A Combinatorial Search of Parameterized Quantum Circuit Learning for Chemical Applications

Grier M. Jones,[†,‡] Nick Taylor,[†] Viki Kumar Prasad,[†,‡] Ulrich Fekl,[*,‡] and Hans-Arno Jacobsen[*,†]

† *The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 Kings College Road, Toronto, Ontario, Canada M5S 3G4*

‡ *Department of Chemical and Physical Sciences, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, Ontario, Canada L5L 1C6*

E-mail: ulrich.fekl@utoronto.ca; jacobsen@eecg.toronto.edu

Abstract

The rapid growth of quantum computing has seen a rise in the development of applicable near-term quantum algorithms. A promising candidate to run on current noisy intermediate-scale quantum (NISQ) hardware is quantum machine learning (QML). Parameterized quantum circuits can be learned as models in a hybrid classical-quantum approach. It has been shown that QML models may need less training data to generalize compared to classical methods. Although QML has been explored largely within classification tasks, regression tasks have received much less attention. Existing studies have only utilized few qubits and lack relevant practical applications. We will discuss our work on the development and comparisons of QML models for highly desired prediction of chemical properties such as bond dissociation energies and barrier height energies. Our work in-

vestigates the trainability, expressibility, and generalization capability of various quantum models for the purpose of regression of molecular properties. Our work also provides a benchmark against classical machine learning models and implications of using error mitigation techniques for QML.

Introduction

12345678

As current quantum hardware grows, much attention is brought to find applicable uses for the potentially powerful new technology while it continues to grow. Current research in the areas of developing quantum algorithms for use with current noisy intermediate-scale quantum (NISQ) hardware lays the foundation for understanding and applying quantum computing. These near-term algorithms have seen a wide deployment across many areas of science and technology as the field grows. Computational chemistry is a natural such candidate for quantum computing applications due to the quantum nature of many problems in the field. An especially important problem is that of calculating the bond separation energies of (molecules?). Efficient and accurate predictions of these values can aid research in material science and drug discovery.

One of the most promising near-term applications of quantum computing within any field has been quantum machine learning (QML). The advantages of QML comes largely from its ability to be readily deployed on quantum computers with few qubits, and its potential to speedup problems with minimal loss of accuracy. The number of qubits needed for a QML model is determined primarily by the number of features in the dataset. As there exist data science techniques developed in the realm of classical machine learning to reduce feature spaces, they can be adopted over to reduce problems to a suitable number of qubits. Further, since QML only needs to train once, it has potential to perform faster when compared to other current NISQ algorithms such as the variational quantum eigensolver, which needs to be trained separately for each problem.

Quantum machine learning has been widely applied to classification tasks with much success. However, little research has been done to understand its capabilities for regression. Previous works have set important groundwork to the understanding of QML for regression by theoretically formulating upper bounds for the generalization error. Most notably, this research suggests that with an efficiently implemented QML model can achieve good generalizability with modest training data sizes. This highlights a potential for QML models to outperform and out-generalize classical models. Further understanding and verification of this, however, requires numerical validation and

demonstration, and while some experimental work has been done, it remains limited to few qubits with little practical application.

Our work expands on the current knowledge base by providing a comprehensive exploration of QML for regression. Specfically, we will by applying QML to the problem of calculating bond dissociation enthalpies of various molecules. This demonstrates a practical application of regression in QML with many applications. The efficient and accurate prediction of these chemical properties can aid research in material science and drug discovery with far reaching implications. Our work serves to benchmark QML and test its capabilities within this context.

Methods

Quantum machine learning is a hybrid classical-quantum approach utilizing parameterized quantum circuits. Typical hybrid algorithms consist of shallow circuit depth of one qubit rotations and two qubit entangling gates. In the context of machine learning, we can subdivide our circuit into three components: the encoder, the variational circuit, and the act of measuring to estimate predictions. The encoder takes our n-dimensional data $\boldsymbol{x} \in \mathbb{R}^n$ and encodes it into a unitary feature map $U(\boldsymbol{x})$ in our working Hilbert Space. This unitary acts on our initialized state $|0\rangle^{\otimes n'}$ to produce the state $U(\boldsymbol{x})|0\rangle^{\otimes n'}$. It is usual to take $n' = n$ and for each feature x_i to correspond to a qubit. With $n < n'$, features can be repeated across the qubits present in the system. Further elaboration will be discussed with specific encoding circuits. After the encoder produces the state $U(\boldsymbol{x})|0\rangle^{\otimes n}$, we then tune the state with a parameterized unitary $U(\boldsymbol{\theta})$ to transform the state into a meaningful prediction. Finally, some configuration of qubits are measured to produce expectation values $\hat{y}_k$ for our data labels y_i which we then use to minimize a loss function. An obstacle in creating quantum models is there is no way to determine the best encoding or variational circuit for the problem. As such, this paper explores many different encoding and variational circuits.

Encoder Circuit

Representation of the data in any machine learning context is a crucial factor in its success. In a classical sense, this can be thought of as generating features and selecting the appropriate features to input into your model. In the quantum context, however, there is the additional question of how to translate the data into the Hilbert space. Different proposed methods include angle encoding, and amplitude encoding. Following the work done by Suzuki and Katouda, we explore three base encoders composed together with entangling operations. Additionally, we explore the IQP encoding as given in PennyLane.

We composed circuits out of primary rotation blocks and entangling blocks. These rotation blocks include the one proposed by Mitrai et al.,[3] a single angle rotation, and a double rotation encoder:

$$U_{A1} = \prod_{i=1}^{n} R_i^Y(x_i) \tag{1}$$

$$U_{A2} = \prod_{i=1}^{n} R_i^Y(x_i) R_i^Z(x_i) \tag{2}$$

$$U_M = \prod_{i=1}^{n} R_i^Z(\arccos x_i^2) R_i^Y(\arcsin x_i^2) \tag{3}$$

Where R_i^j is a rotation about the j axis on qubit i. It's clear to see that this method of encoding imposes a restriction on the number of features to the number of qubits. With this encoding method, it is possible to use more features than qubits and to re-encode the features along the remaining qubits. Then, to compose these encoders together we use entangling circuits that entangle adjacent qubits using either CNOT or CZ entangling operations. These encoding circuits can be seen presented in Figure **??**.

Variational Circuit

After encoding the data, we have a quantum circuit in a state $U_{\text{enc}}(\boldsymbol{x})\,|0\rangle^{\otimes n}$. We need to act on this state with a unitary that has parameters which can be tuned in our machine learning model.

Choosing this circuit ansatz can greatly dictate the performance of a given model. Model performance can be greatly influenced by the variational circuit chosen as well as the depth of the circuit as we will later show. It is not obvious which circuit we should pick for our given problem, especially given that our algorithm is not physically motivated like other popular parameterized quantum circuit algorithms. We have gathered a list of 12 variational circuit blocks from previous works and libraries which we will then benchmark for our particular problem. These circuits each can be decomposed into individual rotation gates and two-qubit entangling gates.

After our variational circuit acts on our state we have the state:

$$|\psi\rangle = U_{\text{var}}(\boldsymbol{\theta})U_{\text{enc}}(\boldsymbol{x})\,|0\rangle^{\otimes n} \tag{4}$$

where $\boldsymbol{\theta}$ is our parameter vector to be tuned in learning. We can further modify this state; by repeating variational circuit blocks we create a more rich and complex model with a better ability to explore the Hilbert space. This would give us the state:

$$|\psi\rangle = U_{\text{var}}(\boldsymbol{\phi})U_{\text{var}}(\boldsymbol{\theta})U_{\text{enc}}(\boldsymbol{x})\,|0\rangle^{\otimes n} \tag{5}$$

However, as we are interested in NISQ applications, it should be noted that increasing layer depths creates decoherence in quantum devices. Additionally, we can use the idea of data re-uploading to create redundancies in the encoding of the model to explore a richer feature space.

$$|\psi\rangle = U_{\text{var}}(\boldsymbol{\phi})U_{\text{enc}}(\boldsymbol{x})U_{\text{var}}(\boldsymbol{\theta})U_{\text{enc}}(\boldsymbol{x})\,|0\rangle^{\otimes n} \tag{6}$$

This paper uses and compares both the data re-uploading and increased depth approaches of creating more complicated models.

Measurement

After our system is in state $|\psi\rangle$ as given by either (4), (5) or (6) we can measure our system to generate a predicted value $\hat{y}_i$. At this point in the algorithm, any loss function can be used. In our simplest system, we measured the first qubit so that,

$$\hat{y}_i = \langle Z_0 \rangle_\psi \tag{7}$$

for a ψ generated by the kth data point. Then, we have the loss function we are trying to minimize as,

$$\mathscr{L}(\hat{\boldsymbol{y}}) = \sum_i^N \frac{(y_i - \hat{y}_i)^2}{N} \tag{8}$$

Implementation

Our work was implemented in Python using PennyLane for constructing the quantum circuits, and Qulacs as a primary backend for simulating the circuits. The optimization was implemented using the scipy.optimize module from the SciPy library. Classical models were implemented using scikit-learn while processing of the data was handled with Pandas, NumPy, and SciPy libraries.

Simulated using FakeCairo backend

Dataset

We trained our models using three datasets: a function fitting dataset, consisting of a noisy linear, quadratic, and sine function, used for model calibration; a dataset consisting of electronic structure features to predict wavefunctions using the data-driven coupled-cluster scheme of Townsend and Vogiatzis;[9] and a dataset of bond separation energies (BSE) of molecules, where the feature set encodes structural information of each molecule.

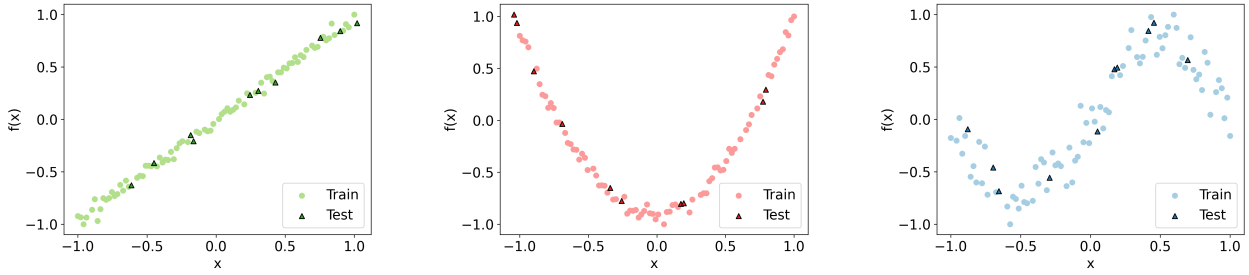
For the function fitting dataset, there is only one feature per target value. For the DDCC dataset, there are initially X features, reduced down to 5 and 16 features, respectively, using SHapley

Additive exPlanation (SHAP) values.[10]

$$t^{ab}_{ij(\mathrm{MP2})} = \frac{\langle ij||ab\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \tag{9}$$

For each $t^{ab}_{ij(\mathrm{CCSD})}$, orbital energies $(\varepsilon_i, \varepsilon_j, \varepsilon_a, \varepsilon_b)$, Coulomb and exchange integrals $(J^i_a, J^j_b, K^a_i, K^b_j)$, binary feature whether two electrons are promoted to the same virtual orbital, the initial MP2 amplitudes, along with the numerator (two-electron integrals ($\langle ij||ab\rangle$)) and denominator $\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b$.

For the BSE dataset, various molecular representations are explored, and we settled on using Morgan fingerprints with X parameters. For the feature reduction, we use PCA, as explained in Section .



(a) (b) (c)

Figure 1

Results

Classical Testing

Function Fitting

We found that on average the best models for the linear function is KNN (Fig. S1), KRR for quadratic (Fig. S2), and KNN for sine (Fig. S3).

DDCC

t_{ij}^{ab}

Dimensions reduced using SHapley Additive Explanation (SHAP) values

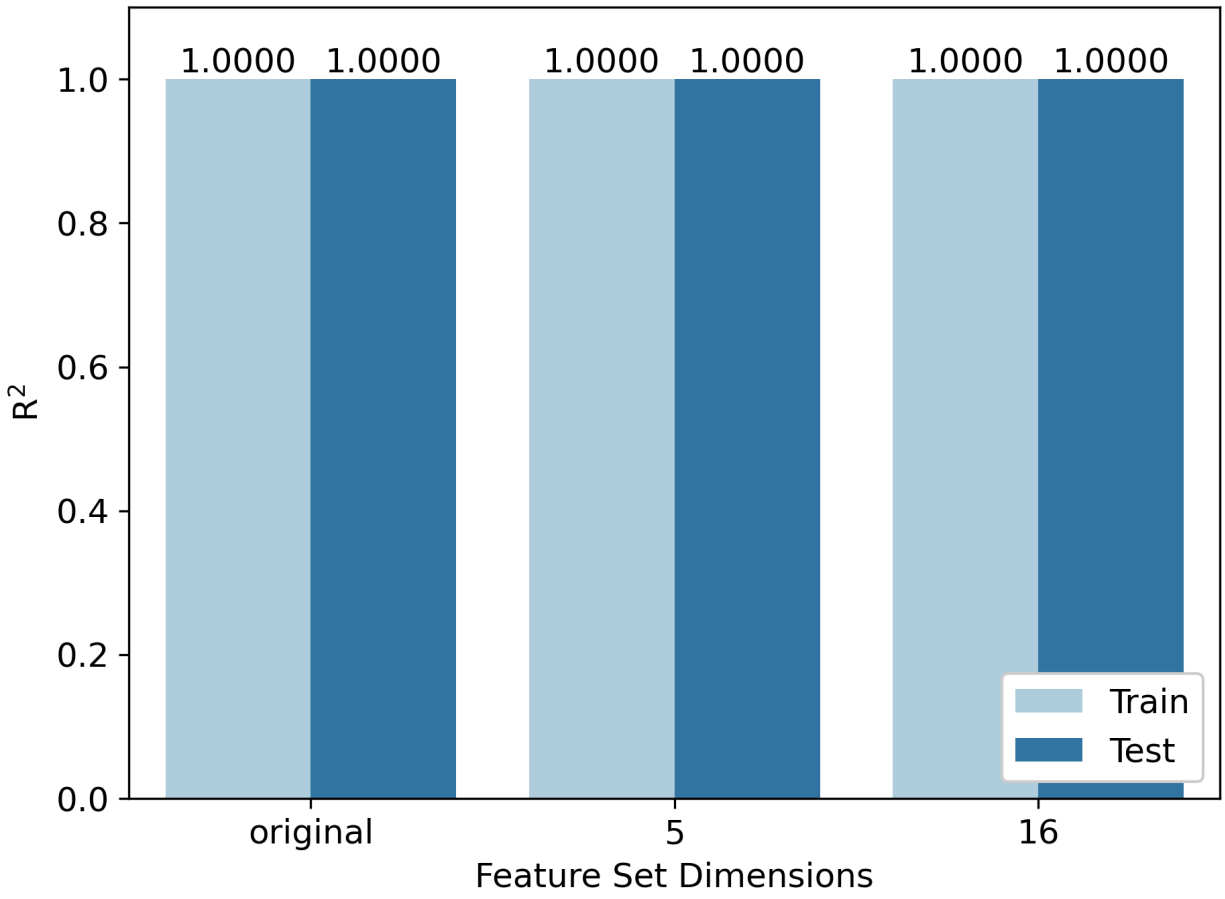


Figure 2: DDCC feature set

BSE

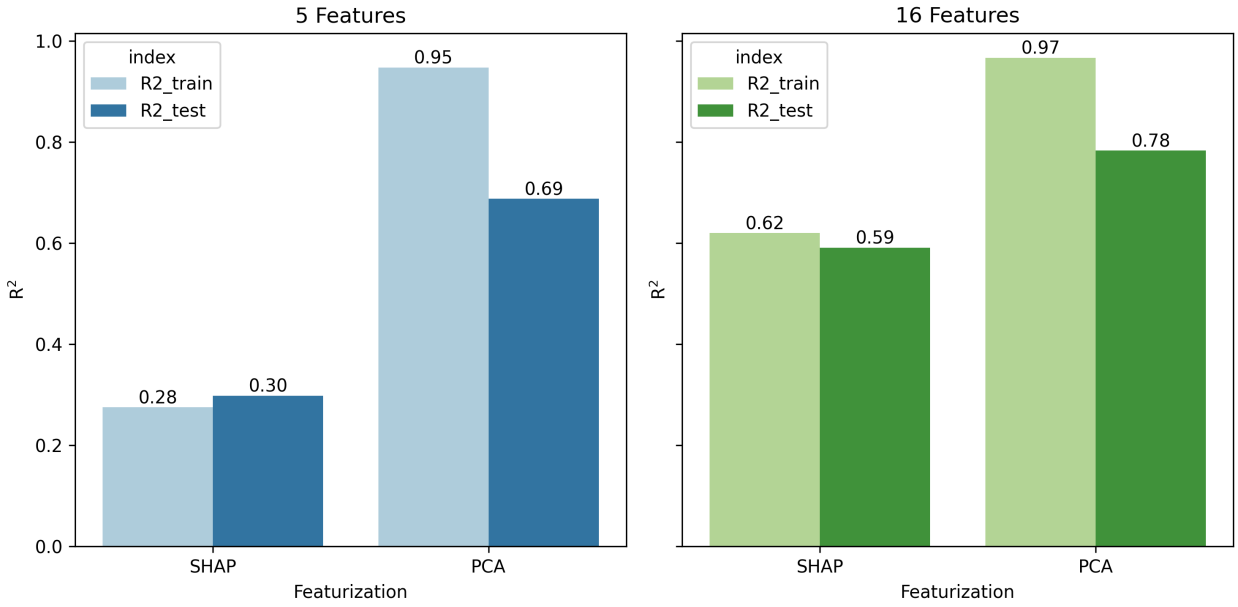


Figure 3: Feat redR2

5 qubit function fitting

Fig. S4 Fig. S5 Fig. S6

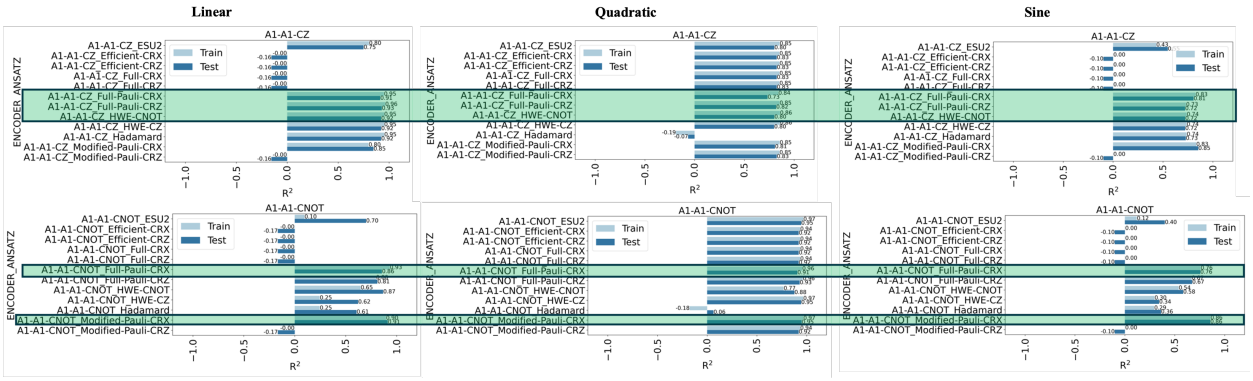


Figure 4: REMAKE

16 qubit function fitting

Fig. S7 Fig. S8 Fig. S9

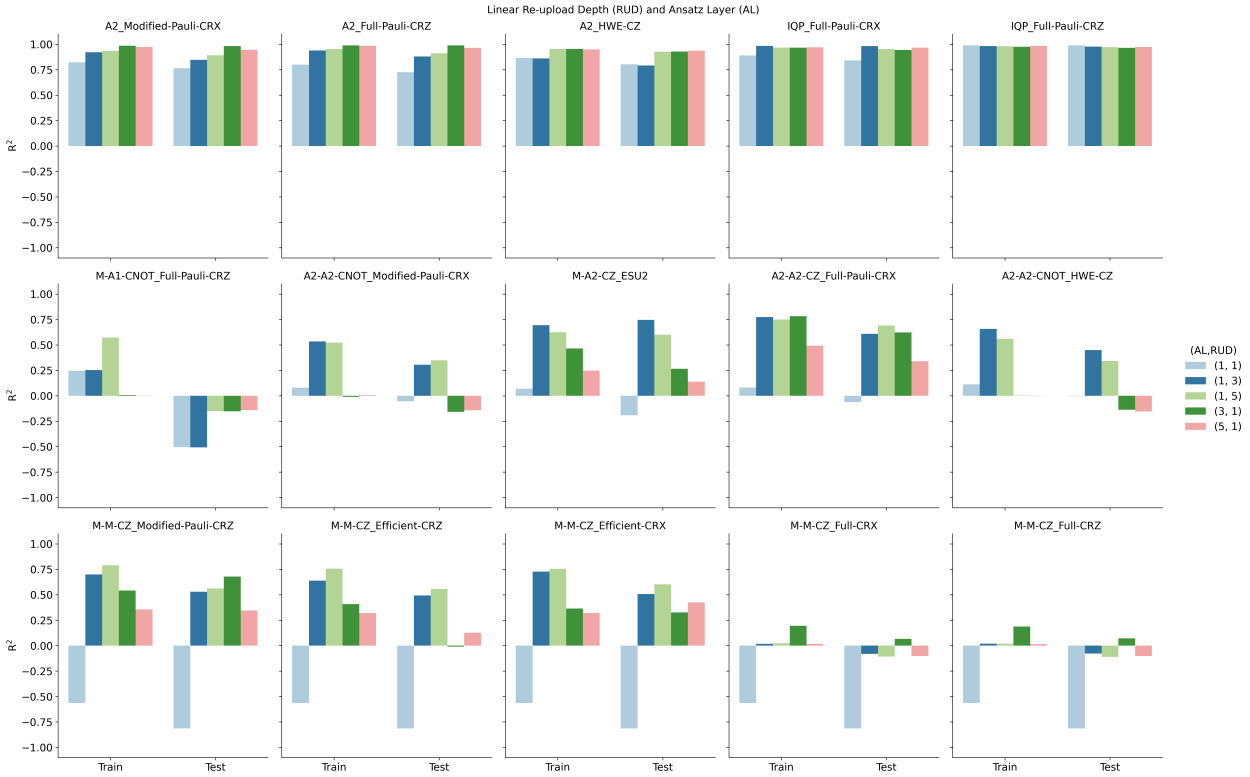


Figure 5: REMAKE

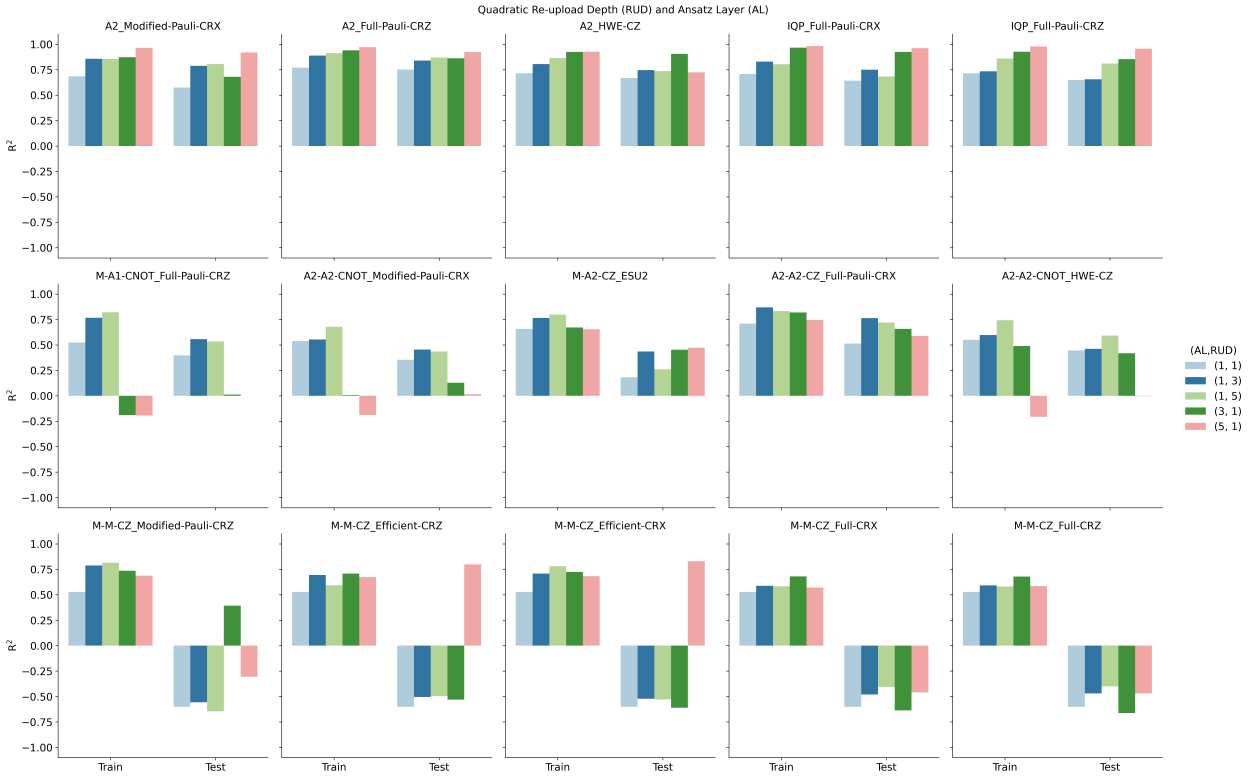


Figure 6: REMAKE

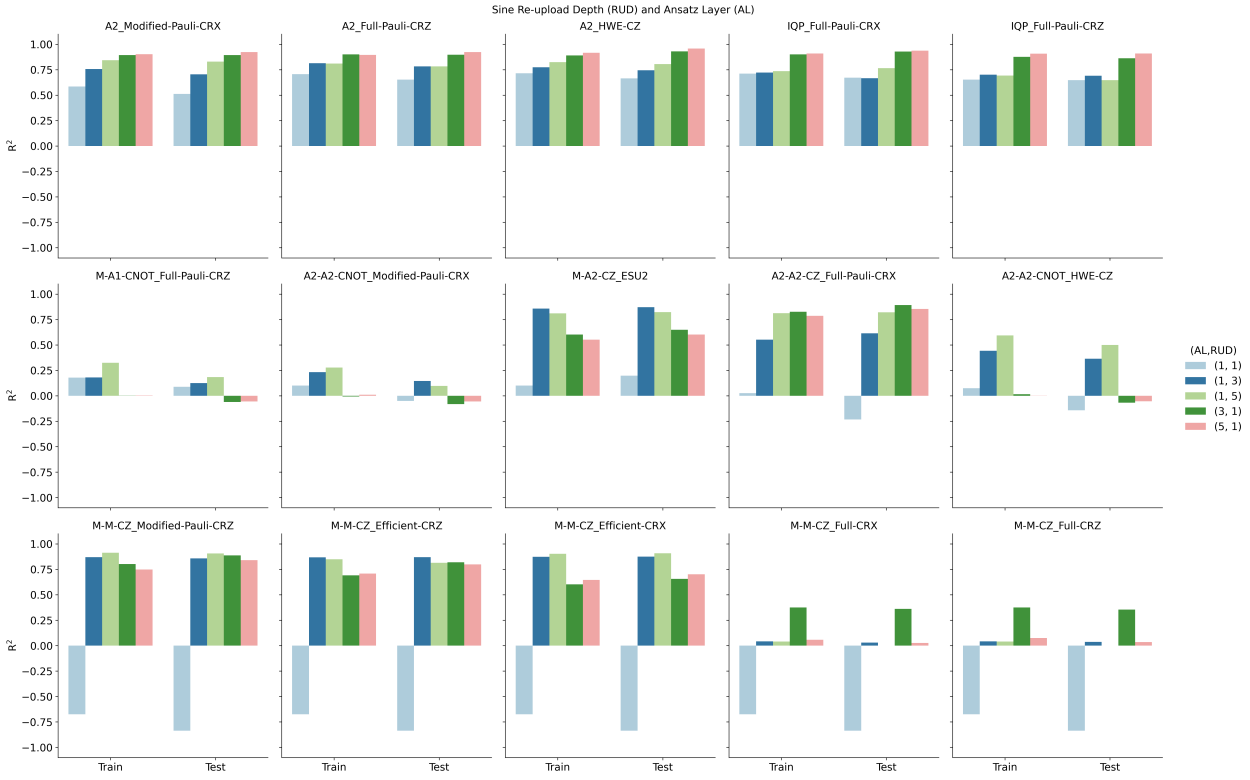
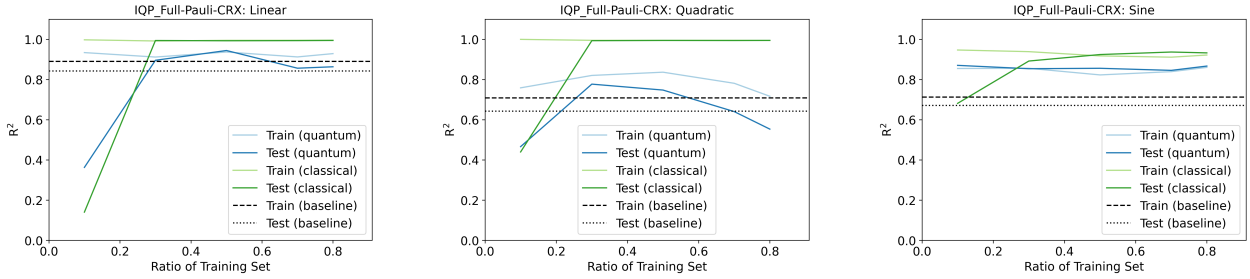


Figure 7: REMAKE

The best is IQP_Full-Pauli-CRX with 5AL and 5 RUD

Missing 0.8 quadratic



(a) (b) (c)

Figure 8: Baseline denotes IQP_Full-Pauli-CRX with 1AL and 1RUD

5 qubit BSE

Done

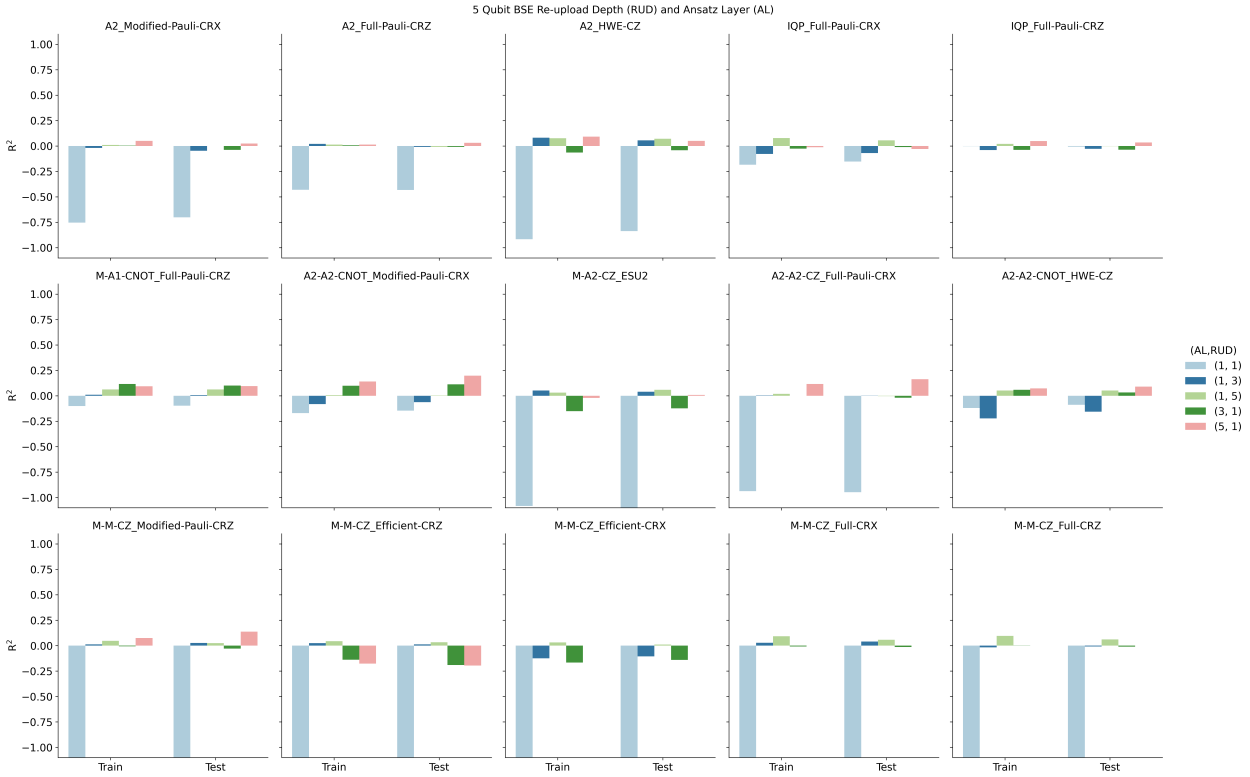


Figure 9: 5qubit BSE RUD AL

16 qubit BSE

Missing data

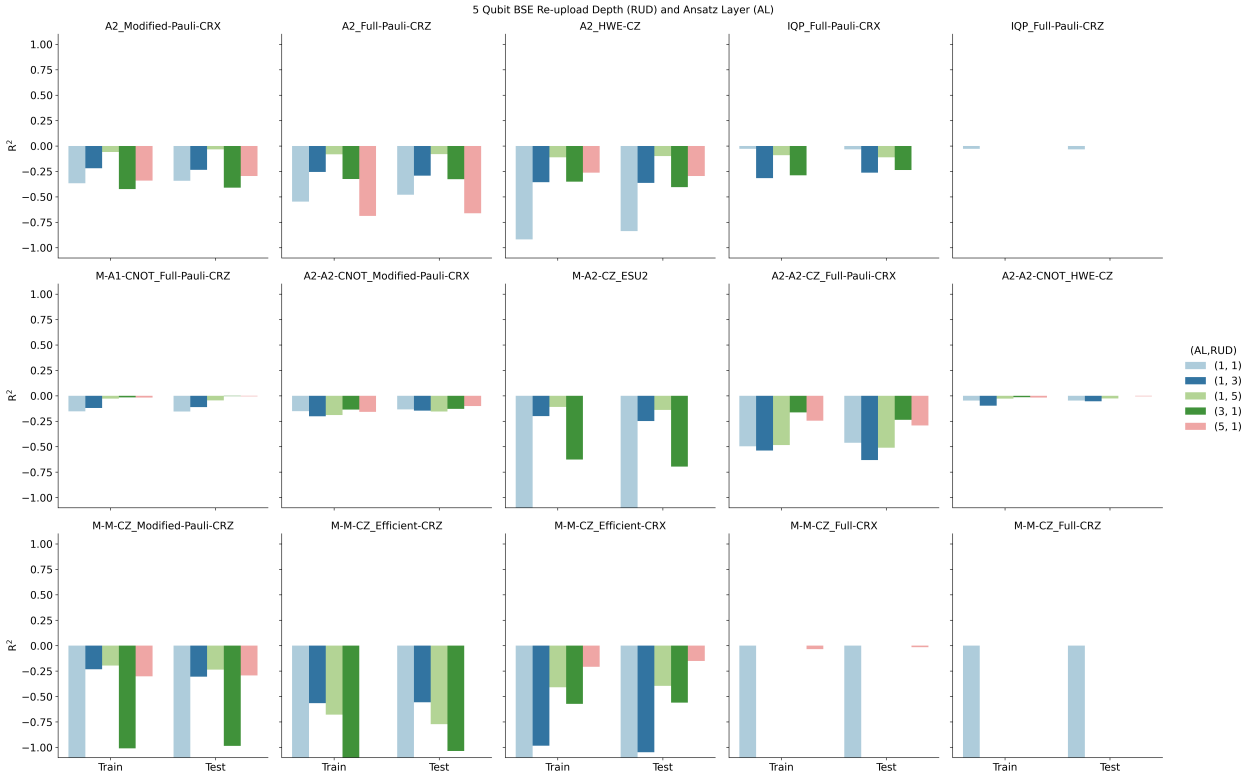


Figure 10: 16qubit BSE RUD AL

Error Correction

TREX/MITIQ ZNE LINEAR/ MITIQ ZNE Richardson

Function Fitting

Running

DDCC

Not ran yet

BSE

Not ran yet

IBM Device

References

(1) Prasad, V. K.; Cheng, F.; Fekl, U.; Jacobsen, H.-A. Applications of noisy quantum computing and quantum error mitigation to "adamantaneland": a benchmarking study for quantum chemistry. *Phys. Chem. Chem. Phys.* **2024**, *26*, 4071–4082.

(2) Suzuki, T.; Katouda, M. Predicting toxicity by quantum machine learning. *J. Phys. Commun.* **2020**, *4*, 125012.

(3) Mitarai, K.; Negoro, M.; Kitagawa, M.; Fujii, K. Quantum circuit learning. *Phys. Rev. A* **2018**, *98*, 032309.

(4) Sim, S.; Johnson, P. D.; Aspuru-Guzik, A. Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms. *Advanced Quantum Technologies* **2019**, *2*, 1900070.

(5) Hatakeyama-Sato, K.; Igarashi, Y.; Kashikawa, T.; Kimura, K.; Oyaizu, K. Quantum circuit learning as a potential algorithm to predict experimental chemical properties. *Digital Discovery* **2023**, *2*, 165–176.

(6) Krenn, M.; Landgraf, J.; Foesel, T.; Marquardt, F. Artificial intelligence and machine learning for quantum technologies. *Phys. Rev. A* **2023**, *107*, 010101.

(7) Benedetti, M.; Lloyd, E.; Sack, S.; Fiorentini, M. Parameterized quantum circuits as machine learning models. *Quantum Sci. Technol.* **2019**, *4*, 043001.

(8) Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202.

(9) Townsend, J.; Vogiatzis, K. D. Data-Driven Acceleration of the Coupled-Cluster Singles and

Doubles Iterative Solver. *J. Phys. Chem. Lett.* **2019**, *10*, 4129–4135, Publisher: American Chemical Society.

(10) Lundberg, S. M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems. 2017.