results are expected to offer researchers valuable insights into the behaviour of QFM under noise conditions. However, an analytical correlation between these factors and potential generalisations to larger circuit sizes remains to be investigated in future research. Also how different ansätze or encoding strategies, such as the ones presented in Ref. [19] fit into this pattern remains to be explored.

While we acknowledge that noise will lose its importance once FTQC is established, the road to FTQC remains stony, and as not every hardware architecture is suitable for error correction, the effects of noise on QFM are still relevant as the proposed results are hardware agnostic. In subsequent studies, we intend to perform analytical derivations to further extend the results presented and allow more definitive conclusions to be drawn, also for different ansätze.

REFERENCES

[1] F. Greiwe, T. Krüger, and W. Mauerer, "Effects of imperfections on quantum algorithms: A software engineering perspective," in *2023 IEEE International Conference on Quantum Software (QSW)*, 2023, pp. 31–42. DOI: 10.1109/QSW59989.2023.00014.

[2] M. Schuld and F. Petruccione, *Supervised Learning with Quantum Computers* (Quantum Science and Technology), en. Cham: Springer International Publishing, 2018. DOI: 10.1007/978-3-319-96424-9.

[3] Z. Zimborás *et al.*, *Myths around quantum computation before full fault tolerance: What no-go theorems rule out and what they don't*, arXiv:2501.05694 [quant-ph], Jan. 2025. DOI: 10.48550/arXiv.2501.05694.

[4] M. Ragone *et al.*, "A Lie Algebraic Theory of Barren Plateaus for Deep Parameterized Quantum Circuits," *Nature Communications*, vol. 15, no. 1, p. 7172, Aug. 2024, arXiv:2309.09342 [quant-ph]. DOI: 10.1038/s41467-024-49909-3.

[5] D. Heimann *et al.*, *Learning capability of parametrized quantum circuits*, arXiv:2209.10345 [quant-ph], Mar. 2024. DOI: 10.48550/arXiv.2209.10345.

[6] K. Georgopoulos, C. Emary, and P. Zuliani, "Modeling and simulating the noisy behavior of near-term quantum computers," *Physical Review A*, vol. 104, no. 6, p. 062 432, Dec. 2021, Publisher: American Physical Society (APS). DOI: 10.1103/physreva.104.062432.

[7] K. Wintersperger, H. Safi, and W. Mauerer, "Qpu-system co-design for quantum hpc accelerators," in *Architecture of Computing Systems*, M. Schulz *et al.*, Eds., Cham: Springer International Publishing, 2022, pp. 100–114.

[8] S. Thelen, H. Safi, and W. Mauerer, "Approximating under the influence of quantum noise and compute power," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 02, 2024, pp. 274–279. DOI: 10.1109/QCE60285.2024.10291.

[9] M. Franz *et al.*, "Uncovering instabilities in variational-quantum deep Q-networks," *Journal of the Franklin Institute*, vol. 360, no. 17, pp. 13 822–13 844, 2023. DOI: https://doi.org/10.1016/j.jfranklin.2022.08.021.

[10] M. Franz *et al.*, "Hype or heuristic? Quantum reinforcement learning for join order optimisation," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, 2024, pp. 409–420. DOI: 10.1109/QCE60285.2024.00055.

[11] M. Periyasamy *et al.*, "Guided-spsa: Simultaneous perturbation stochastic approximation assisted by the parameter shift rule," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, 2024, pp. 1504–1515. DOI: 10.1109/QCE60285.2024.00177.

[12] "Industry quantum computing applications," *EPJ Quantum Technology*, vol. 8, no. 1, Nov. 2021. DOI: 10.1140/epjqt/s40507-021-00114-x.

[13] C. Carbonelli *et al.*, "Challenges for quantum software engineering: An industrial application scenario perspective," in *Quantum Software: Aspects of Theory and System Design*, I. Exman *et al.*, Eds. Cham: Springer Nature Switzerland, 2024, pp. 311–335. DOI: 10.1007/978-3-031-64136-7_12.

[14] M. Gogeissl, H. Safi, and W. Mauerer, "Quantum data encoding patterns and their consequences," in *Proceedings of the 1st Workshop on Quantum Computing and Quantum-Inspired Technology for Data-Intensive Systems and Applications*, ser. Q-Data '24, Santiago, AA, Chile: Association for Computing Machinery, 2024, pp. 27–37. DOI: 10.1145/3665225.3665446.

[15] H. Mhiri *et al.*, *Constrained and Vanishing Expressivity of Quantum Fourier Models*, arXiv:2403.09417, Mar. 2024. DOI: 10.48550/arXiv.2403.09417.

[16] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, "Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms," en, *Advanced Quantum Technologies*, vol. 2, no. 12, p. 1 900 070, 2019. DOI: 10.1002/qute.201900070.

[17] M. Schuld, R. Sweke, and J. J. Meyer, "The effect of data encoding on the expressive power of variational quantum machine learning models," *Physical Review A*, vol. 103, no. 3, p. 032 430, Mar. 2021, arXiv: 2008.08605. DOI: 10.1103/PhysRevA.103.032430.

[18] A. Pérez-Salinas *et al.*, "Data re-uploading for a universal quantum classifier," *Quantum*, vol. 4, p. 226, Feb. 2020, arXiv: 1907.02085. DOI: 10.22331/q-2020-02-06-226.

[19] B. Casas and A. Cervera-Lierta, "Multidimensional Fourier series with quantum circuits," *Physical Review A*, vol. 107, no. 6, p. 062 612, Jun. 2023, arXiv: 2302.03389. DOI: 10.1103/PhysRevA.107.062612.

[20] M. C. Caro *et al.*, "Encoding-dependent generalization bounds for parametrized quantum circuits," *Quantum*, vol. 5, p. 582, Nov. 2021, arXiv: 2106.03880. DOI: 10.22331/q-2021-11-17-582.

[21] N. A. Nemkov, E. O. Kiktenko, and A. K. Fedorov, "Fourier expansion in variational quantum algorithms," *Physical Review A*, vol. 108, no. 3, p. 032 406, Sep. 2023, arXiv:2304.03787 [quant-ph]. DOI: 10.1103/PhysRevA.108.032406.

[22] M. Wiedmann, M. Periyasamy, and D. D. Scherer, *Fourier Analysis of Variational Quantum Circuits for Supervised Learning*, arXiv:2411.03450, Nov. 2024. DOI: 10.48550/arXiv.2411.03450.

[23] B. Jaderberg *et al.*, "Let Quantum Neural Networks Choose Their Own Frequencies," *Physical Review A*, vol. 109, no. 4, p. 042 421, Apr. 2024, arXiv:2309.03279 [quant-ph]. DOI: 10.1103/PhysRevA.109.042421.

[24] M. Cerezo *et al.*, "Cost Function Dependent Barren Plateaus in Shallow Parametrized Quantum Circuits," *Nature Communications*, vol. 12, no. 1, p. 1791, Mar. 2021, arXiv: 2001.00550. DOI: 10.1038/s41467-021-21728-w.

[25] M. Larocca *et al.*, "Barren plateaus in variational quantum computing," en, *Nature Reviews Physics*, pp. 1–16, Mar. 2025, Publisher: Nature Publishing Group. DOI: 10.1038/s42254-025-00813-9.

[26] S. Wang *et al.*, "Noise-Induced Barren Plateaus in Variational Quantum Algorithms," *Nature Communications*, vol. 12, no. 1, p. 6961, Nov. 2021, arXiv: 2007.14384. DOI: 10.1038/s41467-021-27045-6.

[27] M. Cerezo *et al.*, "Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing," arXiv, Tech. Rep., Dec. 2023, arXiv: 2312.09121. DOI: 10.48550/arXiv.2312.09121.

[28] F. J. Schreiber, J. Eisert, and J. J. Meyer, "Classical surrogates for quantum learning models," arXiv, Tech. Rep., Jun. 2022, arXiv: 2206.11740. DOI: 10.48550/arXiv.2206.11740.

[29] E. Fontana *et al.*, "Classical simulations of noisy variational quantum circuits," arXiv, Tech. Rep., Jun. 2023, arXiv: 2306.05400. DOI: 10.48550/arXiv.2306.05400.

[30] J. Landman *et al.*, "Classically Approximating Variational Quantum Machine Learning with Random Fourier Features," arXiv, Tech. Rep., Oct. 2022, arXiv: 2210.13200. DOI: 10.48550/arXiv.2210.13200.

[31] R. Sweke *et al.*, "Potential and limitations of random Fourier features for dequantizing quantum machine learning," *Quantum*, vol. 9, p. 1640, Feb. 2025, arXiv:2309.11647 [quant-ph]. DOI: 10.22331/q-2025-02-20-1640.

[32] K. Sharma *et al.*, "Reformulation of the No-Free-Lunch Theorem for Entangled Datasets," *Physical Review Letters*, vol. 128, no. 7, p. 070 501, Feb. 2022, Publisher: American Physical Society. DOI: 10.1103/PhysRevLett.128.070501.

[33] X. Wang *et al.*, "Transition role of entangled data in quantum machine learning," en, *Nature Communications*, vol. 15, no. 1, p. 3716, May 2024, Publisher: Nature Publishing Group. DOI: 10.1038/s41467-024-47983-1.

[34] J. Bowles, S. Ahmed, and M. Schuld, "Better than classical? The subtle art of benchmarking quantum machine learning models," arXiv, Tech. Rep., Mar. 2024, arXiv: 2403.07059. DOI: 10.48550/arXiv.2403.07059.

[35] T. Rohe *et al.*, *The Questionable Influence of Entanglement in Quantum Optimisation Algorithms*, arXiv:2407.17204 [quant-ph], Jul. 2024. DOI: 10.48550/arXiv.2407.17204.

[36] A. Joch, G. S. Uhrig, and B. Fauseweh, *Entanglement-informed Construction of Variational Quantum Circuits*, arXiv:2501.17533 [quant-ph], Jan. 2025. DOI: 10.48550/arXiv.2501.17533.

[37] E. Fontana *et al.*, "Spectral analysis for noise diagnostics and filter-based digital error mitigation," arXiv, Tech. Rep., Nov. 2022, arXiv: 2206.08811. DOI: 10.48550/arXiv.2206.08811.

[38] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951, Publisher: Institute of Mathematical Statistics. DOI: 10.1214/aoms/1177729694.

[39] D. A. Meyer and N. R. Wallach, "Global entanglement in multiparticle systems," *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4273–4278, Sep. 2002. DOI: 10.1063/1.1497700.

[40] G. K. Brennen, "An observable measure of entanglement for pure states of multi-qubit systems," arXiv, Tech. Rep., Nov. 2003. DOI: 10.48550/arXiv.quant-ph/0305094.

[41] W. K. Wootters, "Entanglement of Formation of an Arbitrary State of Two Qubits," *Physical Review Letters*, vol. 80, no. 10, pp. 2245–2248, Mar. 1998, arXiv:quant-ph/9709029. DOI: 10.1103/PhysRevLett.80.2245.

[42] M. B. Plenio and S. Virmani, *An introduction to entanglement measures*, arXiv:quant-ph/0504163, Jun. 2006. DOI: 10.48550/arXiv.quant-ph/0504163.

[43] M. Schuld *et al.*, "Circuit-centric quantum classifiers," *Physical Review A*, vol. 101, no. 3, p. 032 308, Mar. 2020, arXiv:1804.00633 [quant-ph]. DOI: 10.1103/PhysRevA.101.032308.

[44] W. Mauerer and S. Scherzinger, "1-2-3 reproducibility for quantum software experiments," in *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, 2022, pp. 1247–1248. DOI: 10.1109/SANER53432.2022.00148.

[45] M. Franz and M. Strobl, *Reproduction Package for "Out of Tune: Demystifying Noise-Effects on Quantum Fourier Models"*, Apr. 2025. DOI: 10.5281/zenodo.15211318.

[46] M. Strobl, M. Franz, and P. Schillinger, *QML Essentials - A framework for working with Quantum Fourier Models*, Apr. 2025.