Prediction of Hydrogen Storage of Metal Organic Frameworks by Quantum Neural Network in Small Data Scenario

Christiano J. G. Pinheiro[1a]; Tharso D. Fernandes[2b]; Anderson A. M. Meneses[3]; Dennis R. Salahub[4] and Maicon P. Lourenço[5c];

Abstract— **Quantum neural network (QNN) regression models were created for metal organic frameworks (MOFs) hydrogen storage prediction, envisioning small data. This is a first step in the implementation of quantum active learning (QAL) algorithms – an iterative global optimization method – using QNN regressors.**

Keywords— **Quantum Neural Network, Metal Organic Frameworks, Hydrogen Storage.**

I. INTRODUCTION

Quantum computing has already met artificial intelligence (AI), resulting in the novel quantum AI (QAI) field[1]. In this work, quantum neural network models (QNN) were developed to predict metal organic framework (MOF)[2] hydrogen storage in a noise free quantum computing environment. The QNN models were designed using different quantum circuits (QC) layers. The sQUlearn[3] quantum computing library was used in this study. MOFs are important materials with high surface area and porous volume[2]. The aim of this work is to deal with small data for MOFs prediction[2], which is a common scenario in most of chemistry and materials science problems. Thus, it is a first step to study and understand QNNs in order to use them in quantum active learning (QAL), a field where quantum machine learning (QML) models are built iteratively in order to find new solutions[1], [4]. As perspective, this QNNs study aims at motivating the interface of QNNs regression models in the MLChem4D software[4]. MLChem4D implements several classical AL and quantum AL (QAL) methods, aiming at probing into the uncharted chemical space using QNN regressors with their epistemic uncertainty for new molecular discoveries by considering just few data in the sampling process. The full MOF data set used here can be found in Ref[2].

II. QUANTUM NEURAL NETWORK ENGINE

Quantum Neural Networks (QNNs) are variational models that optimize a cost function by computing expectation values[3]. Classical data $\mathbf{x}$ is embedded into a quantum state $\rho(\mathbf{x}, \boldsymbol{\theta})$ via an encoding circuit:

$$U(\mathbf{x}, \boldsymbol{\theta}) = \prod_{l=1}^{L} S_l(\mathbf{x}) W_l(\boldsymbol{\theta}_l),$$

where S_l are data-dependent operations and W_l parameterized gates. The output is the expectation value of an observable O:

$$f(\mathbf{x}, \boldsymbol{\theta}) = \mathrm{tr}\left[\rho(\mathbf{x}, \boldsymbol{\theta}) O\right].$$

The resulting state $\rho(x, \theta)$ is measured repeatably to calculate the expectation value of observable O.

Parameters $\boldsymbol{\theta}$ are optimized by minimizing the empirical loss $\mathcal{L}(\boldsymbol{\theta})$ using gradient-based methods. Gradients are computed via the *parameter-shift rule*:

$$\frac{\partial f}{\partial \theta} = \frac{1}{2}\left[\langle O \rangle^{+} - \langle O \rangle^{-}\right],$$

where $\langle O \rangle^{\pm}$ are expectation values evaluated at $\theta \pm \pi/2$. QNNs exhibit unique properties in expressivity and data efficiency.

III. RESULTS

The idea of this work is to create Quantum Neural Network (QNN) regression models[3] for MOF materials[2] (which are represented by the characteristics: surface area, pore volume, pressure, and temperature) and their hydrogen storage. From that, the QNN models designed using small data can predict the hydrogen storage of unlabeled MOFs. Different quantum circuit (QC) layers (or depth) as well as MOFs data size to create QNNs regression models were investigated.

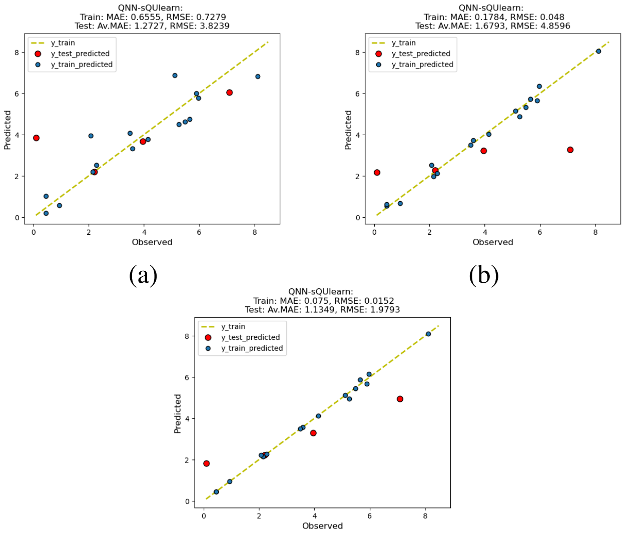
Fig. 1 shows the results of the QNN regression for 20 MOFs data. The 4:1 MOFs data ratio was considered to create the training and testing set. The red points in Fig. 1 are the testing set and the blue ones are the training set. Note in Fig. 1a (QC with 1 layer, Fig. 3a) the test$_{MAE}$/train$_{MAE}$ is equal to 1.9 while this ratio for the QNN model created using 4 QC layers (Fig. 3b, Fig. 3b) was 9.3. It indicates that for 20 MOFs data, increasing the depth of the QCs for building QNN regression models results in overfitting.

As we increase the data set to 60 MOFs, we can see in Fig. 2 an overall improvement of the QNNs models accuracy, when comparing, for instance, the test$_{MAE}$/train$_{MAE}$ ratio. As for 20 MOFs data, the QNNs for 60 data considered the 4:1 ratio to define the training (blue points in Fig. 2) and testing set (red points in Fig. 2). The QNNs regression models presented the following test$_{MAE}$/train$_{MAE}$ ratios: 1.6 for QC with 1

---

[1] Departamento de Engenharia Rural, UFES, [a] e-mail: christiano.pinheiro@ufes.br; [2] Departamento de Matemática Pura e Aplicada, UFES, [b] e-mail: tharso.fernandes@ufes.br; [3] Instituto de Engenharia e Geociências, UFOPA; [4] Department of Chemistry, University of Calgary. [5] Departamento de Química e Física, UFES, [c] e-mail: maiconpl01@gmail.com.
Este trabalho foi parcialmente financiado por FAPES, CNPQ, CAPES.

layer (Fig. 3a and Fig. 2a); 2.9 for QC with 4 layers (Fig. 3b and Fig. 2b); 9.5 for QC with 6 layers (Fig. 3c and Fig. 2c). Also for 60 MOFs data, increasing the depth of the QCs for building QNN regression models results in overfitting.



(a) (b)

(c)

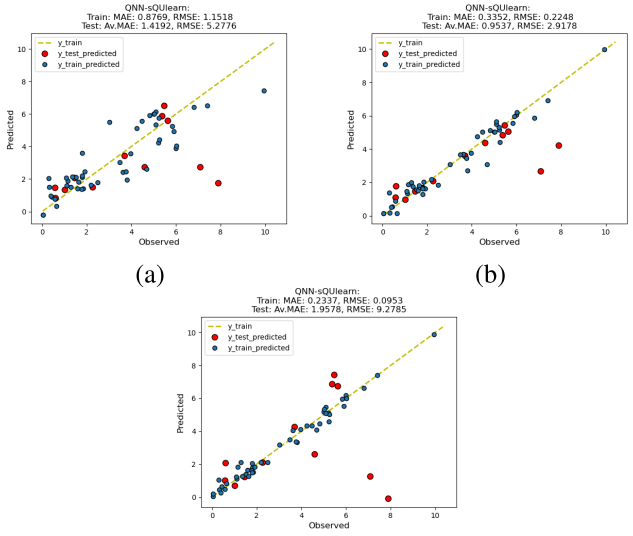
Fig. 1: QNN regressions of MOF hydrogen adsorption using 1 (a) 4 (b) and 6 (c) QC layers or depth. 20 MOFs data were used and the training (blue points) and testing (red points) set was split in 4:1. Observed axis: hydrogen adsorption in % obtained experimentally. Predicted axis: hydrogen adsorption in % estimated by QNN. MAE: mean absolute error. RMSE: root mean squared error. The diagonal line (green curve) has the x and y axis with the values equal to the observed axis, meaning an ideal line to be predicted by the QNN.

IV. CONCLUSION

This works shows that QNN, as implemented in sQUlearn, is feasible for performing predictions of MOFs hydrogen adsorption for small data, aiming at performing QAL for material design using QNN regressors, since in QAL the epistemic uncertainty is considered in the decision process. QNN models were created for 20 and 60 MOFs data. Also, the effect of different circuits layers in the regression accuracy was investigated. The results showed that increasing the number of layers (from 1 to 6) improved the training set but the test set, resulting in overfitting. Besides, the QNNs models with 1 QC layer (Fig. 3a) presented the better accuracy, as indicated by the test$_{MAE}$/train$_{MAE}$ ratio. As expected, as more data are used to train the QNN models, better is the regression accuracy.

As a perspective, new observables and quantum circuits for data encoding (as implemented in sQUlearn[3]) will be used in order to improve the QNN models.

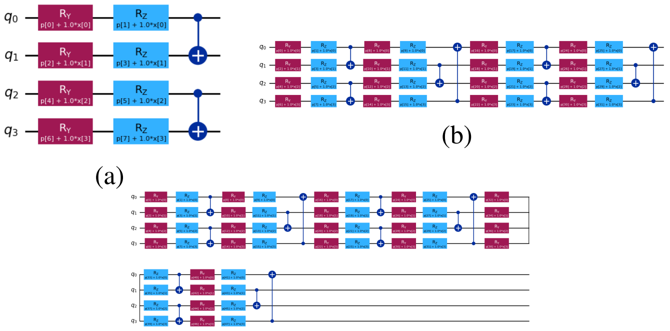
Finally, this QNN experience will allow the QNN implementation in MLChem4D code for QAL. The QNN models developed for MOFs (in small data scenario) can be applied to other materials, such as perovskites[4].



(a) (b)

(c)

Fig. 2: QNN regressions of MOF hydrogen adsorption using 1 (a) 4 (b) and 6 (c) QC layers or depth. 60 MOFs data were used and the training (blue points) and testing (red points) set was split in 4:1. Observed axis: hydrogen adsorption in % obtained experimentally. Predicted axis: hydrogen adsorption in % estimated by QNN. MAE: mean absolute error. RMSE: root mean squared error. The diagonal line (green curve) has the x and y axis with the values equal to the observed axis, meaning an ideal line to be predicted by the QNN.



(a) (b)

(c)

Fig. 3: Quantum circuit "YZ_CX_EncodingCircuit"[3] used to create the QNN models with 1 (a), 4 (b) and 6 (c) layers.

ACKNOWLEDGEMENT

The support of the Brazilian agencies: Fundação de Amparo à Pesquisa do Espírito Santo (FAPES), Conselho Nacional para o Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES).

REFERENCES

[1] Lourenço, M., Naseri, M., Herrera, L., Zadeh-Haghighi, H., Gaur, D., Simon, C. & Salahub, D. Quantum Active Learning for Structural Determination of Doped Nanoparticles-A Case Study of 4Al@ Si11. *Journal Of The Brazilian Chemical Society*. **36**, e-20250054 (2025)

[2] Salehi, K., Rahmani, M. & Atashrouz, S.. Machine learning assisted predictions for hydrogen storage in metal-organic frameworks. *International Journal of Hydrogen Energy*. **48**, (2023)

[3] Kreplin, D., Willmann, M., Schnabel, J., Rapp, F., Hagelüken, M. & Roth, M. sQUlearn: a Python library for quantum machine learning. *IEEE Software*. (2025)

[4] Lourenço, M., Tchagang, A., Shankar, K., Thangadurai, V. & Salahub, D. Active learning for optimum experimental design—insight into perovskite oxides. *Canadian Journal Of Chemistry*. **101**, 734-744 (2023)