

# A Combinatorial Search of Parameterized Quantum Circuit Learning for Chemical Applications

Grier M. Jones,<sup>†,‡</sup> Nick Taylor,<sup>†</sup> Viki Kumar Prasad,<sup>†,‡</sup> Ulrich Fekl,<sup>\*,‡</sup> and Hans-Arno Jacobsen<sup>\*,†</sup>

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

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

E-mail: [ulrich.fekl@utoronto.ca](mailto:ulrich.fekl@utoronto.ca); [jacobsen@eecg.toronto.edu](mailto:jacobsen@eecg.toronto.edu)

## Abstract

Within the quantum machine learning (QML) field, parameterized quantum circuits (PQCs), built using fixed and parameterized gates, offer a hybrid approach for complex machine learning tasks. While many potential use cases have been proposed, the exploration of relevant datasets for chemists is lacking. Our study seeks to understand the possible advantages and disadvantages of PQCs for two chemically relevant datasets: one based on the bond separation energies of 49 different classes of bonds, called the BSE49 dataset, and another consisting of water confirmations, where coupled-cluster singles and doubles (CCSD) wave functions are predicted using electronic structure theory data from lower-level methods using the data-driven coupled-cluster (DDCC) method. In our study, we examine a combinatorial space of 14 data encoding layers and 12 variational (ansatz) layers,

for a combined total of 168 PQCs. To calibrate our PQCs, we utilize a dataset of noisy linear, quadratic, and sine functions to explore the effects of the circuit width and depth, the effects of the feature set size, and various error mitigation techniques. Following this step, we similarly examine our chemically relevant datasets. Our work highlights the difficulties in encoding classical molecular representations in a PQC for predicting bond separation energies and the aptitude for PQCs for predicting molecular wave functions.

- Abstract
- Introduction
- Methods
- Datasets
- Results and Discussion
  - Function fitting
    - \* All (5 + 16 qubit)
    - \* RUD + AL (5 + 16 qubit)
    - \* Error mitigation on FakeQuebec
  - BSE
    - \* All 5 qubit
    - \* Truncated 16 qubit set (cost analysis 5 qubit)
    - \* RUD + AL (5 + 16 qubit)
    - \* Real device (without error mitigation and with whatever the best is from function fitting)
  - DDCC (5 qubit only because classically it can be done with 5 features!)
    - \* Truncated 5 qubit set (based on BSE 5 qubit cost analysis)
    - \* RUD + AL (5 qubit)
    - \* Real device (without error mitigation and with whatever the best is from function fitting)
- Conclusion

# 1 Introduction

Within the field of chemistry, tools such as machine learning and quantum computing have become popular due to the promise of discovering new molecules and materials. Machine learning in chemistry has already become a broadly applicable tool for exploring chemical compound space,<sup>1,2</sup> accelerating wave functions,<sup>3,4</sup> and for applications, such as catalysis and drug discovery.<sup>5?–8</sup> [Sentence about quantum computing] Another popular avenue, which combines quantum computing and machine learning, is quantum machine learning (QML). QML utilizes quantum algorithms during some portion of the machine learning process in an attempt to outperform classical algorithms.<sup>9</sup> On the current era of quantum computers, parameterized quantum circuits (PQCs) offer an example of QML that is performed using a hybrid quantum-classical feedback loop, where the variational parameters are optimized classical, while the circuit is ran on quantum hardware.<sup>10,11</sup>

PQCs are often composed using two different subcircuits: the encoding layer, which is used to encode the features into the quantum circuit, and variational layers, which are composed of parameterized gates which can be optimized classical.

PQCs are composed of data-encoding and variational subcircuits, which can be repeated multiple times.

While several examples exist for PQCs applied to chemical applications, they are largely relegated to drug discovery<sup>11–18</sup> and experimental properties.<sup>19</sup> Herein, we explore two different chemical datasets: one consisting of bond separation energies of 49 unique bond types, called the BSE49 dataset,<sup>20</sup> and another which utilizes data-driven coupled cluster (DDCC)<sup>3</sup> to predict the coupled-cluster singles and doubles wave function parameters using lower-level quantum chemistry methods.

In this study, we create a robust Python code base for exploring a set of 14 data-encoding and 12 variational (ansatz) subcircuits for a total of 168 combined PQCs. Our work includes the single and double encoding layers found in the paper of Suzuki and Katouda,<sup>11</sup> including the encoding of Mitarai *et al.*,<sup>21</sup> along with using instantanious quantum polynomial (IQP) as an

encoding layer.<sup>25</sup> All of our variational (ansatz) layers can be found in the study of Sim *et al.*,<sup>22</sup> where the expressibility and entanglement capability of these circuits is analyzed. We also examine the effects of the number of re-upload and ansatz layers, the effects of feature set size, and various error mitigation techniques.

<sup>23</sup>

QML:<sup>9</sup>

- “Quantum machine learning software makes use of quantum algorithms as part of a larger implementation. By analysing the steps that quantum algorithms prescribe, it becomes clear that they have the potential to outperform classical algorithms for specific problems (that is, reduce the number of steps required). This potential is known as quantum speedup.”
- “The notion of a quantum speedup depends on whether one takes a formal computer science perspective—which demands mathematical proofs—or a perspective based on what can be done with realistic, finite-size devices—which requires solid statistical evidence of a scaling advantage over some finite range of problem sizes. For the case of quantum machine learning, the best possible performance of classical algorithms is not always known.”
- DO WE SEE THIS: “Determination of a scaling advantage contrasting quantum and classical machine learning would rely on the existence of a quantum computer and is called a ‘benchmarking’ problem. Such advantages could include improved classification accuracy and sampling of classically inaccessible systems. Accordingly, quantum speedups in machine learning are currently characterized using idealized measures from complexity theory: query complexity and gate complexity (see Box 1 and Box 1 Table). Query complexity measures the number of queries to the information source for the classical or quantum algorithm. A quantum speedup results if the number of queries needed to solve a problem is lower for the quantum algorithm than for the classical algorithm. To determine the gate complexity, the number of elementary quantum operations (or gates) required to obtain the desired result are counted.”

- Input problem: “Classical data must be input before being processed on a quantum computer. This ‘input problem’ often has little overhead but can present a serious bottleneck for certain algorithms. Likewise, the ‘output problem’ is faced when reading out data after being processed on a quantum device. Like the input problem, the output problem often causes a noticeable operational slowdown.”
- QML challenges: “These hardware challenges are technical in nature, and clear paths exist towards overcoming them. They must be overcome, however, if quantum machine learning is to become a ‘killer app’ for quantum computers. As noted previously, most of the quantum algorithms that have been identified face a number of caveats that limits their applicability. We can distill the caveats mentioned above into four fundamental problems. (1) The input problem. Although quantum algorithms can provide dramatic speedups for processing data, they seldom provide advantages in reading data. This means that the cost of reading in the input can in some cases dominate the cost of quantum algorithms. Understanding this factor is an ongoing challenge. (2) The output problem. Obtaining the full solution from some quantum algorithms as a string of bits requires learning an exponential number of bits. This makes some applications of quantum machine learning algorithms infeasible. This problem can potentially be sidestepped by learning only summary statistics for the solution state. (3) The costing problem. Closely related to the input/output problems, at present very little is known about the true number of gates required by quantum machine learning algorithms. Bounds on the complexity suggest that for sufficiently large problems they will offer huge advantages, but it is still unclear when that crossover point occurs. (4) The benchmarking problem. It is often difficult to assert that a quantum algorithm is ever better than all known classical machine algorithms in practice because this would require extensive benchmarking against modern heuristic methods. Establishing lower bounds for quantum machine learning would partially address this issue.”

PQCs:<sup>10</sup>

- “Parameterized quantum circuits (PQCs) offer a concrete way to implement algorithms and demonstrate quantum supremacy in the NISQ era. PQCs are typically composed of fixed gates, e.g. controlled NOTs, and adjustable gates, e.g. qubit rotations. Even at low circuit depth, some classes of PQCs are capable of generating highly non-trivial outputs. For example, under well-believed complexity-theoretic assumptions, the class of PQCs called instantaneous quantum polynomial-time cannot be efficiently simulated by classical resources (see Lund et al [3] and Harrow and Montanaro [4] for accessible Reviews of quantum supremacy proposals). The demonstration of quantum supremacy is an important milestone in the development of quantum computers. In practice, however, it is highly desirable to demonstrate a quantum advantage on applications. The main approach taken by the community consists in formalizing problems of interest as variational optimization problems and use hybrid systems of quantum and classical hardware to find approximate solutions. The intuition is that by implementing some subroutines on classical hardware, the requirement of quantum resources is significantly reduced, particularly the number of qubits, circuit depth, and coherence time. Therefore, in the hybrid algorithmic approach NISQ hardware focuses entirely on the classically intractable part of the problem.”

PQCs:<sup>11</sup>

- “More recently, quantum machine learning (QML) [26–30] is a rapidly growing research field that combines near-term quantum algorithms and machine learning techniques. In particular, parameterized quantum circuits (PQCs) have been considered as machine learning models with high expressive power within the hybrid quantum–classical framework [31, 32]. PQCs are typically composed of fixed quantum gates (e.g., qubit rotations and entangling gates) in a shallow circuit layout, with variable parameters optimized in a classical feedback loop.”
- “To our knowledge, however, the application of QML to regression tasks has not been fully investigated in the literature. It remains unclear what kinds of quantum states should be used

in order to generate the feature map with high expressibility that is suited for real-world data sets.”

- “quantitative structure–toxicity relationship (QSTR) models for predicting the toxicity of 221 phenols. While there are a variety of QSAR/QSTR models (e.g., 3D-QSAR [1, 4]), as a first step we employ QSAR/QSTR models including molecular descriptors such as hydrophobicity, acidity constant, and frontier orbital energies. There have been quantum computations in biochemical and pharmaceutical areas, such as protein folding [43–45], molecular similarity [46], and biological data [47]; yet, there has been no study on quantum application to QSAR modeling, albeit an important part of ligand-based computer-aided drug design.”
- “PQCs have been regarded as machine learning models with high expressive power within the framework of the hybrid quantum–classical approach. PQCs are usually composed of one-qubit rotations and two-qubit entangling operations in a shallow circuit layout, with parameters optimized in a feedback loop.”
- “Combining near-term quantum algorithms and machine learning, QML using the framework of PQCs is sometimes referred to as quantum circuit learning (QCL)”
- “PQCs consist of three components: the encoder circuit, the variational circuit, and the measurement for the estimation of the loss function”

PQCs:<sup>19</sup>

- What they do: “introduce quantum circuit learning (QCL) as an emerging regression algorithm for chemo- and materials-informatics. The supervised model, functioning on the rule of quantum mechanics, can process linear and smooth non-linear functions from small datasets (<100 records). Compared with conventional algorithms, such as random forest, support vector machine, and linear regressions, the QCL can offer better predictions with some one-dimensional functions and experimental chemical databases”

- “Here, we introduce quantum circuit learning (QCL), an emerging algorithm for supervised learning.<sup>15–18</sup> QCL works on the rule of quantum mechanics. It can predict various parameters likewise to classical models. The current quantum systems (noisy intermediate-scale quantum computers: NISQ)<sup>19,20</sup> face the problems of calculation noise and the limited processable number of information units (qubits).”
- “Since QCL is a frontier for machine learning, few reports have been published on authentic regression tasks. The success of prediction with a toxicity dataset of organic molecules was reported,<sup>17</sup> whereas the study was still conceptual. The prediction processes and advantages have been unclear, especially from chemical and material viewpoints. Here, we conducted a more comprehensive study on QCL with standard datasets of one-dimensional functions and chemical properties. Both simulated and actual quantum computing was undertaken to clarify the challenges of QCL. Various hyperparameters were optimized for higher accuracy. The comparison with conventional models contrasted the benefits of the new approach: capable of learning both linear and non-linear functions even from small datasets. The property was also favorable for predicting the so-called extrapolating data region, which is essential for chemical and material research.”
- They cite<sup>11,21,24</sup> as examples of PQCs/quantum circuit learning for regression task
- Justification for starting with 1D regression: “Preliminary optimization of quantum circuits<sup>15,17</sup> and our optimization revealed that the following configuration was concise and practical for regression tasks (for details, see results and explanations in Fig. S1”

## 2 Methods

PQCs are often constructed of three parts: encoding layers that are used to encode the features onto a quantum circuit, variational layers which include parameters that are optimized classically, and measurements which provide numerical estimations of the regression target values.<sup>11</sup> In this

study, we utilize the Mitarai ( $M$ ),<sup>21</sup> single- ( $A1$ ) and double-angle ( $A2$ ) encoding layers found in Ref.<sup>11</sup>, along with the instantaneous quantum polynomial (IQP) circuit found in Refs.<sup>25</sup> and<sup>26</sup>. In the following section, we follow the notations derived from Ref.<sup>11</sup>. Encoding layers work mapping a  $d$ -dimensional feature vector,  $\mathbf{x} = (x_1, x_2, \dots, x_d)^T \in \mathbb{R}^d$ , normalized on the range  $[-1, 1]$ , onto a quantum circuit using a unitary matrix, denoted as  $U_{\Phi(\mathbf{x})}$ , to produce the quantum state  $U_{\Phi(\mathbf{x})}|0\rangle^{\otimes n}$ , where  $n$  are the number of qubits. The encoding layer takes the following general form,

$$U_{\Phi(\mathbf{x})} = \prod_l E_{\text{ent}}^l U_{\phi_l(\mathbf{x})} \quad (1)$$

where,  $E_{\text{ent}}^l$  denotes the entangling gates, which can be a CNOT, CZ, or identity ( $I$ ) gates,  $U_{\phi_l(\mathbf{x})}$  denotes the choice of encoding unitaries. Like in Ref.<sup>11</sup>, we choose  $l \in \{1, 2\}$ , such that when  $l = 1$ ,  $E_{\text{ent}}$  corresponds to an identity matrix and  $U_{\Phi(\mathbf{x})} = U_{\phi_1(\mathbf{x})}$ .

When  $l = 1$ ,  $U_{\Phi(\mathbf{x})}$  can be one of the following four encoding layers:  $U_{A1}$ ,  $U_{A2}$ ,  $U_M$ , or  $U_{\text{IQP}}$ . The single-angle encoding (Fig. 1 (a)) is the simplest and takes the following form,

$$U_{A1} = \prod_{i=0}^n R_i^Y(x_i), \quad (2)$$

where  $R_i^Y$  denotes a parameterized Y rotation gate on qubit  $i$ . Like the single-angle encoding, the double-angle encoding (Fig. 1 (b)) utilizes a parameterized Y rotation gate on qubit  $i$ , with the addition of a parameterized Z rotation gate on qubit  $i$ , denoted as

$$U_{A2} = \prod_{i=0}^n R_i^Z(x_i) R_i^Y(x_i). \quad (3)$$

The Mitarai encoding layer (Fig. 1 (c)) is a double-angle encoding layer with the addition of an arccosine function on the parameterized Z gate and arcsine on the parameterized Y gate,

$$U_M = \prod_{i=0}^n R_i^Z(\arccos(x_i^2)) R_i^Y(\arcsin(x_i^2)). \quad (4)$$

Following the formulation provided in the PennyLane<sup>27</sup> software package, the IQP encoding layer (Fig. 1 (d)) is defined as,

$$U_{\text{IQP}} = \prod_{i=0}^n H_i R_i^Z(x_i) \prod_{i < j} ZZ_{ij}, \quad (5)$$

where  $H_i$  denotes a Hadamard gate on qubit  $i$  and  $ZZ_{ij}$  denotes a two-qubit entanglement gate defined as  $ZZ_{ij} = e^{-ix_i x_j \sigma_z \otimes \sigma_z}$ .

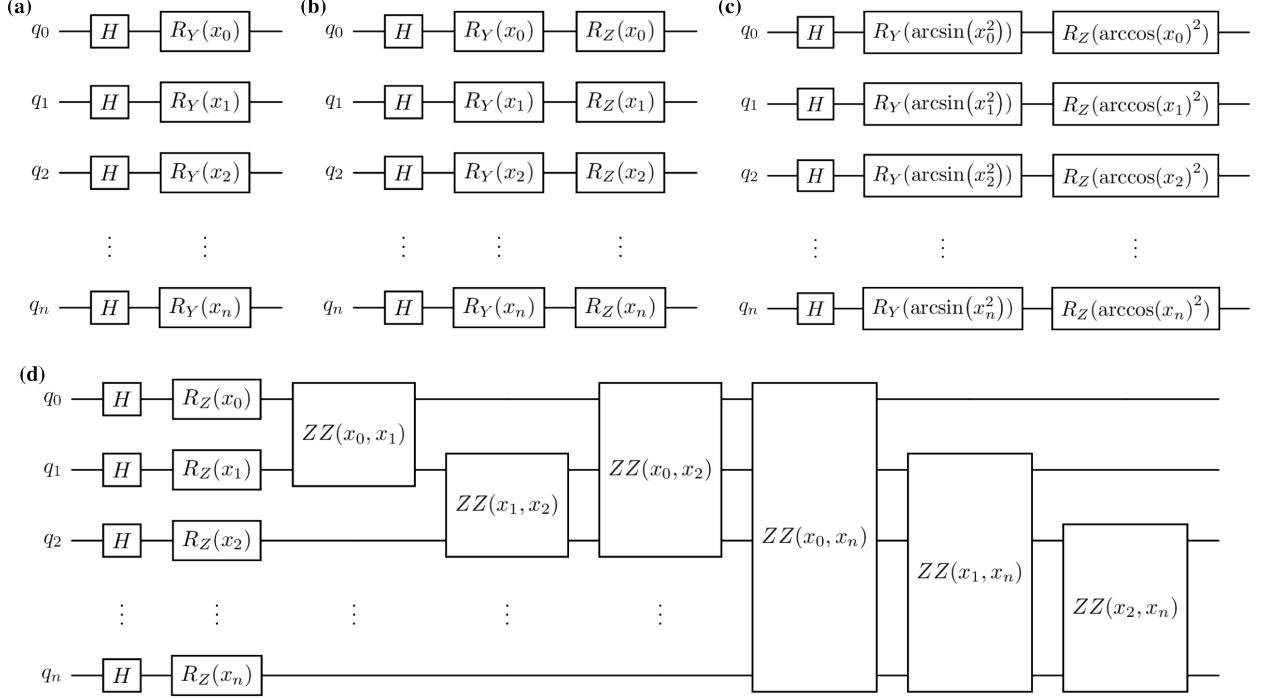


Figure 1: (a) Single angle (A1) encoding, (b) double angle (A2) encoding, (c) Mitarai (M) encoding, and (d) Instantaneous Quantum Polynomial (IQP) encoding

When  $l = 2$ , like in Ref.<sup>11</sup> we choose entanglement gates,  $E_{\text{ent}}^1$  and  $E_{\text{ent}}^2$  to be equivalent, and the encoding layer takes the following form,  $U_{\Phi(x)} = E_{\text{ent}} U_{\phi_2(\mathbf{x})} E_{\text{ent}} U_{\phi_1(\mathbf{x})}$ . We also exclude IQP encoding when  $l = 2$  due to the increased circuit depth, when compared to A1, A2, and M encoding. Therefore, there are five unique combinations of  $U_{\phi_1(\mathbf{x})}$  and  $U_{\phi_2(\mathbf{x})}$  (M-M, A1-A1, A2-A2, M-A1, and M-A2) and two different entanglement layer options (CNOT and CZ) for a total of 10 encoding circuits. These circuits are denoted as  $U_{\phi_1(\mathbf{x})} - U_{\phi_2(\mathbf{x})} - E_{\text{ent}}$ , for example, two example encoding circuits are M-M-CNOT and M-A1-CNOT. Table 1 shows all fourteen encoding

circuits examined in this study.

Table 1: Add something smart

Name	$U_{\phi_1(\mathbf{x})}$	$U_{\phi_2(\mathbf{x})}$	$E_{\text{ent}}$
A1	$U_{A1}$	—	—
A2	$U_{A2}$	—	—
M	$U_M$	—	—
IQP	$U_{\text{IQP}}$	—	—
A1–A1–CNOT	$U_{A1}$	$U_{A1}$	$E_{\text{CNOT}}$
A2–A2–CNOT	$U_{A2}$	$U_{A2}$	$E_{\text{CNOT}}$
M–M–CNOT	$U_M$	$U_M$	$E_{\text{CNOT}}$
M–A1–CNOT	$U_M$	$U_{A1}$	$E_{\text{CNOT}}$
M–A2–CNOT	$U_M$	$U_{A2}$	$E_{\text{CNOT}}$
A1–A1–CZ	$U_{A1}$	$U_{A1}$	$E_{\text{CZ}}$
A2–A2–CZ	$U_{A2}$	$U_{A2}$	$E_{\text{CZ}}$
M–M–CZ	$U_M$	$U_M$	$E_{\text{CZ}}$
M–A1–CZ	$U_M$	$U_{A1}$	$E_{\text{CZ}}$
M–A2–CZ	$U_M$	$U_{A2}$	$E_{\text{CZ}}$

Following the encoding layers, variational (or ansatz) layers are used to introduce trainable parameters into the quantum circuit. We use a mixed notation from Refs.<sup>11</sup> and<sup>22</sup>, since Ref.<sup>22</sup> contains all of the variational layers used within this work. We relegate the discussion of the expressibility and entanglement examined in that work to Section 3. A general variational layer can be denoted as,

$$U(\boldsymbol{\theta}) = \prod_v U_v(\boldsymbol{\theta}_v), \quad (6)$$

where  $\boldsymbol{\theta}$  denotes the variational parameters and  $v$  denotes the number of times that the layer is repeated within the circuit. As  $v$  increases and the number of trainable parameters ( $\boldsymbol{\theta}$ ) increase, the theoretical assumption is that the model expressibility should also increase. In our study, we choose  $v \in \{1, 3, 5\}$  and refer to this as the number of ansatz layers (ALs). We examine 12 different variational circuits, as shown in Fig. 2, which are denoted using the following labels: Modified-Pauli-CRZ (Fig. 2(a)), Modified-Pauli-CRX (Fig. 2(b)), Efficient-CRZ (Fig. 2(c)), Efficient-CRX (Fig. 2(d)), HWE-CNOT (Fig. 2(e)), HWE-CZ (Fig. 2(f)), ESU2 (Fig. 2(g)), Full-Pauli-CRZ (Fig. 2(h)), Full-Pauli-CRX (Fig. 2(i)), Hadamard (Fig. 2(j)), Full-CRZ (Fig. 2(k)), and Full-

CRX (Fig. 2(l)),

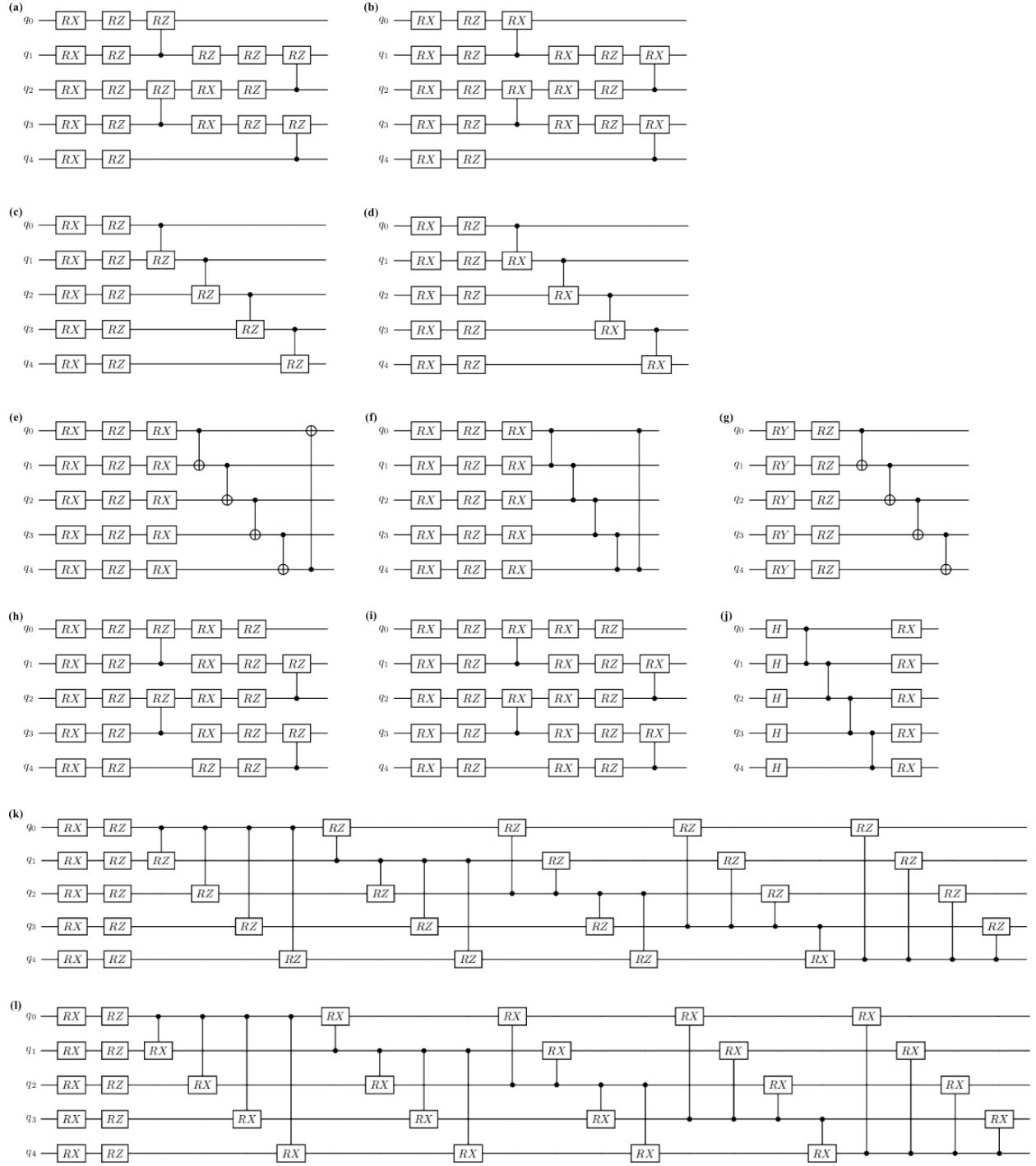


Figure 2: (a) Modified-Pauli-CRZ, (b) Modified-Pauli-CRX, (c) Efficient-CRZ, (d) Efficient-CRX, (e) HWE-CNOT, (f) HWE-CZ, (g) ESU2, (h) Full-Pauli-CRZ, (i) Full-Pauli-CRX, (j) Hadamard, (k) Full-CRZ, and (l)Full-CRX

Now that we have define the encoding (Eq. 1) and variational (Eq. 6) circuits, we can then

combine them to denote a general, complete circuit as,

$$|\Psi\rangle = U(\theta)U_{\Phi(\mathbf{x})}|0\rangle^{\otimes n} = \prod_k \left( \prod_v U_v(\theta_v) \prod_l E_{\text{ent}}^l U_{\phi_l(\mathbf{x})} \right) |0\rangle^{\otimes n}, \quad (7)$$

where we choose  $k \in \{1, 3, 5\}$ , which denotes the re-upload depth (RUD) of the circuit. When a sufficient number of data re-uploading occur, it has been shown by Pérez-Salinas *et al.* that data re-uploading is equivalent to the Universal Approximation Theorem for artificial neural networks.<sup>28</sup>

Lastly, to recover the predicted target values,  $\hat{y}_i$ , from our quantum circuits, measurement of the quantum state,  $|\Psi\rangle$ , must be performed. To perform this operation, we apply the Pauli Z operator on the first qubit denoted as,

$$\hat{y}_i = \langle \Psi | Z_0 | \Psi \rangle_i. \quad (8)$$

The set of predicted target values,  $\hat{\mathbf{y}} = (\hat{y}_1, \dots, \hat{y}_N) \in \mathbb{R}^N$ , where  $N$  is the number of samples, is then passed to the loss function,  $\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}})$ , where  $y_i$  belongs to the set of true target values  $\mathbf{y} = (y_1, \dots, y_N) \in \mathbb{R}^N$ . In practice,  $\mathcal{L}$  can be any loss function but we choose to use the mean square error loss function denoted as,

$$\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2. \quad (9)$$

## Implementation

We perform all simulation calculations using PennyLane,<sup>27</sup> either using Qulacs<sup>29</sup> for state vector calculations, while noisy calculations were performed using *qiskit-aer* with the *FakeQuebec* backend as implemented in the PennyLane-Qiskit plugin.<sup>30</sup> We perform calculations in the *ibm\_quebec* device using circuits implemented using Qiskit,<sup>30</sup> due to issues we initially faced with running experiments on the real device using the PennyLane-Qiskit plugin. For the experiments using PennyLane, we utilize the Simultaneous Perturbation Stochastic Approximation method (SPSA) as implemented in PennyLane, while for the experiments run on *ibm\_quebec* utilizes the Constrained Optimization By Linear Approximation (COBYLA) optimizer as implemented in SciPy.<sup>31</sup> Each

optimizer was chosen based on the performance for the given task. All features ( $\mathbf{x}$ ) and target values ( $\mathbf{y}$ ) were scaled using the MinMaxScaler in Scikit-learn,<sup>32</sup> such that all features and target values are  $\mathbb{R} \in [-1, 1]$ . For the simulations using *FakeQuebec* and experiments on *ibm\_quebec* we utilize Twirled Readout Error eXtinction (TREX) error mitigation.

## Dataset

In this study, we explore 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;<sup>3</sup> and a dataset of bond separation energies (BSE) of molecules, where the feature set encodes structural information of each molecule.

The noisy function fitting dataset is generated by  $f(x) = x + \varepsilon$  such that  $\varepsilon \in \mathcal{N}(0, \sigma^2)$ , where  $\sigma \approx 0.34$ . The linear function is  $f(x) = x$ , quadratic is  $f(x) = 0.3967x^2 - 3.2807x + 6.7569$ , and sine function is  $f(x) = 0.9913 \sin(2.3838x + 0.0344)$ . The normalized versions are shown in Fig. 3.

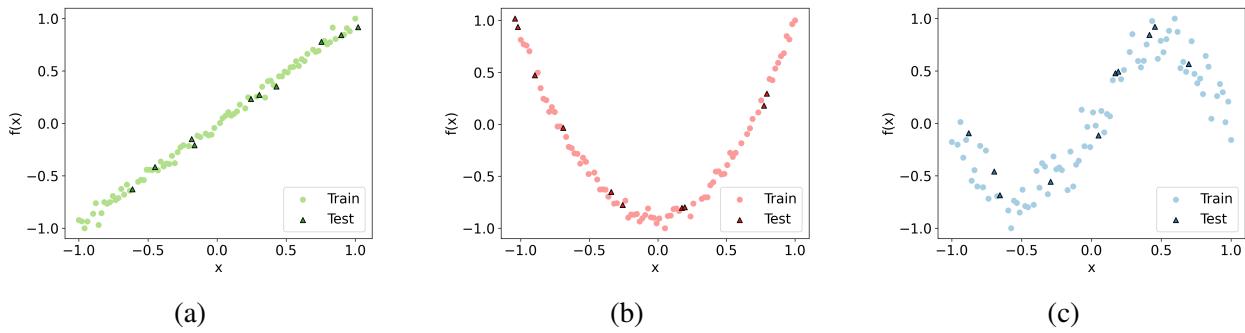


Figure 3

33 34

The first chemical dataset we explore is the BSE49 dataset which consists of homolytic bond cleavages, such as  $A-B \longrightarrow A\cdot + B\cdot$ , of X molecules with 49 unique bond types. One important aspect of this dataset is the choice of molecular representation, or how the molecule is represented in the machine learning models. To find the optimal molecular representation, we examined five

different molecular representations including Coulomb Matrices,<sup>33</sup> Molecular ACCess Systems fingerprint (MACCS),<sup>35</sup> persistence images (PIs),<sup>36–38</sup> RDKit [cite], and Morgan or extended-connectivity fingerprints.<sup>34,39</sup> For each representation we tested two different choices of representation, one based on the products minus the reactants, denoted by *sub*, similar to the method used in [cite], and one based on just the reactants, denoted as *AB*. We use the same 80% train/20% test split as we use throughout our studies. We found that the best molecular representation across all models test, as shown in Fig. 4a, was Morgan fingerprints using the *sub* formulation.

One challenge of applying classical molecular representations for quantum machine learning models is mapping the classical features, often containing hundreds or thousands of features per sample, to the number of qubits used. Using the *sub*-Morgan fingerprints, we explore mapping these features to 5 and 16 qubits, the number of qubits used throughout this study, using SHapley Additive ExPlanation analysis (SHAP)<sup>40</sup> and principal component analysis (PCA). We found that the most accurate models, using both 5 and 16 features, were those generated by PCA, as shown in Fig. 4c and 4b.

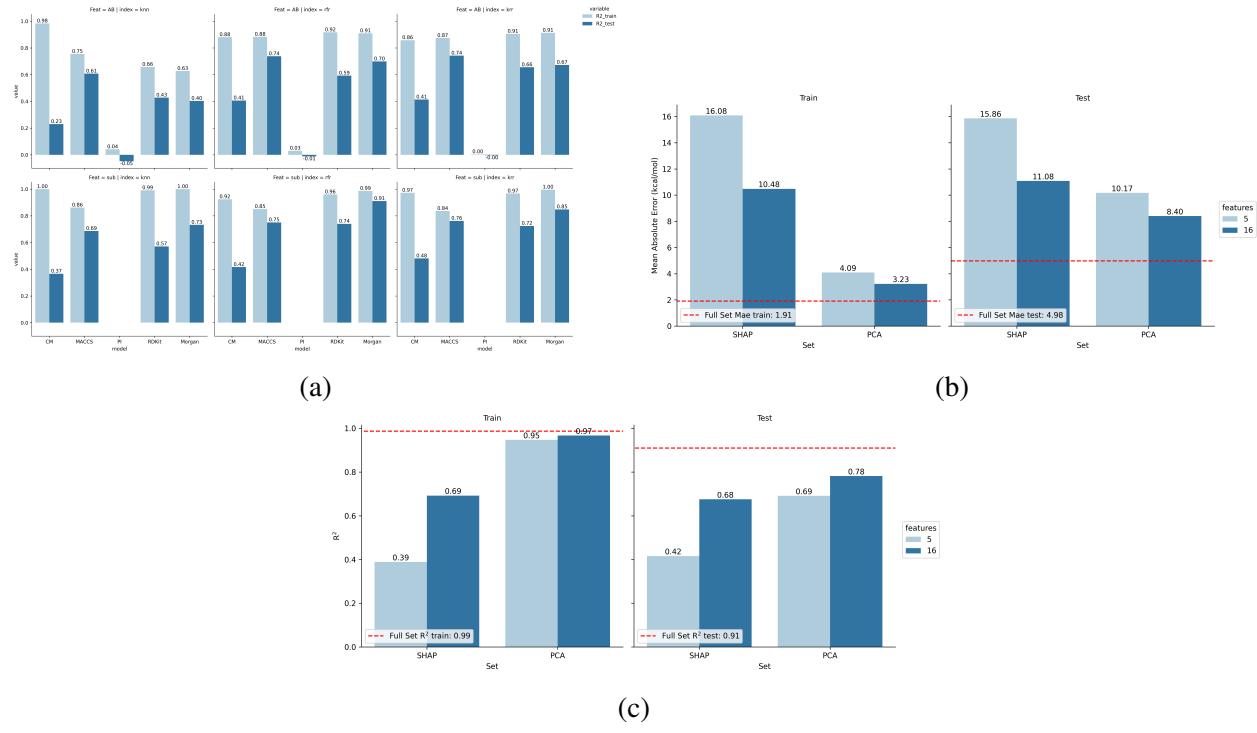


Figure 4

The next chemical dataset explored is based on the Data-Driven Coupled-Cluster (DDCC) methods found in Ref.<sup>3,4</sup> DDCC is a machine learning-based method for accelerating coupled-cluster singles and doubles (CCSD) calculations by predicting the CCSD T2-amplitudes using features generated by a lower-level method, typically Hartree-Fock (HF) and Møller-Plesset second-order perturbation theory (MP2). We generate a database consisting X water molecules with y features for each molecule, using Psi4<sup>41</sup> and Psi4Numpy.<sup>42</sup>

Like with the BSE dataset, we explore how we can reduce the initial y features down to 5 and 16 features using SHAP and PCA.

$$t_{ij(\text{MP2})}^{ab} = \frac{\langle ij||ab\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \quad (10)$$

For each  $t_{ij(\text{CCSD})}^{ab}$ , orbital energies ( $\varepsilon_i, \varepsilon_j, \varepsilon_a, \varepsilon_b$ ), Coulomb and exchange integrals ( $J_a^i, J_b^j, K_i^a, K_j^b$ ), 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  $\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b$ .

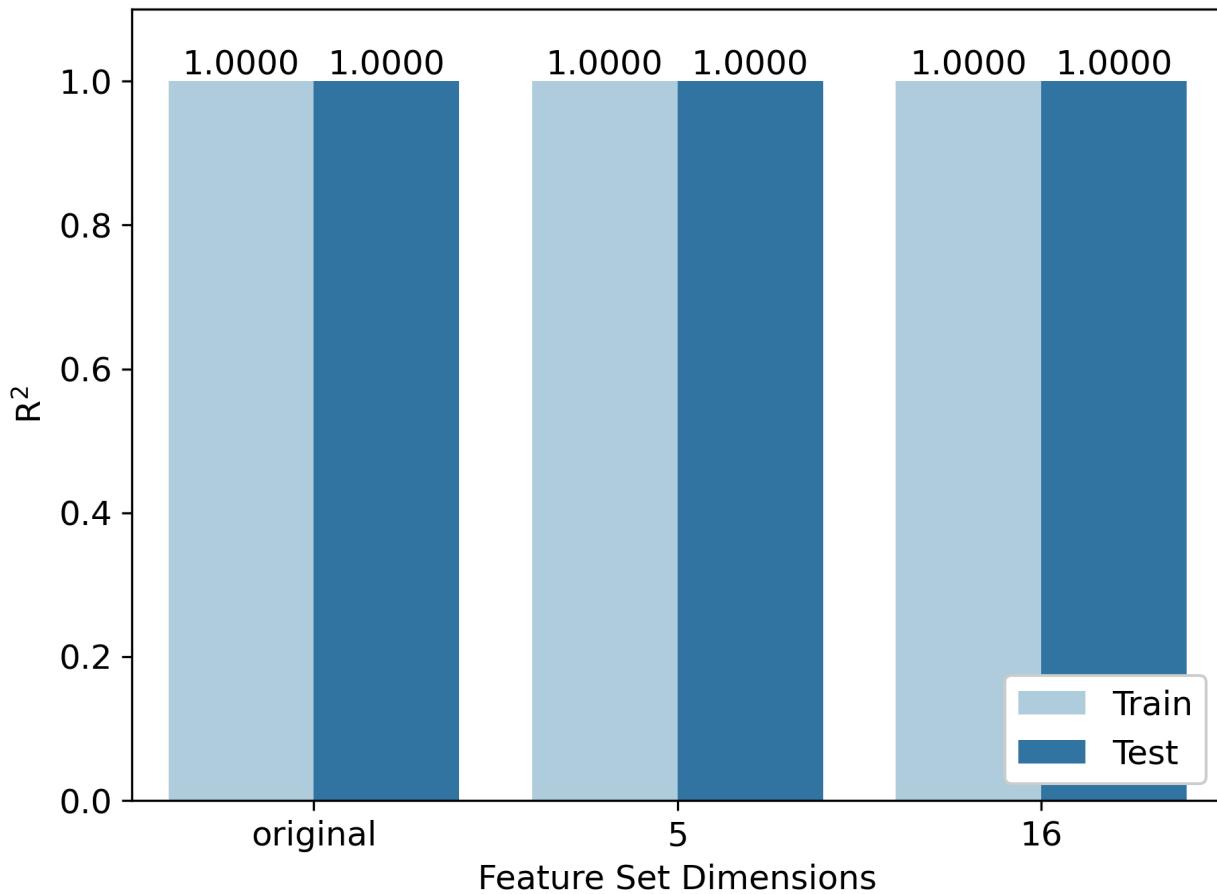


Figure 5

We utilize the BSE49 and DDCC databases for two different reasons: the BSE49 database consists of a hard chemical property to predict using few features, while the DDCC dataset can be predicted easily using few features classically but is data intensive in the sense that for every molecule you will have  $N_{occ}^2 \cdot N_{virt}^2$  entries.

### 3 Results and Discussion

#### Function Fitting

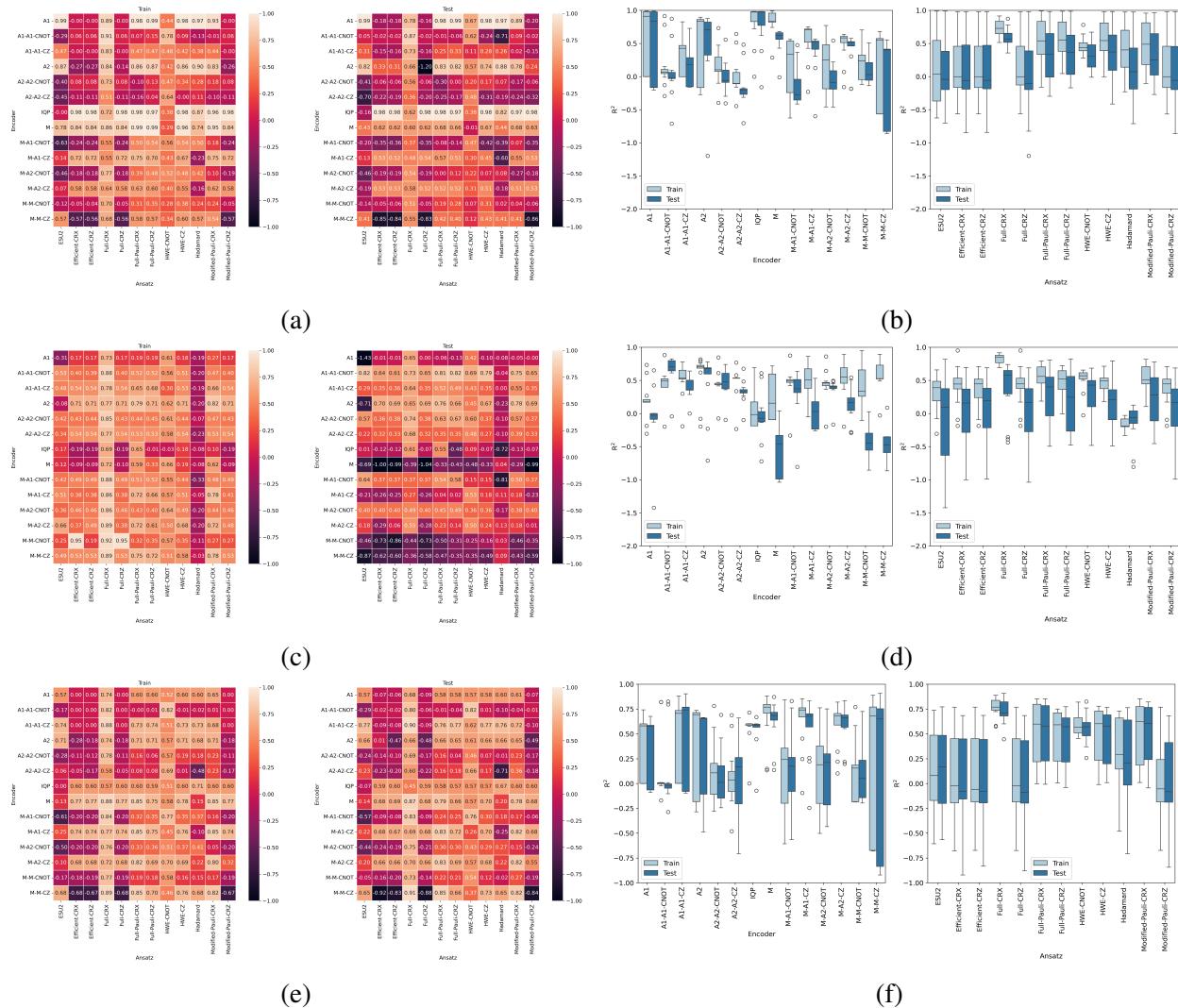


Figure 6

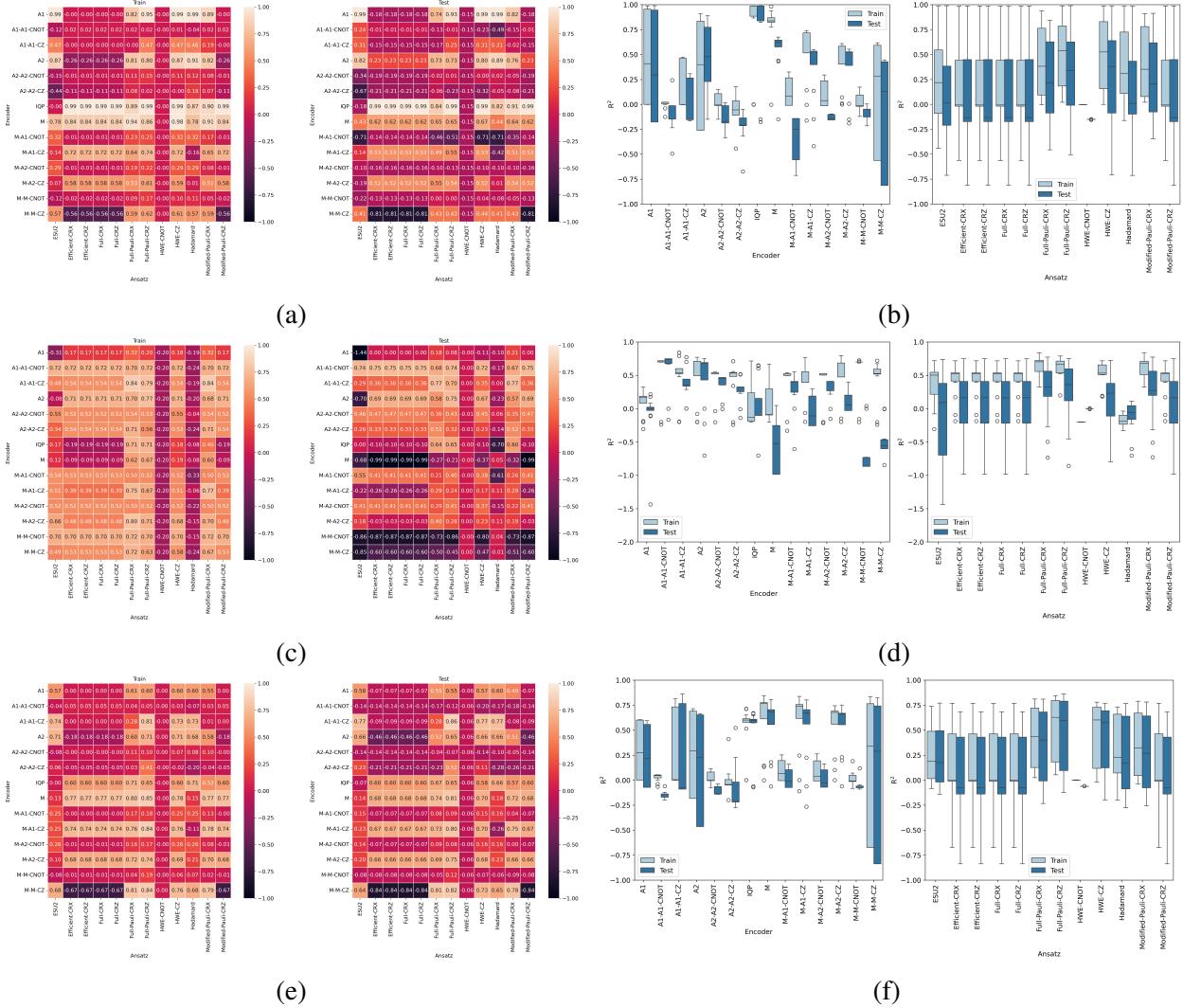


Figure 7

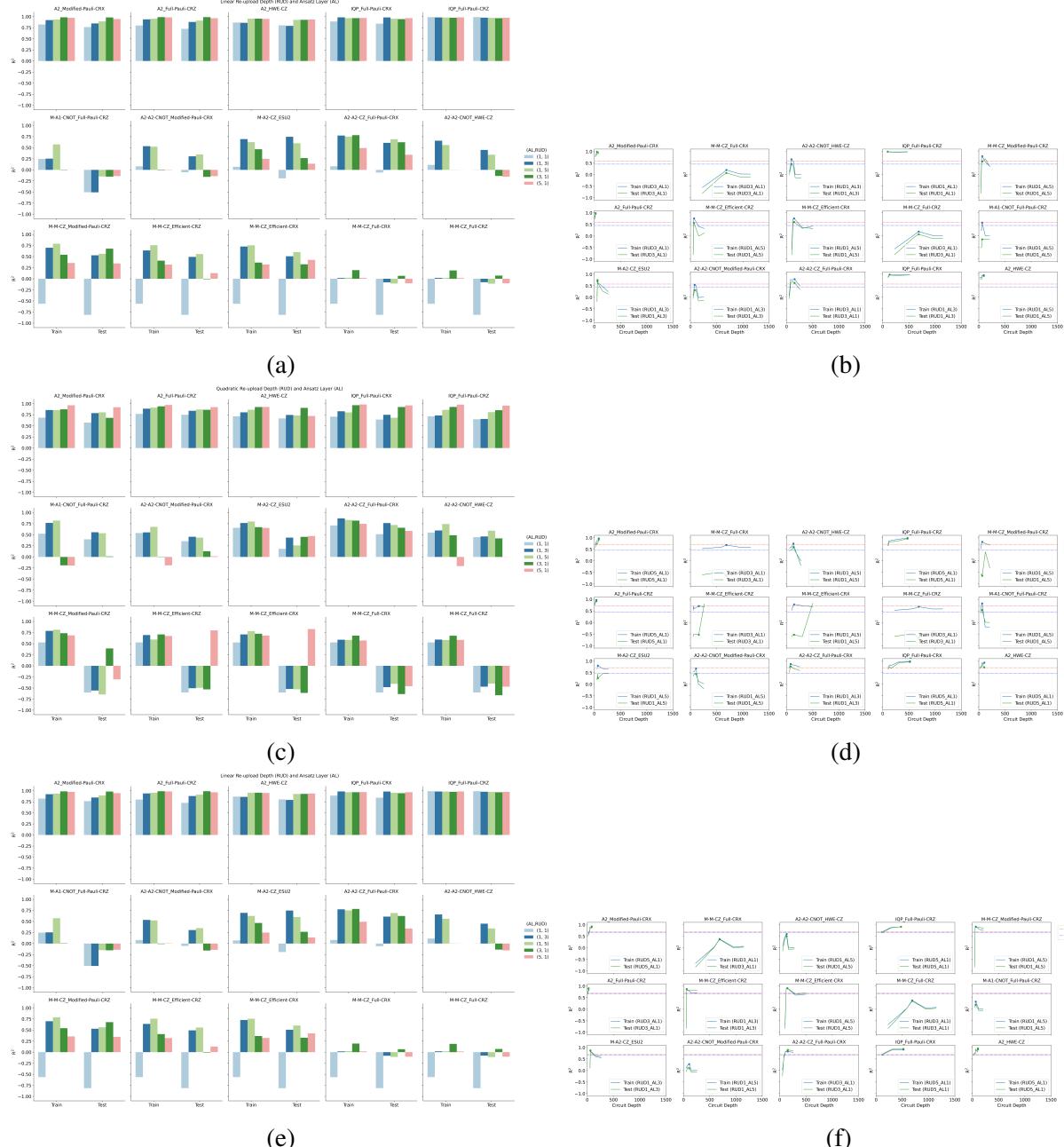


Figure 8

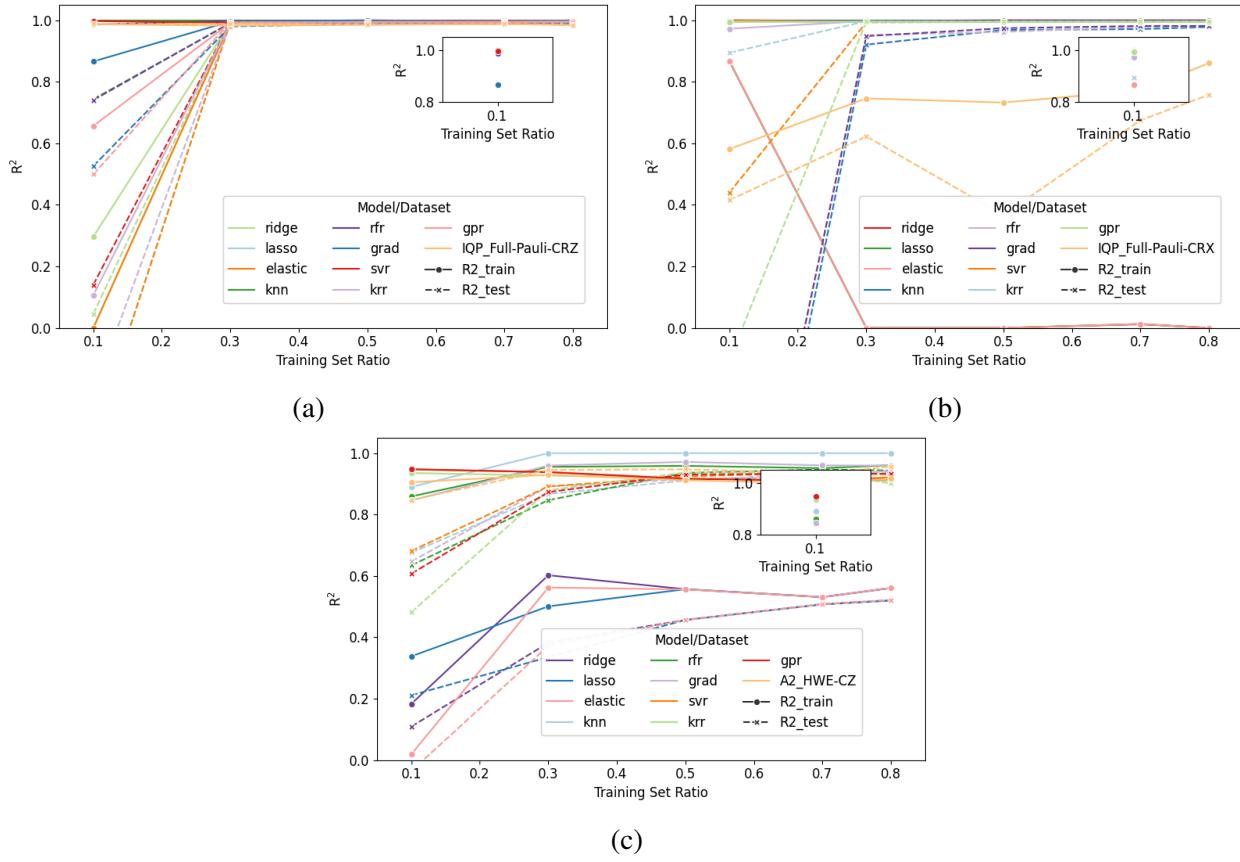


Figure 9

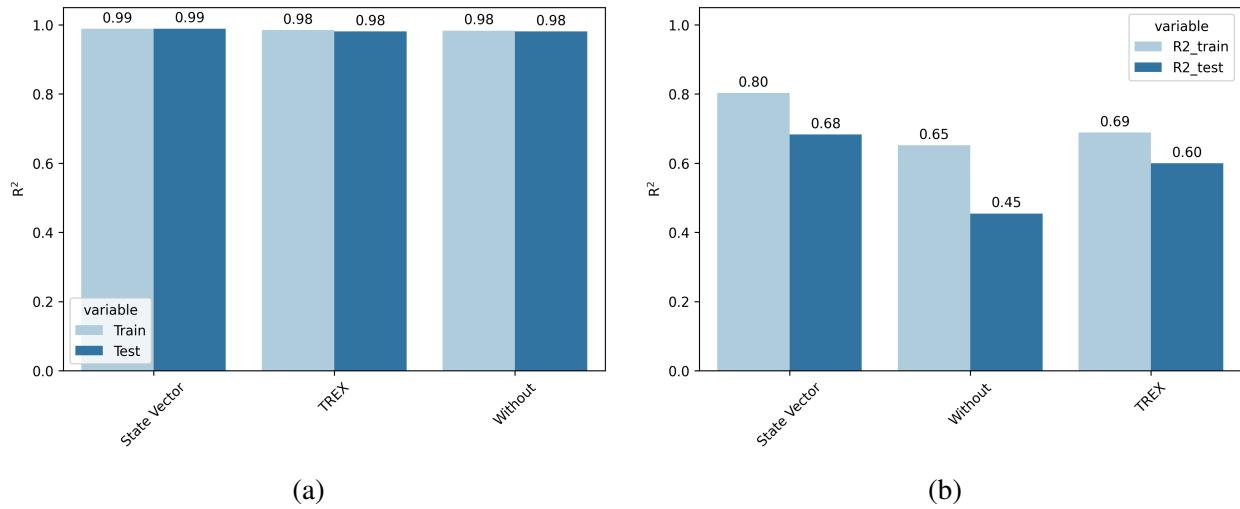


Figure 10

## BSE

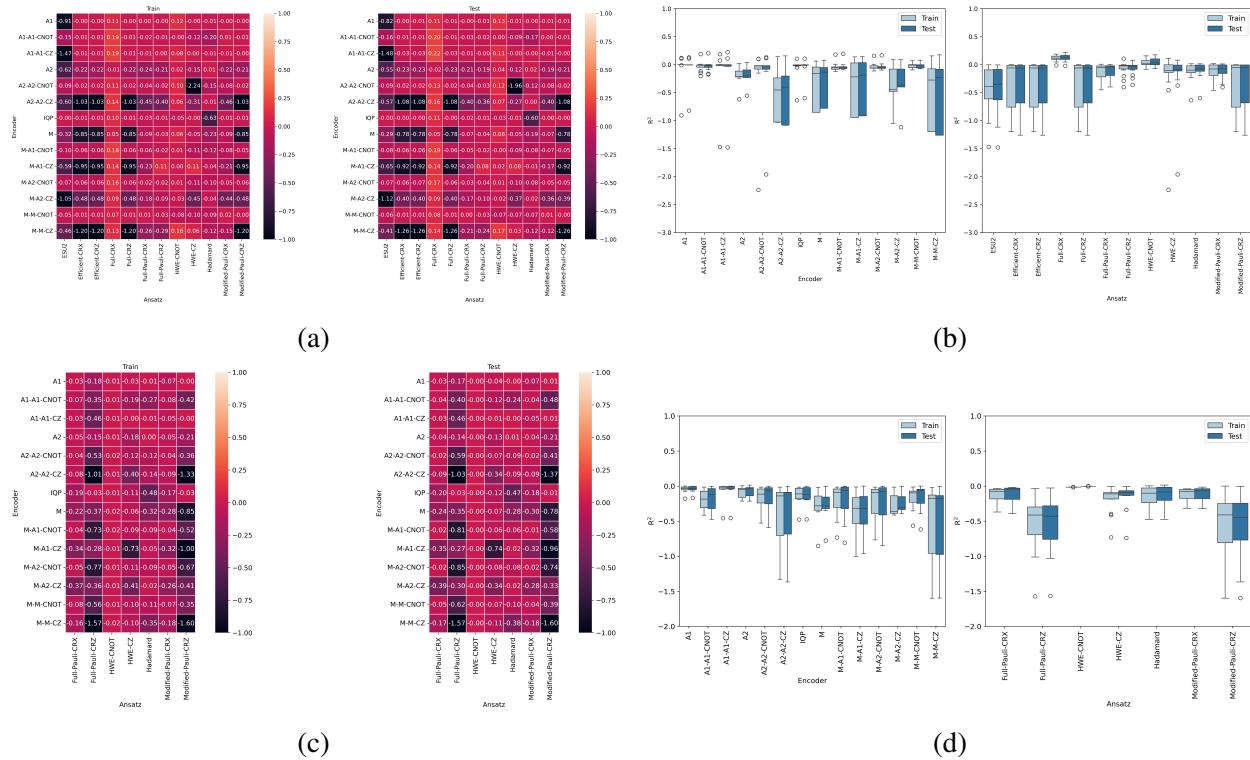


Figure 11

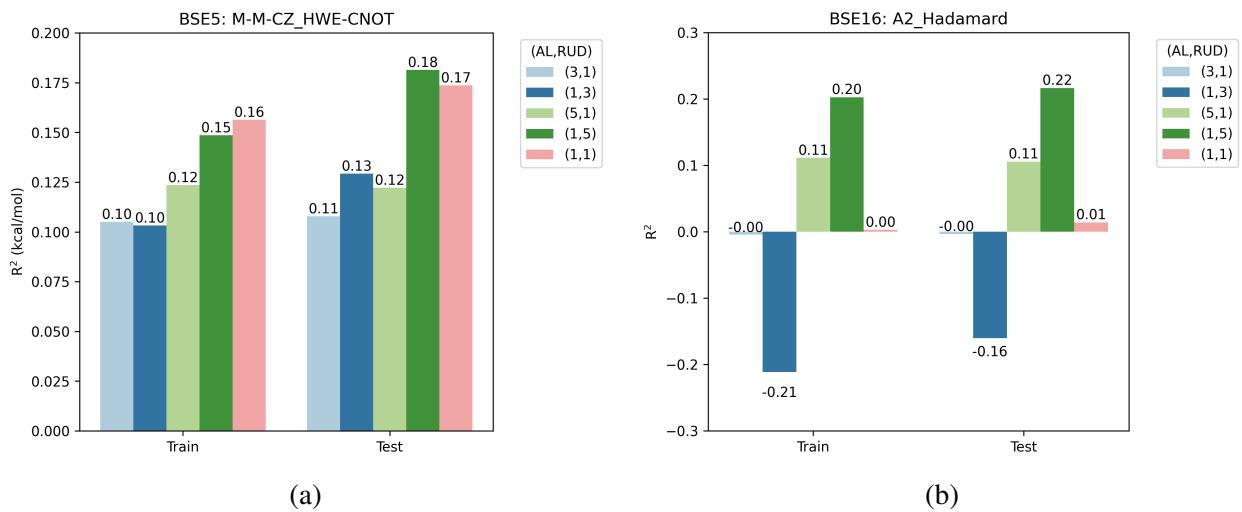


Figure 12

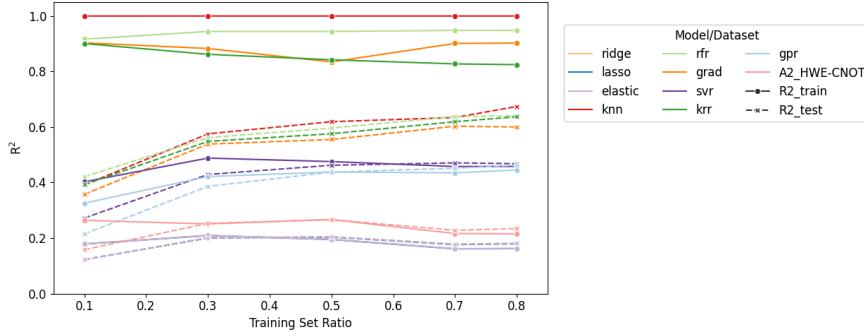


Figure 13

## DDCC

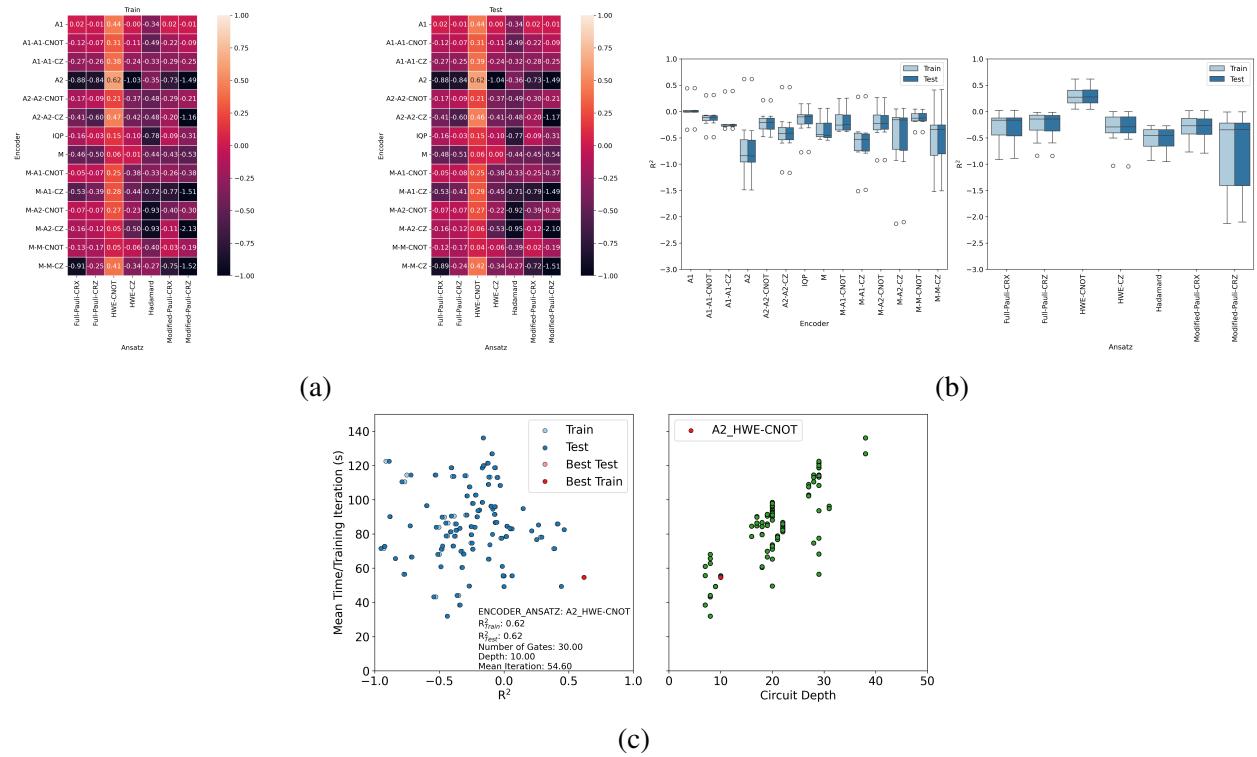


Figure 14

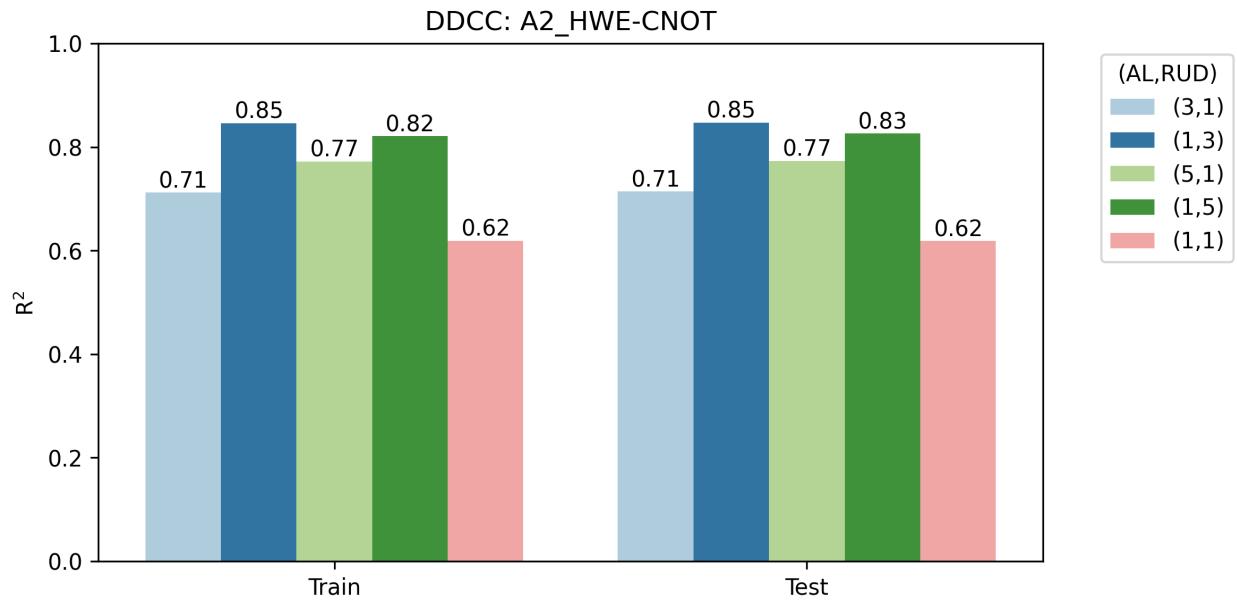


Figure 15

Ansaetze analysis<sup>22</sup> “In particular, a substantial improvement in performance of two-qubit gates in a ring or all-to-all connected arrangement, compared to that of those on a line, is observed.”

“Furthermore, improvement in both descriptors is achieved by sequences of controlled X-rotation gates compared to sequences of controlled Z-rotation gates.”

“investigated how expressibility “saturates” with increased circuit depth, finding that the rate and saturated value appear to be distinguishing features of a PQC”

## 4 Conclusion

Depth is not always better! Molecular representations specifically for QML Distributed QC to incorporate more features

## References

- (1) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *SCIENTIFIC DATA* **2014**, *1*.

- (2) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The  $\Delta$ -Machine Learning Approach. *JOURNAL OF CHEMICAL THEORY AND COMPUTATION* **2015**, *11*, 2087–2096.
- (3) 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.
- (4) Jones, G. M.; S. Pathirage, P. D. V.; Vogiatzis, K. D. In *Quantum Chemistry in the Age of Machine Learning*; Dral, P. O., Ed.; Elsevier, 2023; pp 509–529.
- (5) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *JOURNAL OF CHEMICAL PHYSICS* **2016**, *145*.
- (6) Goh, G. B.; Hodas, N. O.; Vishnu, A. Deep learning for computational chemistry. *JOURNAL OF COMPUTATIONAL CHEMISTRY* **2017**, *38*, 1291–1307.
- (7) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *NATURE* **2018**, *559*, 547–555.
- (8) Janet, J. P.; Kulik, H. J. *Machine Learning in Chemistry*; ACS In Focus; American Chemical Society, 2020.
- (9) Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202.
- (10) Benedetti, M.; Lloyd, E.; Sack, S.; Fiorentini, M. Parameterized quantum circuits as machine learning models. *Quantum Sci. Technol.* **2019**, *4*, 043001.
- (11) Suzuki, T.; Katouda, M. Predicting toxicity by quantum machine learning. *J. Phys. Commun.* **2020**, *4*, 125012.
- (12) Smaldone, A. M.; Batista, V. S. Quantum-to-Classical Neural Network Transfer Learning Applied to Drug Toxicity Prediction. *J. Chem. Theory Comput.* **2024**, *20*, 4901–4908, Publisher: American Chemical Society.

- (13) Ishiyama, Y.; Nagai, R.; Mieda, S.; Takei, Y.; Minato, Y.; Natsume, Y. Noise-robust optimization of quantum machine learning models for polymer properties using a simulator and validated on the IonQ quantum computer. *Sci Rep* **2022**, *12*, 19003, Publisher: Nature Publishing Group.
- (14) Ranga, D.; Rana, A.; Prajapat, S.; Kumar, P.; Kumar, K.; Vasilakos, A. V. Quantum Machine Learning: Exploring the Role of Data Encoding Techniques, Challenges, and Future Directions. *Mathematics* **2024**, *12*, 3318, Number: 21 Publisher: Multidisciplinary Digital Publishing Institute.
- (15) Alam, M.; Ghosh, S. QNet: A Scalable and Noise-Resilient Quantum Neural Network Architecture for Noisy Intermediate-Scale Quantum Computers. *Front. Phys.* **2022**, *9*, Publisher: Frontiers.
- (16) Avramouli, M.; Savvas, I.; Vasilaki, A.; Garani, G.; Xenakis, A. Quantum Machine Learning in Drug Discovery: Current State and Challenges. Proceedings of the 26th Pan-Hellenic Conference on Informatics. New York, NY, USA, 2023; pp 394–401.
- (17) Avramouli, M.; Savvas, I. K.; Vasilaki, A.; Garani, G. Unlocking the Potential of Quantum Machine Learning to Advance Drug Discovery. *Electronics* **2023**, *12*, 2402, Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.
- (18) Bhatia, A. S.; Saggi, M. K.; Kais, S. Quantum Machine Learning Predicting ADME-Tox Properties in Drug Discovery. *J. Chem. Inf. Model.* **2023**, *63*, 6476–6486, Publisher: American Chemical Society.
- (19) 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.
- (20) Prasad, V. K.; Khalilian, M. H.; Otero-de-la Roza, A.; DiLabio, G. A. BSE49, a diverse,

- high-quality benchmark dataset of separation energies of chemical bonds. *Sci Data* **2021**, *8*, 300, Publisher: Nature Publishing Group.
- (21) Mitarai, K.; Negoro, M.; Kitagawa, M.; Fujii, K. Quantum circuit learning. *Phys. Rev. A* **2018**, *98*, 032309.
- (22) 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.
- (23) Krenn, M.; Landgraf, J.; Foesel, T.; Marquardt, F. Artificial intelligence and machine learning for quantum technologies. *Phys. Rev. A* **2023**, *107*, 010101.
- (24) Takaki, Y.; Mitarai, K.; Negoro, M.; Fujii, K.; Kitagawa, M. Learning temporal data with a variational quantum recurrent neural network. *Phys. Rev. A* **2021**, *103*, 052414.
- (25) Bremner, M. J.; Montanaro, A.; Shepherd, D. J. Average-case complexity versus approximate simulation of commuting quantum computations. *Phys. Rev. Lett.* **2016**, *117*, 080501, arXiv:1504.07999 [quant-ph].
- (26) Havlicek, V.; Córcoles, A. D.; Temme, K.; Harrow, A. W.; Kandala, A.; Chow, J. M.; Gambetta, J. M. Supervised learning with quantum enhanced feature spaces. *Nature* **2019**, *567*, 209–212, arXiv:1804.11326 [quant-ph].
- (27) Bergholm, V. et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. 2022; <http://arxiv.org/abs/1811.04968>, arXiv:1811.04968 [quant-ph].
- (28) Pérez-Salinas, A.; Cervera-Lierta, A.; Gil-Fuster, E.; Latorre, J. I. Data re-uploading for a universal quantum classifier. *Quantum* **2020**, *4*, 226, Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.
- (29) Suzuki, Y. et al. Qulacs: a fast and versatile quantum circuit simulator for research purpose. *Quantum* **2021**, *5*, 559, arXiv:2011.13524 [quant-ph].

- (30) Javadi-Abhari, A.; Treinish, M.; Krsulich, K.; Wood, C. J.; Lishman, J.; Gacon, J.; Martiel, S.; Nation, P. D.; Bishop, L. S.; Cross, A. W.; Johnson, B. R.; Gambetta, J. M. Quantum computing with Qiskit. 2024; \_eprint: 2405.08810.
- (31) Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272.
- (32) Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
- (33) Rupp, M.; Tkatchenko, A.; Müller, K. R.; Lilienfeld, O. A. v. Fast and accurate modeling of molecular atomization energies with machine learning. *Physical Review Letters* **2012**, *108*.
- (34) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling* **2010**, *50*, 742–754, Publisher: American Chemical Society.
- (35) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL Keys for Use in Drug Discovery. *Journal of Chemical Information and Computer Sciences* **2002**, *42*, 1273–1280.
- (36) Adams, H.; Emerson, T.; Kirby, M.; Neville, R.; Peterson, C.; Shipman, P.; Chepushtanova, S.; Hanson, E.; Motta, F.; Ziegelmeier, L. Persistence Images: A Stable Vector Representation of Persistent Homology. *Journal of Machine Learning Research* **2017**, *18*, 1–35.
- (37) Townsend, J.; Micucci, C. P.; Hymel, J. H.; Maroulas, V.; Vogiatzis, K. D. Representation of molecular structures with persistent homology for machine learning applications in chemistry. *Nature Communications* **2020**, *11*, 3230, Publisher: Nature Research.
- (38) Schiff, Y.; Chenthamarakshan, V.; Hoffman, S. C.; Natesan Ramamurthy, K.; Das, P. Augmenting Molecular Deep Generative Models with Topological Data Analysis Representations. ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2022; pp 3783–3787.

- (39) Morgan, H. L. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *Journal of Chemical Documentation* **1965**, *5*, 107–113, Type: Journal Article.
- (40) Lundberg, S. M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems. 2017.
- (41) Parrish, R. M. et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197, Publisher: American Chemical Society.
- (42) Smith, D. G. A. et al. Psi4NumPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. *J. Chem. Theory Comput.* **2018**, *14*, 3504–3511, Publisher: American Chemical Society.