QUANTUM ALGORITHM OPTIMIZATION ON MANIFOLD: RETHINKING QUANTUM COORDINATE BEYOND EUCLIDEAN

Nam Nguyen, M.A
Department of Statistics
University of South Florida
Tampa, FL 33620
namphuongnguyen510@gmail.com

July 16, 2023

ABSTRACT

This work introduces a new approach for quantum algorithm optimization performed directly on the manifold of quantum Hilbert space. The optimization problem is formulated and solved using group arithmetic, which provides a coherency between the initial quantum state, unitary transformation, and measured operator through the concept of finite Galois field. The numerical demonstrations include (1) synthesized datasets, (2) COVID-19 growth dynamics modeling, (3) tumor burden modeling, and (4) molecular energy prediction. Finally, we establish the connection of the proposed algorithm to re-normalization in Quantum Field Theory, Quantum Computing, and Knots.

Contents

1 Introduction

Quantum computing is a type of computation that leverages the principles of quantum mechanics to process information. It represents a fundamental departure from classical computing and holds the potential to revolutionize a variety of fields. Quantum computing has potential advantages over classical computing, including data processing acceleration and problem-solving (machine learning (ML) [30, 31], quantum chemistry [1, 2, 4], simulation and cryptography [8–10]).

Quantum ansatz optimization, also often referred to within the context of the variational quantum eigensolver (VQE) algorithm or quantum approximate optimization algorithm (QAOA), is an approach in quantum computing that uses a hybrid quantum-classical scheme to solve problems [16, 25, 27, 28, 33, 40]. The protocol follows:

1. Choosing Ansatz Candidate: The term "ansatz" is borrowed from physics and denotes an educated guess or an assumed form for a function optimized to fit a particular scenario. In quantum computing, a quantum ansatz is a guess for the system's quantum state. The ansatz is typically chosen based on some physical insight into the problem.
2. Preparing the Quantum State: Once a model architecture candidate is chosen, it is used to prepare a quantum state on the quantum computer. This is typically done by applying a sequence of quantum gates based on the parameters defined in the ansatz.
3. Measurement and Feedback: Measurements are taken once the quantum state has been prepared. These measurements are then fed back into a classical computer.
4. Classical Optimization: The classical computer takes the results of these measurements and uses them to adjust the parameters in the ansatz. This process is essentially an optimization loop, and classical optimization algorithms can be used to find the best set of parameters.

Quantum ansatz optimization is an important topic in the early literature of Quantum Computing and Quantum Machine Intelligence. Finding smaller complex circuits with high performance can significantly reduce the computation cost for evaluating quantum algorithms.

1.1 Related Works

There are two principal classifications of the ansatz optimization problem. The first category focuses on "circuit simplification", aimed at minimizing the computational burden on quantum hardware. This is achieved by optimizing the local or global structure of the ansatz and replacing it with equivalent but more computationally efficient architectures. This approach has been reported in several studies. The second category of ansatz optimization targets the identification of the best ansatz that delivers optimal performance for specific tasks. This heuristic search focuses on deriving a highly-performing ansatz for specific tasks rather than minimizing computational requirements. The primary objective of our proposed work falls under this category, aiming to identify the optimal ansatz for quantum machine learning problems.

In recent works, frontiers in quantum layout synthesis [22, 23, 36, 37] introduced the optimization framework based on the coordinate of quantum ansatz, for which the space coordinate represents the number of prepared qubits (width) and the time coordinate represents the circuit depth. Although these techniques achieved remarkable effectiveness and efficiency, they could still be completed. Specifically, the quantum coordinate in reference is Euclidean vector space, which could meet a high complexity when addressing the permutation of quantum gates. The quantum operator generally does not commute, meaning $AB \neq BA$. Thus, given a search space of (Space = 1, Time = n) with the sequential of gates $A_1 \cdot A_2 \cdot A_3 \dots A_n$, there will be $n!$ permutations; i.e., neural candidates from the symmetric group of n-order.

1.2 Contribution

In this study, we introduce a novel framework that enables quantum algorithm optimization directly on the manifold of the quantum Hilbert space. We introduce the formalism in Section 2 to establish several important preliminary for the model construction. Then, Section 3 will introduce the quantum algorithm alignment technique via group arithmetic. Our proposed model assumes the coherency between three quantum entities: (1) initial quantum states, (2) unitary transformation, and (3) measured operator via finite Galois field (Prime field). In other words, the configuration of unitary transformation and measured operator only differ from the configuration of initial state by shifting operators; however, these architecture embeddings with the proposed shift-operator form a closed, coherent mathematics structure of the group (Appendix A.1).

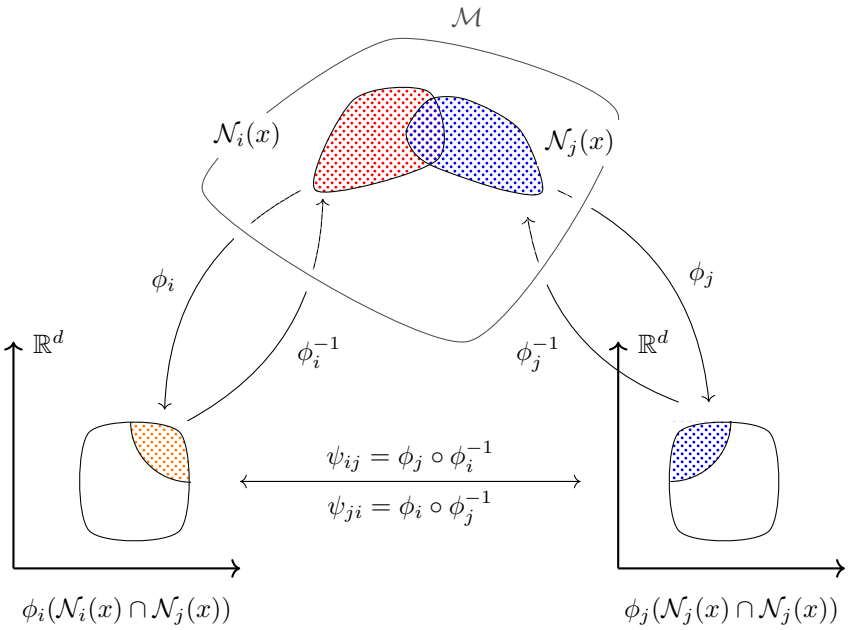


Figure 1: Illustration of Manifold $\mathcal{M}$, with $\phi_i(U_i)$ and $\phi_j(U_j)$ are the isomorphism mapping neighborhoods U_i and U_j on $\mathcal{M}$ onto Euclidean $\mathbb{R}^d$. The function ψ_{ij} is a transformation from $\phi_i(U_i) \mapsto \phi_j(U_j)$, written as the composition of $\psi_{ij} = \phi_j \circ \phi_i^{-1}$.

In the line of applications, we validate the proposed algorithm on four case-studies: (1) synthesized dataset (Section 4.1) and three actual applications, which are (2) modeling COVID-19 growth dynamics (Section 4.2.1), (3) modeling of TMB with data scarcity (Section 4.2.2) and (4) molecular energy prediction (Section 4.2.3).

Finally, we will discuss the theory of this research under a bigger scope in Section 5, for which we refer to a *LangLands* bridge connecting the re-normalization of quantum fields and the links from knots to quantum computing.

2 Formalism

Definition 1 Any Hilbert Space $\mathcal{H}$ with equipped metric measurement $d(x, y)$ - the norm induced by the inner product defines a metric on the Hilbert space, is a Hausdorff space. Given a Hausdorff space $\mathcal{M} \equiv \langle \mathcal{H}, d \rangle$, $\forall x \in \mathcal{M}$, $\exists \mathcal{N}(x)$ - the neighborhood of $\mathcal{H}$ that is homeomorphic to an open set in $\mathbb{R}^d$. Then $\mathcal{M}$ defines a d-dimensional, topological manifold.

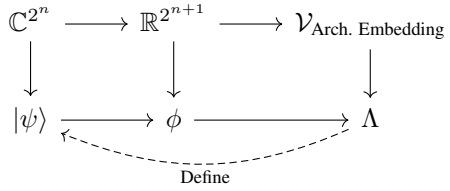
Example 1 Given the Hilbert space $\mathcal{H}^{2n}$ of the quantum wavefunction from n-qubit system. An 1-qubit system is associated with $\mathbb{C}^2$, for which the electronic wavefunction is

$$\begin{aligned} |\psi\rangle &= \alpha\,|0\rangle + \beta\,|1\rangle\,; \alpha, \beta \in \mathbb{C}. \\ \alpha &= a + bi \\ \beta &= c + di \end{aligned} \tag{1}$$

Because of the constrain $|\alpha|^2 + |\beta|^2 = 1$, we have $a^2 + b^2 + c^2 + d^2 = 1$, which defines a 4-sphere $\in \mathbb{R}^4$. Given a 2-qubit system, the quantum state is presented on a standard computational basis as

$$\begin{aligned} |\psi\rangle &= \alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle \\ \alpha, \beta, \gamma, \delta &\in \mathbb{C}^4 \end{aligned} \tag{2}$$

Thus, the associating Hausdorff space is $\mathbb{R}^8$. We can generalize as:



In this study, one of the main aims is to find the representation Λ associating with wavefunction through a manifold that enables **optimal** quantum ansatz design.

We find the connection between the wavefunction $|\psi\rangle$ on a finite field $\mathbb{F}_p$ (p is a prime) and unitary transformation as $\mathbb{F}_p \to GL_2(\mathbb{C})$ in Appendix 2. In this proposal, we aim to find the representation for quantum circuit architecture using a sequence of $\mathbb{F}_p$, given as

$$\mathcal{V}_{\text{Arch. Embedding}} = \mathbb{F}_2 \times \mathbb{F}_3 \times \mathbb{F}_5 \times \cdots \times \mathbb{F}_k := \mathcal{G}_p, \tag{3}$$

where k is the n^{th} prime and $\times$ is the Cartesian product.

Definition 2 Given the Hilbert space $\mathbb{C}^{2^n}$ of the wavefunction from n-qubit system, the vector space of architecture embedding can be presented as an n-sphere ($\mathcal{S}_n$):

$$(x - \boxed{2})^2 + (x - \boxed{3})^2 + \cdots + (x - \boxed{k})^2 = r^2, \tag{4}$$

where $\boxed{\eta}$ is the solution of equation $2^{n+1} \equiv \boxed{\eta} \mod p$; i.e., we compute the central of the n-sphere on the associated $\mathbb{F}_p$ or $\mathbb{Z}/p\mathbb{Z}$.

Definition 3 A quantum machine learning (QML) algorithm is generalized as the following information flow:

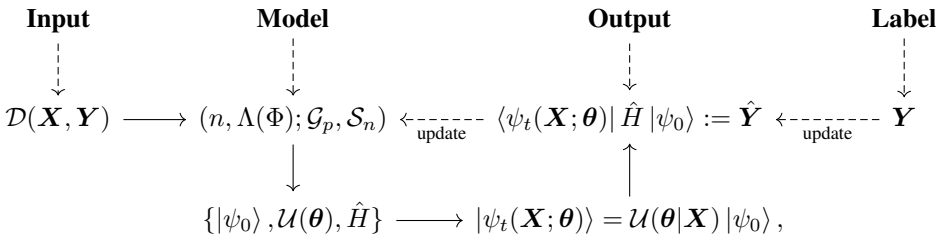


Figure 2: Information Flow of Quantum Machine Learning Models.

where n is the number of prepared qubits, $\Lambda(\Phi)$ is the architecture weight; $|\psi_0\rangle$ and $|\psi_t\rangle$ are initial and final state; $\mathcal{U}(\boldsymbol{\theta}|\boldsymbol{X}) |\psi_0\rangle$ is the quantum perturbation under unitary, parameterized transformation by $\boldsymbol{\theta}$ with realization of observations $\boldsymbol{X}$.

3 Approach

We begin with establishing a criterion for good QML models or optimality in such algorithm design:

- C1 **Optimization of Learning Objectives:** Given the predicted value $\hat{\boldsymbol{Y}}$ from a QML model, we aim to minimize the cost function $\mathcal{L}(\boldsymbol{Y}, \hat{\boldsymbol{Y}})$. This approach is common in ML; thus, further discussion will be neglected in this study.
- C2 **Optimization of Architecture Representation:** The outcome of a QML model is defined by the triplet $|\psi_0\rangle$ (initial state), $\mathcal{U}(\boldsymbol{\theta})$ (unitary transformation) and $\hat{H}$ is an operator for measurement. In common designs, there is no connection among these elements: the unitary transformation is defined by a stack of identical circuit ansatzes, while operators are arbitrary Hermitian matrices. We assume that good model architecture design should have a coherency among the quantum states, unitary transformation, and quantum operators. The optimization protocol is depicted in Diagram 3.

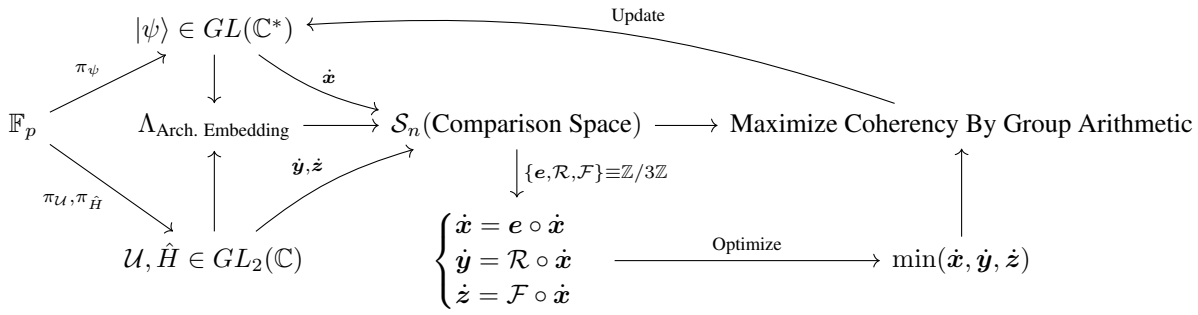


Figure 3: Optimization of Architecture Representation on Λ using discussed n-sphere $\mathcal{S}_n$ using embedded coordinate $\dot{\boldsymbol{x}}$, $\dot{\boldsymbol{y}} = \mathcal{R} \circ \dot{\boldsymbol{x}}$ and $\dot{\boldsymbol{z}} = \mathcal{F} \circ \dot{\boldsymbol{x}}$ where $\mathcal{R}$ and $\mathcal{F}$ are shifting operator to right and left, respective. The use of group operator $\{\boldsymbol{e}, \mathcal{R}, \mathcal{F}\} \equiv \mathbb{Z}/3\mathbb{Z}$ creates the coherency in the architecture representation.

3.1 Quantum Algorithm Alignment via Group Arithmetic

Here, we use the constructed n-sphere as the comparison space for architecture embeddings, including $\pi_\psi, \pi_\mathcal{U}$ and $\pi_{\hat{H}}$ corresponding to the under-scripted quantities. This approach is novel because we measure the similarity of such model architecture representations on the same manifold spanned by sequences of prime fields $\mathbb{F}_p$, which has yet to be used to extend our knowledge. We assume that $\pi_\psi, \pi_\mathcal{U}$ and $\pi_{\hat{H}}$ is elements of symmetric group S_n, represented by n-sphere.

Morevover, these architecture representations $\pi_\psi, \pi_\mathcal{U}$ and $\pi_{\hat{H}}$ only differ by shifting operator $\mathcal{R}$ (shifting right) or $\mathcal{F}$ (shifting left). Specifically, given the coordinates of embeddings on $\mathcal{S}_n$

$$\begin{aligned} \pi_\psi &= (x_1, x_2, \ldots, x_n) = \dot{\boldsymbol{x}} \in \mathcal{S}_n \\ \pi_\mathcal{U} &= (y_1, y_2, \ldots, y_n) = \dot{\boldsymbol{y}} \in \mathcal{S}_n \\ \pi_{\hat{H}} &= (z_1, z_2, \ldots, z_n) = \dot{\boldsymbol{z}} \in \mathcal{S}_n \\ \langle x_i, y_i, z_i \rangle &\in \mathbb{Z}/3\mathbb{Z} \equiv \mathbb{F}_3 \equiv \Delta_3 \end{aligned} \tag{5}$$

we have the collection $\{\pi_\psi, \pi_\mathcal{U}, \pi_{\hat{H}}\} \equiv \mathbb{Z}/3\mathbb{Z}$, which is isomorphic with the triangle group Δ_3. We define the

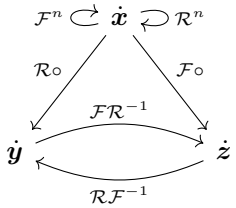


Figure 4: Group Transformation $\{\boldsymbol{e}, \mathcal{R}, \mathcal{F}\} \equiv \mathbb{Z}/3\mathbb{Z} \equiv \Delta_3$, with $\boldsymbol{e} = \mathcal{R}^n = \mathcal{F}^n$.

shift-right operator $\mathcal{R}$ acting $\dot{\boldsymbol{x}}$ gives $\dot{\boldsymbol{y}} = \mathcal{R} \circ \dot{\boldsymbol{x}}$ (similar to rotation but in n-dimension); shift-left operator $\mathcal{F}$ acting $\dot{\boldsymbol{x}}$ gives $\dot{\boldsymbol{z}} = \mathcal{F} \circ \dot{\boldsymbol{x}}$. We use group arithmetic to achieve three goals. First, the group transformation establishes the coherency or dependencies among architecture representations. Second, group structure ensures the closure of the mathematical structure (here is the search space) by the group axioms (Appendix A.1), which could lead to better convergence and optimality of the algorithm. Finally, establishing the dependency means reducing the search space complexity: we only need to optimize for $\dot{\boldsymbol{x}}$ and derive the corresponding $\dot{\boldsymbol{y}}$ and $\dot{\boldsymbol{z}}$ based on the established group transformation. It is easy to imagine that the search space will be triple if we do not consider such a coherency because we now need to optimize for all coordinate $\{\dot{\boldsymbol{x}}, \dot{\boldsymbol{y}}, \dot{\boldsymbol{z}}\}$.

Using group and representation theory is our lessons learned from a novel description of modern AI models, for which any AI algorithm can be described by the group of actions on set [11, 26]. The existing work provides a comprehensive discussion of classical AI architecture but has yet to be specified for quantum AI models. Here, we leverage the theories to learn (1) the representation π_ψ for prepared quantum states, (2) the representation $\pi_\mathcal{U}$ for unitary transformation, and (3) the representation $\pi_{\hat{H}}$ for the measured operator. We highlight that these representations are coherent and being optimized on $\mathcal{S}_n$, constructed by finite fields $\{\mathbb{F}_p\}_{i=1}^n$.

3.2 Algorithm

We describe the proposed algorithm in **Algorithm** 1. First, we prepare an n-qubit system with the set of model variational weight $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}$ and architecture weight $\dot{\boldsymbol{x}}$. Second, we define the comparison space for architecture weight optimization, which is

$$\mathcal{S}_n = \sum_{i=1}^{n} (\boldsymbol{\zeta} - \boxed{\eta})^2 \tag{6}$$

where $\boxed{\eta}$ is the solution of equation $2^{n+1} \equiv \boxed{\eta} \mod p$. We explain why this comparison space is used in Section 5.1. The **for-loop** of model training is discussed with details in **Algorithm** 1. To optimize the algorithm, we synchronously minimize the learning objective (C1) and architecture loss (C2) by

$$\begin{cases} (C1) : \mathcal{L}_{\text{obj}}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) \\ (C2) : \mathcal{L}_{\text{arch.}}(\mathcal{S}_n) = \frac{1}{3} \sum \left(|\mathcal{S}_n(\dot{\boldsymbol{x}}) - \mathcal{S}_n(\dot{\boldsymbol{y}})| + |\mathcal{S}_n(\dot{\boldsymbol{y}}) - \mathcal{S}_n(\dot{\boldsymbol{z}})| + |\mathcal{S}_n(\dot{\boldsymbol{z}}) - \mathcal{S}_n(\dot{\boldsymbol{x}})| \right), \end{cases} \tag{7}$$

thus the final loss function is

$$\arg \min_{\boldsymbol{\alpha}', \boldsymbol{\beta}', \dot{\boldsymbol{x}}'} \left(\mathcal{L}_{\text{obj}} + \mathcal{L}_{\text{arch.}} \right) \tag{8}$$

4 Numerical Demonstration

4.1 Synthesized Study: Learning Non-Linear Dynamics with Limited Data

Experimental Design 1 We validate the effectiveness of the proposed model using synthesized datasets generated by mathematical models in **Table** 1. We inject random noise with standard deviation from 0.1, 0.2 to 0.3. This experimental setting allows us to evaluate the model ability of noise quantification and highlight. We investigate eight mathematical models used for modeling non-linear dynamics of physical systems. Of note, such a class of models is well-known and widely used in biomedical applications [5, 14, 15, 18, 42]. The details for these generated functions are given in Appendix C.1.

We generate 1000 observations from each model, resulting in 24 evaluated datasets (eight models with three noise levels). The monitored time is $t \in [0, 11]$ (in days), including 12 observations. We aim to demonstrate the effectiveness of our proposed algorithm in the data scarcity scenario. For each experiment, we use the first 9 data points ($t \in [0, 8]$) for model training and test on the last three observations. This evaluation has applicable meanings in clinical scenarios where we are interested in predicting the dynamics of physical systems forward in time. The target variable $\boldsymbol{Y}$ is calibrated using min-max normalization as in common practice.

Result 1 We show the fitted dynamics in **Appendix Figure** 1, 2, 3 (4 qubits), 4, 5, 6 (8 qubits) and 7, 8, 9 (12 qubits). In general, the dynamics of the fitted curve following exponential growth were recorded in a major of experimental runs. However, our model can generate different dynamics, such as modeling the linear growth in Appendix Figure 1, where the model dynamics decrease from $t \in [0, 4]$ and increase during the remaining monitored time.

The proposed algorithms effectively model (with noise) increasing dynamics such as linear, exponential, Gompertz, and logistic growth. In contrast, the algorithm failed to approximate monotonous decreasing functions such as exposure-dependent and non-linear dynamics with drug effects. However, we can adjust the predicted outcomes by adopting some class of function, such as:

$$\begin{aligned} \hat{\boldsymbol{Y}} &\leftarrow \frac{1}{\hat{\boldsymbol{Y}}}, \text{or} \\ \hat{\boldsymbol{Y}} &\leftarrow \frac{1}{1 + \hat{\boldsymbol{Y}}} \end{aligned} \tag{9}$$

to generate decreasing dynamics.

The proposed model can effectively perform regression modeling in the low-limited dataset, with output dynamics similar to exponential growth but not restricted to such an exponential family.

Remarks 1 We draw some remarks as follows:

Algorithm 1 Quantum Algorithm Optimization on Manifold

input: $\mathcal{D} = (\boldsymbol{X}, \boldsymbol{Y})$
initiate: n-qubit; model weights $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}$; architecture weight $\dot{\boldsymbol{x}}$; number of epochs N.
define: $\mathcal{S}_n = \sum_{i=1}^{n} (\boldsymbol{\zeta} - \boxed{\eta})^2$ where $\boxed{\eta}$ is the solution of equation $2^{n+1} \equiv \boxed{\eta} \mod p$.
for i in N:

compute:

$|\psi_0\rangle = \otimes \prod_{j=1}^{n} (\alpha_i |0\rangle + \beta_i |1\rangle)$ (Initial Quantum State)

$\dot{\boldsymbol{y}} = \mathcal{R} \circ \dot{\boldsymbol{x}}$

$\dot{\boldsymbol{z}} = \mathcal{F} \circ \dot{\boldsymbol{x}}$

$\mathcal{S}_n(\dot{\boldsymbol{x}})|_{\boldsymbol{\zeta}=\dot{\boldsymbol{x}}}; \mathcal{S}_n(\dot{\boldsymbol{y}})|_{\boldsymbol{\zeta}=\dot{\boldsymbol{y}}}; \mathcal{S}_n(\dot{\boldsymbol{z}})|_{\boldsymbol{\zeta}=\dot{\boldsymbol{z}}}$ (Arch. Representations)

for p in $\{\mathbb{F}_p\}_{p=1}^{n}$:

$|\psi_p\rangle = \alpha_p |0\rangle + \beta_p |1\rangle$

$$\mathcal{U}_p(\boldsymbol{\zeta})(\in GL_2(\mathbb{C})) := \begin{bmatrix} \cos\left(\frac{2\pi(\zeta - \boxed{\eta})}{p}\right) & -\sin\left(\frac{2\pi(\zeta - \boxed{\eta})}{p}\right) \\ \sin\left(\frac{2\pi(\zeta - \boxed{\eta})}{p}\right) & \cos\left(\frac{2\pi(\zeta - \boxed{\eta})}{p}\right) \end{bmatrix} \Bigg|_{\boldsymbol{\zeta}=\dot{\boldsymbol{x}},\dot{\boldsymbol{y}},\dot{\boldsymbol{z}}} \quad \text{(Unitary Transformation)}$$

compute:

$|\psi_t\rangle = \otimes \prod_{p=1}^{n} |\psi_p\rangle$

$\mathcal{U}(\dot{\boldsymbol{y}}) := \otimes \prod_{p=1}^{n} \left(\mathcal{U}_p(\dot{\boldsymbol{y}}) |\psi_p\rangle \right)$

$\hat{H}(\dot{\boldsymbol{z}}) := \otimes \prod_{p=1}^{n} \left(\mathcal{U}_p(\dot{\boldsymbol{z}}) \right)$ (Measurement Operator)

forward pass: (Feature Dependent)

$H_0 = \boldsymbol{X} \cdot |\psi_0\rangle$

$H_t = \boldsymbol{X} \cdot |\psi_t\rangle$

$\hat{\boldsymbol{Y}} = \langle H_t | \hat{H} | H_0 \rangle$ (Quantum Perturbation)

compute objectives:

$$\begin{cases} \mathcal{L}_{\text{obj}}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) \\ \mathcal{L}_{\text{arch.}}(\mathcal{S}_n) = \frac{1}{3} \sum \left(|\mathcal{S}_n(\dot{\boldsymbol{x}}) - \mathcal{S}_n(\dot{\boldsymbol{y}})| + |\mathcal{S}_n(\dot{\boldsymbol{y}}) - \mathcal{S}_n(\dot{\boldsymbol{z}})| + |\mathcal{S}_n(\dot{\boldsymbol{z}}) - \mathcal{S}_n(\dot{\boldsymbol{x}})| \right) \end{cases}$$

optimize:

$$\arg \min_{\boldsymbol{\alpha'}, \boldsymbol{\beta'}, \dot{\boldsymbol{x}}'} \left(\mathcal{L}_{\text{obj}} + \mathcal{L}_{\text{arch.}} \right)$$

update: $(\boldsymbol{\alpha}, \boldsymbol{\beta}; \dot{\boldsymbol{x}}) \leftarrow \boldsymbol{\alpha'}, \boldsymbol{\beta'}, \dot{\boldsymbol{x}}'$;

1. Increasing the number of used qubits does not guarantee the improvement of the model prediction. In contrast, the model performance, like modeling data from linear growth dynamics, sometimes degrades.
2. Model trainability significantly depends on the training set, such as learning rate, weight decay, and the number of epochs. We suggest performing hyper-parameter optimization depending on the problem of interest to select the best model.

4.2 Applications

4.2.1 COVID-19 Growth Dynamics

Experimental Design 2 We validate our algorithm in modeling the growth of new cases in the COVID-19 pandemic, with data from two regions, Vietnam (VNM) and North America (NA) [34]. For the first nation, we use the monitored time in $[0, 70]$ from January to March 2022 to demonstrate the data scarcity case. We use data from January 2021 to April 2022 for the latter region to demonstrate large-scale experiments. We use the same train-test splitting as the

synthesized study: we use the first 47 days to train the model to learn the growth dynamics of COVID-19 in VNM and test on the remaining 23 days. Regarding NA, we use the data in the first year, 365 days for training and testing on the remaining period.

Result 2 The proposed algorithm effectively models the growth dynamics of COVID-19 cases in both quantified regions. In the VNM model, the fitted curve achieves 0.0367 in test MSE, while that in NA is 0.0027, shown in Figure 5.

Remarks 2 We find several interesting behaviors of the proposed algorithm in this case study:

1. In the VNM study, the predicted curve (red, dashed) does not pass through test observations but rather than predicting the upper bound for the number of new cases. Of note, this prediction is learned from retrospective data used for training, depicted in blue. Thus, the algorithm can effectively predict the maximum number of new cases during the pandemic.
2. In the NA study, we see a contrast observation to VNM, as the predicted curve (red, dashed) is relatively close to the actual value. Besides, the fitted dynamics follow exponential growth.

4.2.2 Tumor Burden Modeling

Experimental Design 3 We test the proposed model for applying tumor burden modeling (TMB), for which tumor burden refers to the total amount or extent of cancerous tissue within a person's body. It measures the overall size and distribution of tumors in an individual. The measurement of tumor burden is important for several reasons. It helps determine the stage of cancer, which is a crucial factor in selecting the most appropriate treatment options. Tumor burden can also indicate the progression or regression of the disease over time, helping healthcare professionals monitor the effectiveness of treatment interventions.

We use the data collected from [39], including mouse tumors that received four different treatments: (1) Control, (2) Drug, (3) Radiation, and (4) Drug and Radiation. Of note, this is the real case study associated with the synthesized study, for which the discussed mathematical models are the common tool to address TMB modeling. We have demonstrated that the proposed algorithm effectively approximates the generated dynamics and quantifies the noise from data collection. Thus, we deter the comparison to classical models in further works.

Result 3 We show four random tumors in Figure 6, which indicates that the proposed algorithm can be effectively used for TMB modeling. The best-fitted curve is recorded in the mouse sample receiving the drug, while the worst prediction is in the sample receiving both the drug and radiation. Besides, we highlight that exponential dynamics could be insufficient for this task: TMB could be reduced immediately after the treatment; thus, monotonous increasing dynamics fail to tackle this phenomenon. Our model prediction surface is beyond the exponential family, which resulted in a critical point (when the gradient changes from positive to negative) discussed in the previous experiment.

Remarks 3 We provide a novel tool beyond the classical standpoint on the dynamics of a physical system: Classical mechanism sees a particle as a point in space-time, and the particle's dynamical quantities are known based on the particle's position. In our model, we take the quantum standpoint, in which the tumor is modeled as quantum entities, in which dynamical quantities are derived from the wave function. This approach has yet to be seen in the clinical literature to extend our knowledge.

4.2.3 Molecular Energy Modeling based on Bond Length

Experimental Design 4 We validate the proposed model in predicting molecular energy using its bond length in this case study. We test the algorithm in four molecules: H_2, H_4, H_6, and BeH_2. We re-formula the predicted outcomes as

$$\hat{\boldsymbol{Y}} \leftarrow \frac{1}{\hat{\boldsymbol{Y}}} \tag{10}$$

since the molecular energy is a decreasing function as the bond length increases. We apply the shift operator of the feature set, including only bond length, to generate the input matrix:

$$\boldsymbol{X} = \begin{bmatrix} b_0 & b_1 & b_2 \\ b_1 & b_2 & b_3 \\ \dots & . & . \\ b_k & b_0 & b_1 \end{bmatrix} \tag{11}$$

where b_i is the bond length associated with energy $\boldsymbol{y}_i$.

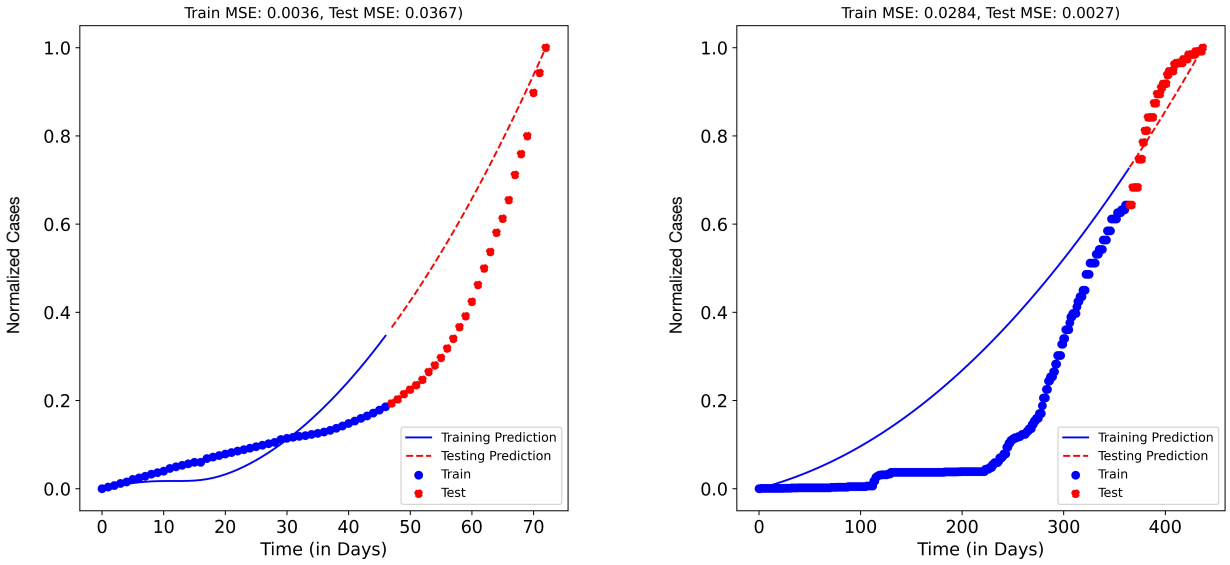


Figure 5: Experimental Result of the Proposed Model in COVID-19 Growth Modeling, concerning Two Regions: Vietnam and North America. The training observations are in blue; the testing data points are in red.

Result 4 Our model architecture can effectively predict molecular energy based on the bond length. The best mode with the lowest MSE is from modeling H_4 and H_2, which are also smaller molecules. In contrast, the algorithm has a larger error rate when modeling the energy of bigger molecules, including H_6 and BeH_2.

Remarks 4 We provide a robust evaluation of the model by investing in its capacity, reported in Figure 8. Here, we generate the $3D$ coordinate of a random molecule with 1000 observation and test on two methods for model outcome adjustment discussed in Equation 9. We observe that:

1. The first adjustment technique $\hat{\boldsymbol{Y}} \leftarrow \frac{1}{\hat{\boldsymbol{Y}}}$ provides a low-energy surface, in which high energy level only appears in a small region.

2. The second adjustment technique $\hat{\boldsymbol{Y}} \leftarrow \frac{1}{1+\hat{\boldsymbol{Y}}}$ provides a wider energy spectrum, which is almost uniformly distributed over the prediction surface.

3. Thus, our model can be used in modeling small molecules (small energy landscape) and larger molecules (wider energy spectrum) with a small adjustment in the model outcomes.

5 Theoretical Implications

5.1 Approaches to Open Problems in Circuit Placement

We will discuss several implications of our proposed algorithm to address existing challenges in circuit placement. In the existing approach, quantum circuit optimization is performed discretely, and the referred quantum coordinate is in Euclidean vector space. Specifically, we observe

1. The common search protocol can be depicted as

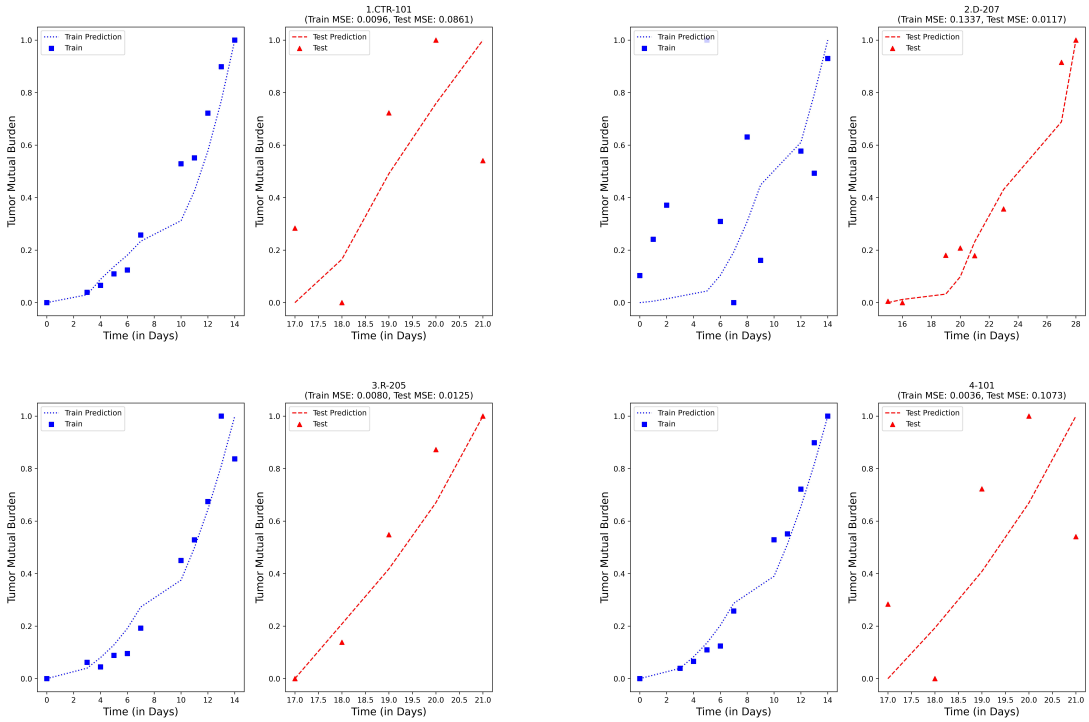
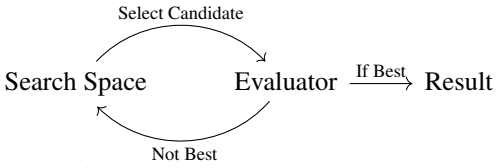
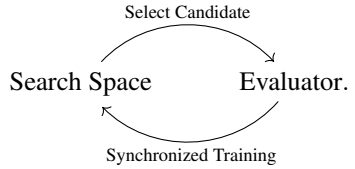


Figure 6: Experimental Result of the Proposed Model in Tumor Burden Modeling, concerning Four Mouse's Tumors Received Different Treatment.



. Our proposed algorithm has two distinct features compared to the existing protocols:

(a) First, the architecture parameter Λ is continuously trained with the model weight, making the algorithm a gradient-based approach. This is our lesson learned from matured Neural Architecture Search (NAS) literature of the classical AI (DART [24] and CSNAS [29]). However, adjusting from the classical to the quantum model is not trivial because quantum machine intelligence has a completely different mathematical abstraction than classical AI. To compare, we depict the optimization process as a continuous loop:



(b) Second, our proposed framework refers to quantum coordinates directly on manifold (Diagram 1), which could derive a better model candidate compared to the Euclidean coordinate of layout synthesis. In other words, we aim to optimize the algorithm at the abstract level, embedded in the representation of the Hilbert space manifold. At the same time, the existing works perform a coarse approximation of the same quantification through circuit placement. The translation of optimal algorithms from abstract to hardware (circuit layout) has yet to be discussed within the scope of this study.

2. **Why do we use the sphere $\mathcal{S}_n$ as referred comparison space for architecture optimization and loss computation?** A main challenge in NAS or ansatz optimization is to find an appropriate and robust metric space for architecture comparison. Recall that the architecture embedding Λ includes all components $\pi_\psi, \pi_{\mathcal{U}}$ and $\pi_{\hat{H}}$ formulated on a series of prime field $\mathbb{F}_p$. Besides, the central of the sphere $\mathcal{S}_n$ in Definition 2 is fixed for pre-defined n as is spanned by $n\#$ (the number of primes up to n). In other word, the boxed-notation $\boxed{\eta}$

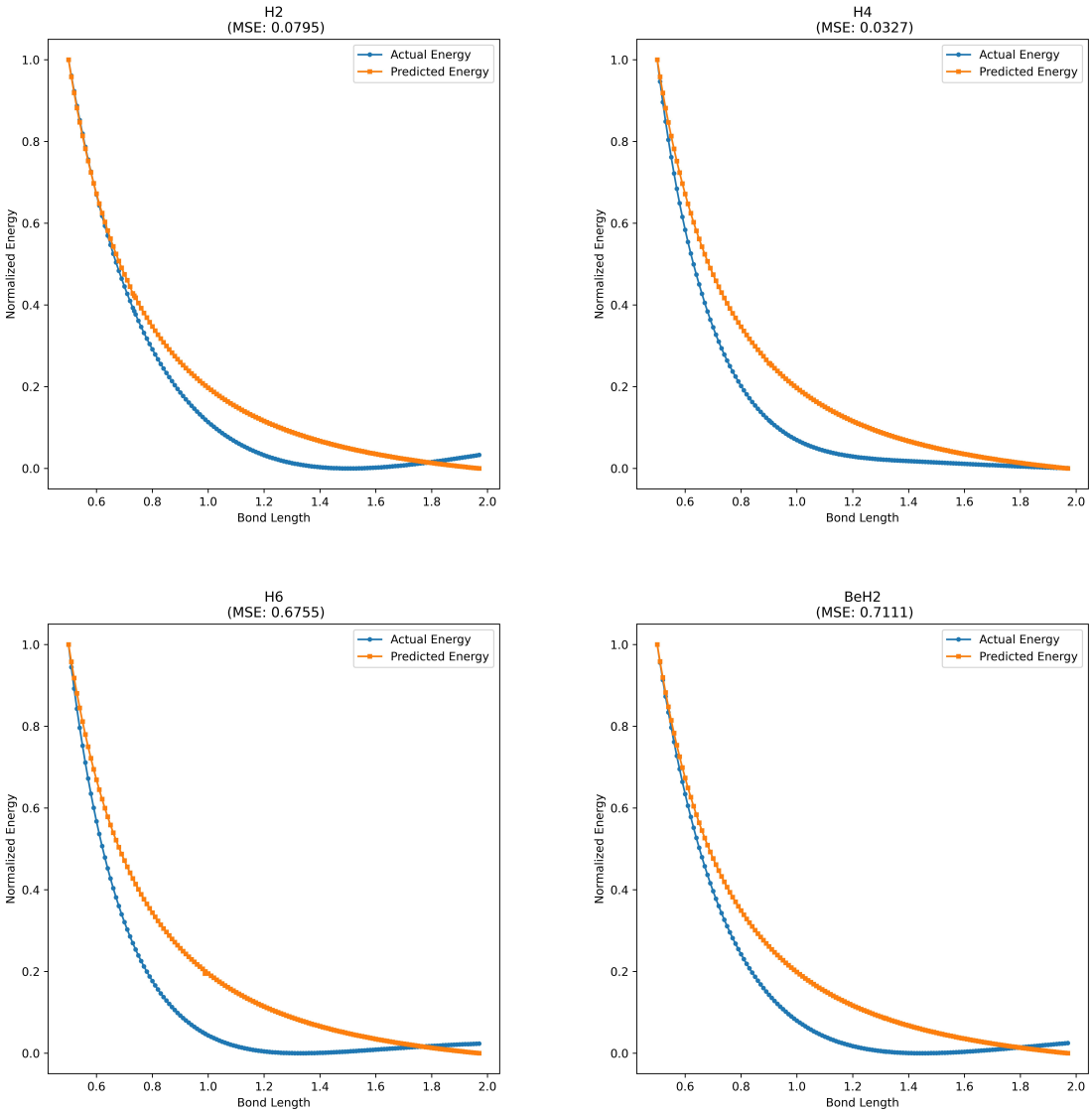


Figure 7: The Experimental Results of Molecular Energy Prediction based on Bond Length using our Proposed Algorithm. We test our model using four small molecules, including H_2, H_4, H_6, and BeH_2.

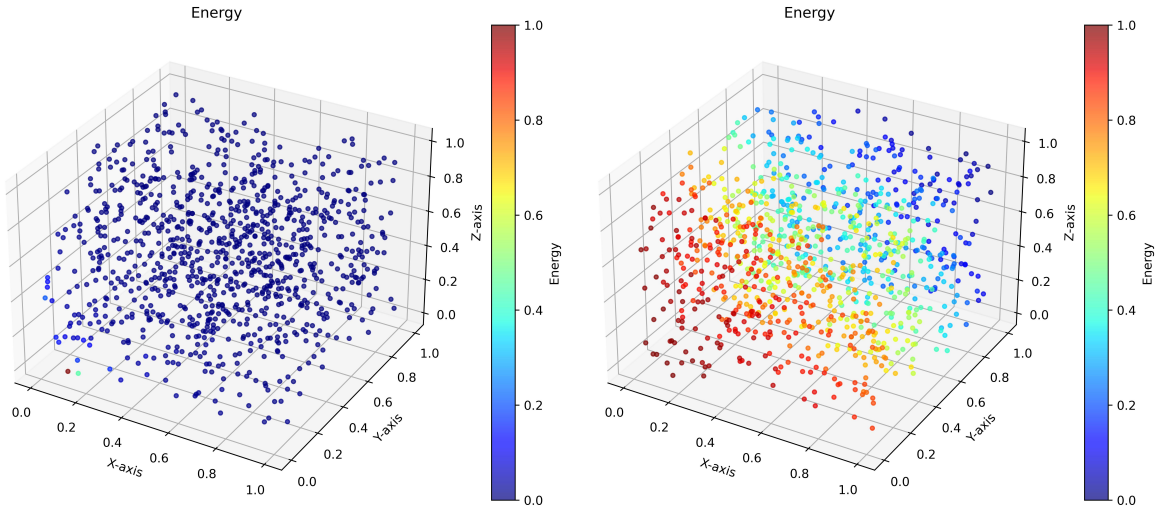


Figure 8: Model Capacity Analysis in the Context of Molecular Energy Prediction. We generate 1000 observations for predicted function: (Left) $\hat{\boldsymbol{Y}} \leftarrow \frac{1}{\hat{\boldsymbol{Y}}}$ and (right) $\hat{\boldsymbol{Y}} \leftarrow \frac{1}{1+\hat{\boldsymbol{Y}}}$.

is invariant, regarding $\pi_\psi, \pi_\mathcal{U}$ and $\pi_{\hat{H}}$. Thus, the proposed architecture loss:

$$\mathcal{L}_{\text{arch.}}(\mathcal{S}_n) = \frac{1}{3}\sum\left(|\mathcal{S}_n(\dot{\boldsymbol{x}}) - \mathcal{S}_n(\dot{\boldsymbol{y}})| + |\mathcal{S}_n(\dot{\boldsymbol{y}}) - \mathcal{S}_n(\dot{\boldsymbol{z}})| + |\mathcal{S}_n(\dot{\boldsymbol{z}}) - \mathcal{S}_n(\dot{\boldsymbol{x}})|\right) \tag{12}$$

encourage architecture embeddings $\pi_\psi, \pi_\mathcal{U}$ and $\pi_{\hat{H}}$ not being too far from each other (coherency) by minimizing their distance to the fixed centroid on the $\mathcal{S}_n$-sphere.

5.2 Approaches to Re-normalization of Quantum Field Theory

Re-normalization refers to various methods employed in quantum field theory (QFT) and the study of self-similar geometric structures. These techniques aim to address the occurrence of infinite values in calculated quantities by adjusting the values themselves, effectively compensating for the impacts of their self-interactions. Here, we propose an approach for re-normalization of the electronic wavefunction through minimization of architecture embeddings, including prepared quantum states π_ψ, unitary transformation $\pi_\mathcal{U}$ and measured operator $\pi_{\hat{H}}$.

These three quantities defining the QFT spanned by electronic wavefunction have self-similar geometric structures compared to the $\mathcal{S}_n$-sphere. Besides, the unitary transformation $\mathcal{U}(.)$ and measured operator $\hat{H}(.)$ are auto-induced by the initial state ψ since the geometric-parameters of these two former quantities $\dot{\boldsymbol{y}}$ and $\dot{\boldsymbol{z}}$ is computed by that of quantum state $\dot{\boldsymbol{x}}$. This is the lesson from our previous experiment: applying parameter-shared scheme and SoftMax function in quantum information aggregation could provide smoother prediction surface [32]. In this research, we investigate the re-normalization beyond model weight, additionally for architecture embedding weights Λ.

5.3 A Link from Quantum Computing to Knots

This work's underlying motivation is to connect concepts in Quantum computing and QFT to knot theory through the Galois field, which could enable a better understanding of Quantum Machine Intelligence. We show the connection as follows:

1. The connection between knot theory and quantum theory has emerged as a new tool to understand the topological structure of quantum entity [43], starting from Jones polynomial [19, 20], which can be computed by the Chern-Simons function [12] for gauge fields.
2. QC and the braid group are two interconnected areas of research that have gained significant attention in recent years. The braid group plays a crucial role in the study of quantum computing, particularly in the context of

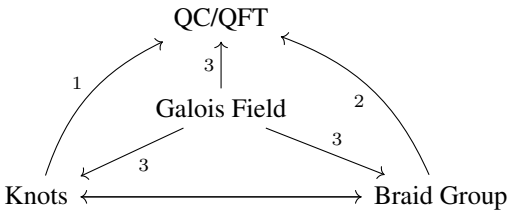


Figure 9: Connection of QC/QFT to Knots and Braid through Galois (Prime, Finite) Field.

topological quantum computation. We show the representation of logical quantum gates as braids in **Appendix Figure** 10.

3. The connection from Galois fields to knots arises using algebraic invariants, such as the Jones polynomial. Besides, the connection between Galois fields and braids can be established through the Artin braid group, a well-defined mathematical structure describing the possible braid operations on a given number of strands.

We illustrate the connections between these theories with this example:

Example 2 The Braid group on n strands, denoted B_n, is indeed generated by $n-1$ elements, usually denoted $\sigma_1, \sigma_2, ..., \sigma_{n-1}$. Each σ_i represents the act of crossing the ith strand over the $(i+1)$th strand. However, the operation of crossing the ith strand under the $(i+1)$th strand is usually denoted by σ_i^{-1}, not σ_{i-1}. The inverse operation σ_i^{-1} undoes the crossing of σ_i. In the case of B_3, the Braid group on three strands, there are two generators: σ_1 and σ_2 (it is isomorphic to the knot group of **the trefoil knot**). The relations are $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$, known as the braid relation. So, to clarify, σ_i^{-1} represents the act of crossing the ith strand under the $(i+1)$th strand, not σ_{i-1}. The index of the σ corresponds to the pair of strands being crossed, not the direction of the crossing.

Definition 4 The trefoil knot can be defined as the curve obtained from the following parametric equations:

$$\begin{aligned} x &= \sin t + 2\sin 2t, \\ y &= \cos t - 2\cos 2t, \\ z &= -\sin 3t. \end{aligned}$$

The $\langle 2, 3\rangle$-torus knot is also a trefoil knot. The following parametric equations give a $\langle 2, 3\rangle$-torus knot lying on the torus $(r-2)^2 + z^2 = 1$:

$$\begin{aligned} x &= (2 + \cos 3t)\cos 2t, \\ y &= (2 + \cos 3t)\sin 2t, \\ z &= \sin 3t. \end{aligned}$$

In algebraic geometry, the trefoil knot can be represented as the intersection of the unit 3-sphere S3 in C^2 with the zero sets of the complex polynomial $z^2 + w^3$, a cuspidal cubic. Galois theory can be applied to the field extensions generated by the solutions of polynomial equations to study their properties, including their symmetries. The field GF(p^n) is finite, while the solutions to the equation $z^2 + w^3 = 0$ in C^2 are infinite. Therefore, there needs to be a more straightforward connection between the Galois fields and the algebraic geometric representation of the trefoil knot. We find that both theories could be unified through the concept of geometric group theory [13] and universal algebra [7], which will be discussed in our future work.

6 Conclusion

To this end, we have introduced a new approach to perform quantum algorithm optimization from an abstract level of the manifold on the quantum Hilbert space (Section 3). The effectiveness and efficiency of our model are demonstrated by four case studies with (1) synthesized dataset (Section 4.1) and three real applications, including (2) modeling COVID-19 growth dynamics (Section 4.2.1), (3) modeling of TMB with data scarcity (Section 4.2.2) and (4) molecular energy prediction (Section 4.2.3).

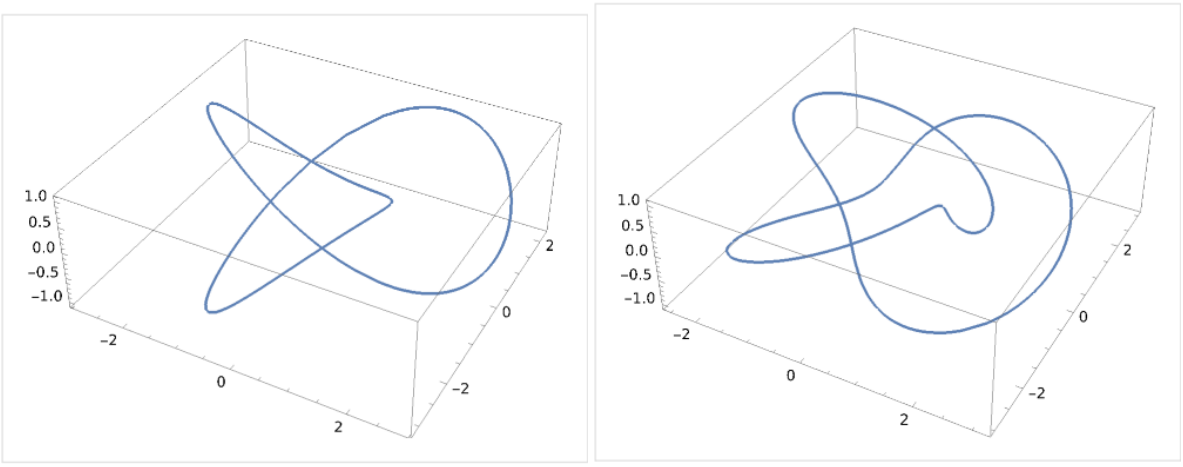


Figure 10: Trefoil Knot presented by Two Classes of Torus $\langle 2, 3 \rangle$.

We also discussed the theoretical implications for our study under a bigger picture, which aims to establish the *Langlands* bridge [41] between QFT, QC, and QMI (Section 5.2) to knots and braid group via Galois field (Section 5.3). Thus, we hope to rethink and propose a new approach for challenges in practice, such as layout synthesis and quantum circuit placement (Section 5.1).

Some limitations of this study including:

1. We have not tested the algorithm on actual quantum hardware such as dynamically field-programmable neutral atoms array processors [38], quantum optical [21] or discrete-variable quantum device [3, 6]. Thus, it is early to compare our proposed algorithm with frontier works in the field.
2. The output dynamics are not universal as we show in numerical demonstrations, for which fitted curves incline to have exponential-alike dynamics (monotonous increasing). Nevertheless, we have tackled this issue by small adjustments in the outcome's functional class, resulting in monotonous decreasing dynamics. Studying additional calibration for model outcomes is problem-dependent and worth studying in further research.
3. The model trainability could be worsened as a new set of architecture weight Λ is added to the model variational weights. However, this is an inevitable trade-off for gradient-based neural optimization: the model parameter is increasing on one side. On the other side, both model weight sets are synchronized during the training stage, which could enable better neural solutions.

Data Availability

The data used in this study is available upon request from the authors.

Code Availability

The code used in this study is available upon request from the authors.

Declarations

Author's Contribution: Nam Nguyen conceptualized the algorithm and performed numerical analysis.

Informed consent: Informed consent was obtained from all of the subjects involved in this study.

Competing interests: The author declares that they have no known competing financial interests reported in this paper.

References

[1] G.-L. R. Anselmetti, D. Wierichs, C. Gogolin, and R. M. Parrish. Local, expressive, quantum-number-preserving vqe ansätze for fermionic systems. *New Journal of Physics*, 23(11):113010, 2021.

[2] J. M. Arrazola, O. Di Matteo, N. Quesada, S. Jahangiri, A. Delgado, and N. Killoran. Universal quantum circuits for quantum chemistry. *Quantum*, 6:742, 2022.

[3] J. M. Arrazola, S. Jahangiri, A. Delgado, J. Ceroni, J. Izaac, A. Száva, U. Azad, R. A. Lang, Z. Niu, O. Di Matteo, et al. Differentiable quantum computational chemistry with pennylane. *arXiv preprint arXiv:2111.09967*, 2021.

[4] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, et al. Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wave-function expansions. *Physical Review A*, 98(2):022322, 2018.

[5] R. A. Bekker, S. Kim, S. Pilon-Thomas, and H. Enderling. Mathematical modeling of radiotherapy and its impact on tumor interactions with the immune system. *Neoplasia*, 28:100796, 2022.

[6] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. Akash-Narayanan, A. Asadi, et al. Pennylane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.

[7] C. Bergman. *Universal algebra: Fundamentals and selected topics*. CRC Press, 2011.

[8] J.-F. Biasse, X. Bonnetain, B. Pring, A. Schrottenloher, and W. Youmans. A trade-off between classical and quantum circuit size for an attack against csidh. *Journal of Mathematical Cryptology*, 15(1):4–17, 2020.

[9] J.-F. Biasse and B. Pring. A framework for reducing the overhead of the quantum oracle for use with grover's algorithm with applications to cryptanalysis of sike. *Journal of Mathematical Cryptology*, 15(1):143–156, 2020.

[10] J.-F. Biasse and F. Song. On the quantum attacks against schemes relying on the hardness of finding a short generator of an ideal in q (ζpn). Technical report, Citeseer, 2015.

[11] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[12] S.-S. Chern and J. Simons. Characteristic forms and geometric invariants. *Annals of Mathematics*, 99(1):48–69, 1974.

[13] C. Druţu and M. Kapovich. *Geometric group theory*, volume 63. American Mathematical Soc., 2018.

[14] H. Enderling and M. Zahid. Systems and methods for predicting individual patient response to radiotherapy using a dynamic carrying capacity model, Feb. 9 2023. US Patent App. 17/786,326.

[15] H. Enderling, M. Zahid, E. Moros, J. Caudell, A. Mohamed, and C. Fuller. Personalization of patient specific radiation dose and dose fractionation using volumetric tumor dynamics. In *MEDICAL PHYSICS*, volume 49, pages E676–E676. WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2022.

[16] T. Fösel, M. Y. Niu, F. Marquardt, and L. Li. Quantum circuit optimization with deep reinforcement learning. *arXiv preprint arXiv:2103.07585*, 2021.

[17] L. S. Georgiev. Topological quantum computation with non-abelian anyons in fractional quantum hall states. In *Quantum Systems in Physics, Chemistry, and Biology: Advances in Concepts and Applications*, pages 75–94. Springer, 2017.

[18] I. Harshe, H. Enderling, and R. Brady-Nicholls. Predicting patient-specific tumor dynamics: How many measurements are necessary? *Cancers*, 15(5):1368, 2023.

[19] V. F. Jones. A polynomial invariant for knots via von neumann algebras. 1985.

[20] V. F. Jones. Hecke algebra representations of braid groups and link polynomials. In *New Developments In The Theory Of Knots*, pages 20–73. World Scientific, 1987.

[21] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook. Strawberry fields: A software platform for photonic quantum computing. *Quantum*, 3:129, 2019.

[22] W.-H. Lin, J. Kimko, B. Tan, N. Bjørner, and J. Cong. Scalable optimal layout synthesis for nisq quantum processors. In *Design Automation Conference (DAC)*, 2023.

[23] W.-H. Lin, B. Tan, M. Y. Niu, J. Kimko, and J. Cong. Domain-specific quantum architecture optimization. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 12(3):624–637, 2022.

[24] H. Liu, K. Simonyan, and Y. Yang. Darts: Differentiable architecture search. *arXiv preprint arXiv:1806.09055*, 2018.

[25] R. Majumdar, D. Madan, D. Bhoumik, D. Vinayagamurthy, S. Raghunathan, and S. Sur-Kolay. Optimizing ansatz design in qaoa for max-cut. *arXiv preprint arXiv:2106.02812*, 2021.

[26] J. Masci, E. Rodolà, D. Boscaini, M. M. Bronstein, and H. Li. Geometric deep learning. In *SIGGRAPH ASIA 2016 Courses*, pages 1–50. 2016.

[27] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Physical Review A*, 98(3):032309, 2018.

[28] Y. Nam, N. J. Ross, Y. Su, A. M. Childs, and D. Maslov. Automated optimization of large quantum circuits with continuous parameters. *npj Quantum Information*, 4(1):1–12, 2018.

[29] N. Nguyen and J. M. Chang. Csnas: Contrastive self-supervised learning neural architecture search via sequential model-based optimization. *IEEE Transactions on Artificial Intelligence*, 3(4):609–624, 2021.

[30] N. Nguyen and K.-C. Chen. Bayesian quantum neural networks. *IEEE Access*, 10:54110–54122, 2022.

[31] N. Nguyen and K.-C. Chen. Quantum embedding search for quantum machine learning. *IEEE Access*, 10:41444–41456, 2022.

[32] N. Nguyen and K.-C. Chen. Translational quantum machine intelligence for modeling tumor dynamics in oncology. *arXiv preprint arXiv:2202.10919*, 2022.

[33] M. Ostaszewski, E. Grant, and M. Benedetti. Structure optimization for parameterized quantum circuits. *Quantum*, 5:391, 2021.

[34] Our World in Data. Coronavirus (COVID-19) Data. `https://ourworldindata.org/coronavirus`, Year Accessed. Accessed: Date Accessed.

[35] B. Steinberg. *Representation theory of finite groups: an introductory approach*. Springer, 2012.

[36] B. Tan, D. Bluvstein, M. D. Lukin, and J. Cong. Qubit mapping for reconfigurable atom arrays. In *Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design*, pages 1–9, 2022.

[37] B. Tan and J. Cong. Layout synthesis for near-term quantum computing: Gap analysis and optimal solution. In *Design Automation of Quantum Computers*, pages 25–40. Springer, 2022.

[38] D. B. Tan, D. Bluvstein, M. D. Lukin, and J. Cong. Compiling quantum circuits for dynamically field-programmable neutral atoms array processors. *arXiv preprint arXiv:2306.03487*, 2023.

[39] M. Varna, P. Bertheau, L. G. Legrès, et al. Tumor microenvironment in human tumor xenografted mouse models. *Journal of Analytical Oncology*, 3(3):159–166, 2014.

[40] G. Verdon, M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni. Learning to learn with quantum neural networks via classical neural networks. *arXiv preprint arXiv:1907.05415*, 2019.

[41] Wikipedia contributors. Langlands program. `https://en.wikipedia.org/wiki/Langlands_program`. [Online; accessed Date Accessed].

[42] N. Wilson, C. S. Drapaca, H. Enderling, J. J. Caudell, and K. P. Wilkie. Modelling radiation cancer treatment with a death-rate term in ordinary and fractional differential equations. *Bulletin of Mathematical Biology*, 85(6):47, 2023.

[43] E. Witten. Knots and quantum theory, 2011.

A Preliminary

A.1 Representation Theory

Group can be described as a **set** and an **operator** acting or (operating) on the set: Let a (finite) set of elements $G = \{g_0, g_1, g_2, \ldots, g_k\} = \{g_i\}_{i=1}^{k}$ and an operator $\circ$ that induces the interaction between G's elements; i.e., $g_i \circ g_j$ means g_i interacts with g_j. A group is denoted as a duality $(G, \circ)$ (mathematical structure) that satisfies these following properties:

1. **C**losure: $\forall g_i, g_j \in G;$ s.t. $g_i \circ g_j \in G$.
2. **A**ssociative: $\forall g_i, g_j, g_k \in G;$ s.t. $g_i \circ (g_j \circ g_k) = (g_i \circ g_j) \circ g_k$.
3. **N**eutral Existence: $\exists e \in G;$ s.t. $a \circ e = e \circ a = a$.
4. **I**nverse Existence: $\exists a^{-1} \in G;$ s.t. $a^{-1} \circ a = a \circ a^{-1} = e$.

We can explain basic concepts in quantum computing using geometric group theory.

1. Given the computational basis of qubits is $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, the electronic-wavefunction is $\psi = \alpha |0\rangle + \beta |1\rangle ; \alpha, \beta \in \mathbb{C}^*$. We have a mapping between a cyclic group $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ into $\mathbb{C}^*$, given as Euler

function

$$\begin{aligned}\psi :&\mathbb{F}_p \to \mathbb{C}^* \\ &m \to \exp\left(\frac{2\pi i m}{p}\right) = \cos\left(\frac{2\pi m}{p}\right) + i\sin\left(\frac{2\pi m}{p}\right) = \xi_p; \\ &m \in \{0, 1, 2, \ldots, p-1\}\end{aligned} \tag{13}$$

A single qubit quantum state can be represented as:

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle$$

where $|\psi\rangle$ is the quantum state, $|0\rangle$ and $|1\rangle$ are the basis states, θ and ϕ are real numbers representing the polar and azimuthal angles on the Bloch sphere, and i is the imaginary unit. The term $e^{i\phi}$ is a root of unity because for any integer n, $(e^{i\phi})^n = e^{in\phi} = 1$ if $n\phi = 2\pi m$ for some integer m. Thus, we can re-write

$$\begin{cases} |\psi\rangle = \xi_p\,|0\rangle + \xi_q\,|1\rangle \\ |\xi_p|^2 + |\xi_q|^2 = 1 \\ (\xi_p, \xi_q) \in \mathbb{F}_p \times \mathbb{F}_q; \end{cases} \tag{14}$$

meaning a quantum state $|\psi\rangle$ is defined by the coordinate (p, q) of two primes via finite fields $\mathbb{F}_p$ (Galois Field).

2. Each rotation $\boldsymbol{\theta}$ can be encoded by parameterized quantum gates corresponding to Ox, Oy, Oz, given by $R_{\sigma_x}(\boldsymbol{\theta}) = \begin{bmatrix} \cos(\boldsymbol{\theta}/2) & -i\sin(\boldsymbol{\theta}/2) \\ -i\sin(\boldsymbol{\theta}/2) & \cos(\boldsymbol{\theta}/2) \end{bmatrix}$, $R_{\sigma_y}(\boldsymbol{\theta}) = \begin{bmatrix} \cos(\boldsymbol{\theta}/2) & -\sin(\boldsymbol{\theta}/2) \\ \sin(\boldsymbol{\theta}/2) & \cos(\boldsymbol{\theta}/2) \end{bmatrix}$, and $R_{\sigma_z}(\boldsymbol{\theta}) = \begin{bmatrix} e^{-i\frac{\boldsymbol{\theta}}{2}} & 0 \\ 0 & e^{i\frac{\boldsymbol{\theta}}{2}} \end{bmatrix}$. We have the equivalent class of representations $\mathbb{F}_p \to GL_2(\mathbb{C})$ [35]

$$\begin{aligned}\mathcal{U}_y(\hat{\boldsymbol{\theta}}) &= \begin{bmatrix} \cos\left(\frac{2\pi m}{p}\right) & -\sin\left(\frac{2\pi m}{p}\right) \\ \sin\left(\frac{2\pi m}{p}\right) & \cos\left(\frac{2\pi m}{p}\right) \end{bmatrix} \\ &\equiv \mathcal{U}_z(-\hat{\boldsymbol{\theta}}) \begin{bmatrix} \exp(\frac{2\pi i m}{p}) & 0 \\ 0 & \exp(-\frac{2\pi i m}{p}) \end{bmatrix}.\end{aligned} \tag{15}$$

Here, $\mathcal{U}_y(\hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}} = \frac{2\pi m}{p}$ is the discrete approximation of parameterized quantum gates $R_{\sigma_y}(\boldsymbol{\theta}); \boldsymbol{\theta} \in \mathbb{R}$. It is because we only have finite value of m from $\{1, \ldots, p-1\}$ for $\hat{\boldsymbol{\theta}}$ instead of infinite number of rotations $\boldsymbol{\theta} \in \mathbb{R}$.

3. A time-evolution of the wavefunction from n-qubit $|0\rangle^n$ is given as

$$\psi(\hat{\boldsymbol{\theta}}) = \mathcal{U}_\sigma(\hat{\boldsymbol{\theta}}) \circ |0\rangle^n. \tag{16}$$

B Pseudo-Code

```
class Ansatz(nn.Module):
    def __init__(self, n_states):
        super(Ansatz, self).__init__()
        self.n_states = n_states
        self.A = nn.Parameter(torch.rand(n_states)+1j*torch.rand(n_states),
    requires_grad=True)
        self.B = nn.Parameter(torch.rand(n_states)+1j*torch.rand(n_states),
    requires_grad=True)
        self.Xdot = nn.Parameter(torch.rand(n_states), requires_grad=True)
        self.Ydot = self.R(self.Xdot)
        self.Zdot = self.F(self.Xdot)
        self.xi = torch.tensor([self.compute_p_root(self.n_states, p) for p in range
    (1,self.n_states+1)])
        # GEOMETRIC TENSORS
        primes = [sympy.prime(p) for p in range(1,self.n_states+1)]
```

```
        self.arch_norm = self.compute_Sn()
        self.GL_Xdot = [self.GL2(self.Xdot, self.xi, p, i) for i, p in enumerate(
    primes)]
        self.GL_Ydot = [self.GL2(self.Ydot, self.xi, p, i) for i, p in enumerate(
    primes)]
        self.GL_Zdot = [self.GL2(self.Zdot, self.xi, p, i) for i, p in enumerate(
    primes)]

    def forward(self,x):
        h = [self.compute_reps(x[i,:]).unsqueeze(0) for i in range(x.size(0))]
        h = torch.cat(h, dim = 0)
        h = h.unsqueeze(1)
        h_abs = h.abs()
        return h_abs

    def arch_loss(self):
        l0 = (self.arch_norm[0] - self.arch_norm[1]).abs()
        l1 = (self.arch_norm[0] - self.arch_norm[2]).abs()
        l2 = (self.arch_norm[1] - self.arch_norm[2]).abs()
        aloss = (l0 + l1 + l2).mean()
        return aloss

    def compute_reps(self, x):
        psi_0 = self.init_state()
        psi_t = self.unitary_trans()
        H = self.operator().type(torch.complex128)
        H0 = torch.mul(x.type(torch.complex128),
                       psi_0.squeeze(1).type(torch.complex128))
        H1 = torch.mul(x.type(torch.complex128),
                       psi_t.squeeze(1).type(torch.complex128))
        z = torch.matmul(torch.matmul(H1, H), H0)
        return z

    # OPERATOR MEASUREMENT
    def operator(self):
        h = self.GL_Zdot[0]
        for i in range(1, self.n_states):
            h = torch.kron(h,self.GL_Zdot[i])
        return h

    # UNITARY TRANSFORMATION
    def unitary_trans(self):
        f = [self.psi(self.A[i], self.B[i]) for i, _ in enumerate(range(self.n_states)
    )]
        H = []
        for i in range(self.n_states):
            U = torch.matmul(self.GL_Ydot[i].type(torch.complex128),
                             f[i].type(torch.complex128))
            H.append(U)
        h = H[0]
        for i in range(1, self.n_states):
            h = torch.kron(h,H[i])
        return h

    # COMPUTING ARCHITECTURE DISTANCE ON S_n[x-[p]]-sphere
    def compute_Sn(self):
        x = (self.xi - self.Xdot)**2
        y = (self.xi - self.Ydot)**2
        z = (self.xi - self.Zdot)**2
        return torch.sqrt(x), torch.sqrt(y), torch.sqrt(z)

    # CONSTRUCTING GL2 (Ry)
    def GL2(self, z, xi, p, i):
        a00 = torch.cos(2*torch.pi*(z[i]-xi[i])/p)
        a01 = -torch.sin(2*torch.pi*(z[i]-xi[i])/p)
```

```
        X = torch.tensor([a00,a01,-a01,a00])
        X = X.reshape(2,2)
        return X

    def R(self,x):
        y = torch.roll(x, shifts=1, dims = 0)
        return y

    def F(self,x):
        y = torch.roll(x, shifts=-1, dims = 0)
        return y

    def init_state(self):
        f = [self.psi(self.A[i], self.B[i]) for i,_ in enumerate(range(self.n_states))
    ]
        h = f[0]
        for i in range(1, self.n_states):
            h = torch.kron(h,f[i])
        return h

    def psi(self, a, b):
        x = a*torch.tensor([1,0]) + b*torch.tensor([0,1])
        x = x.reshape(2,1)
        return x

    def compute_p_root(self, n, prime_index):
        prime = sympy.prime(prime_index)  # Find the nth prime number
        power = 2**(n+1)  # Compute 2^(n+1)
        remainder = power % prime  # Take the remainder (2^(n+1) mod p)
        return remainder
```

Listing 1: The Pseudo-code for Algorithm 1

C Supplemental Results

C.1 Synthesized Study

Table 1: **Growth Models and Equations Used in Synthesized Study**. The exposure of tumors to the drug is assumed to follow: Exposure $=$ EpMax$(t^{1/2}/(\text{Ept50}^{1/2}+t^{1/2}))$. We used EpMax $= 50$, Ept50 $= 10$, decay $= 1$ and IC50 $= 5$ for data synthesis.

Model	**Equation**
Linear Growth	$dV/dt = \alpha$
Exponential Growth	$dV/dt = \alpha V$
Logistic Growth	$dV/dt = \alpha V(1 - V/V_{\text{Max}})$
Gompertz Growth	$dV/dt = \alpha V \log(V_{\text{Max}}/V)$
Exposure-Dependent FOTE	$dV/dt = \alpha V - \beta \cdot \text{Exposure} * V$
Exposure-Dependent FOTE Resistance	$dV/dt = \alpha V - \beta \exp(-\text{decay} \cdot t) \cdot \text{Exposure} * V$
First Order Treatment Effect	$dV/dt = \alpha V - \beta V$
Non-Linear Drug Exposure Effect	$dV/dt = \alpha V(1 - \text{EpMax}) \cdot \text{Exposure}/(\text{IC50} + \text{Exposure})$

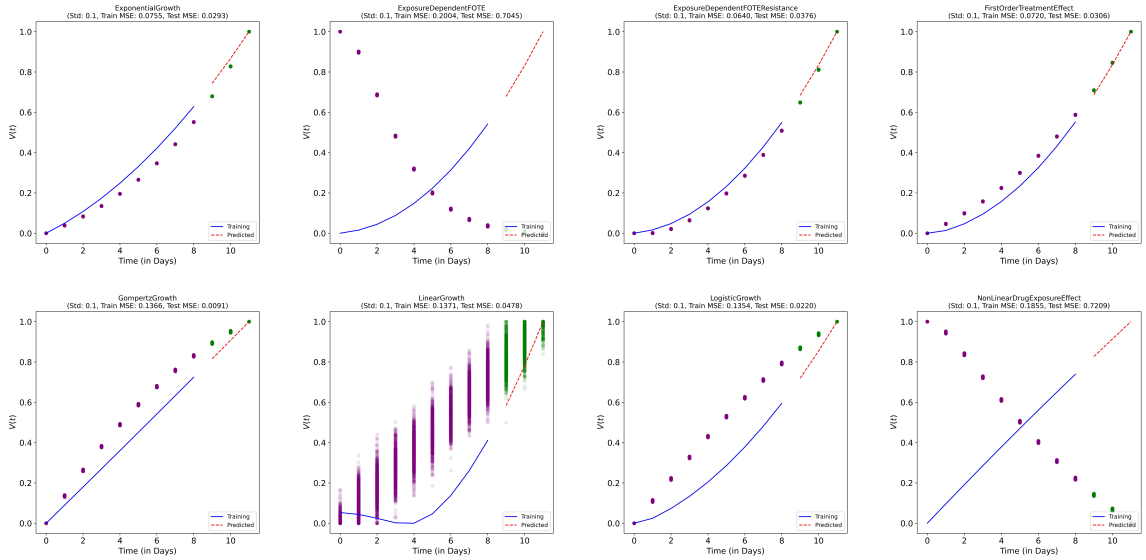


Figure 1: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.1$ using 4 qubits.

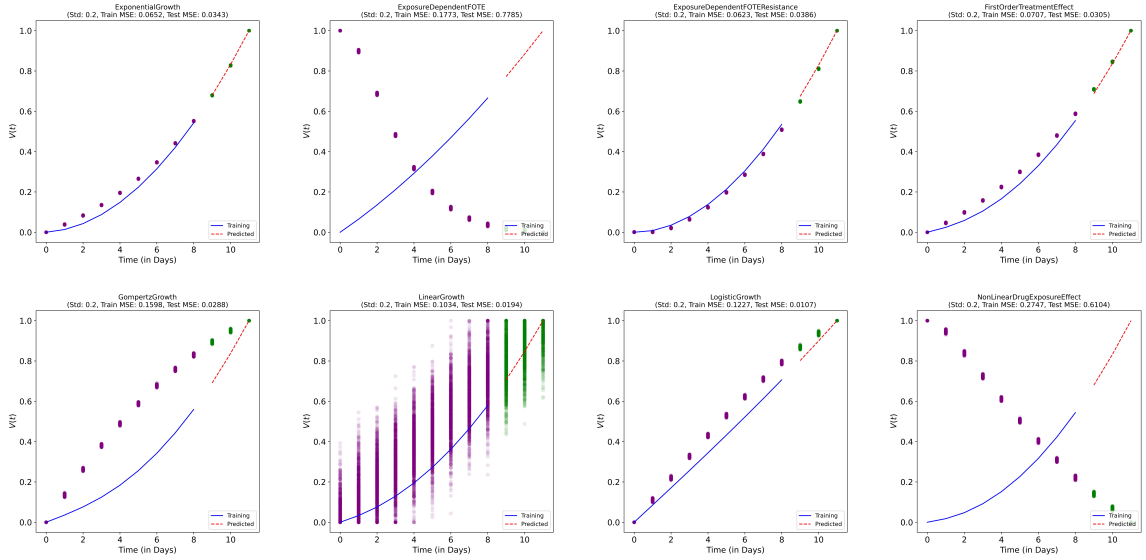


Figure 2: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.2$ using 4 qubits.

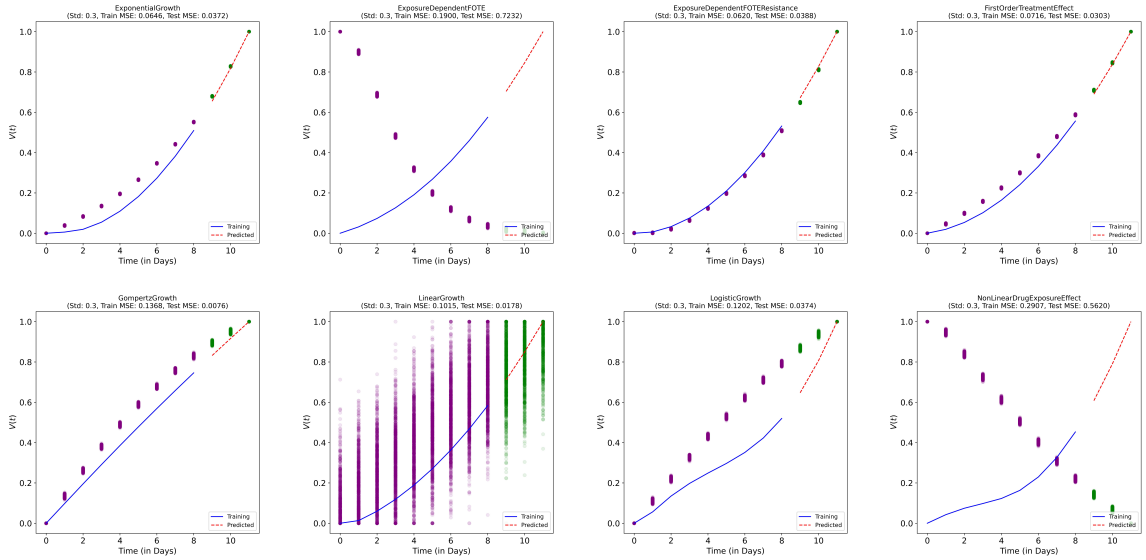


Figure 3: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.3$ using 4 qubits.

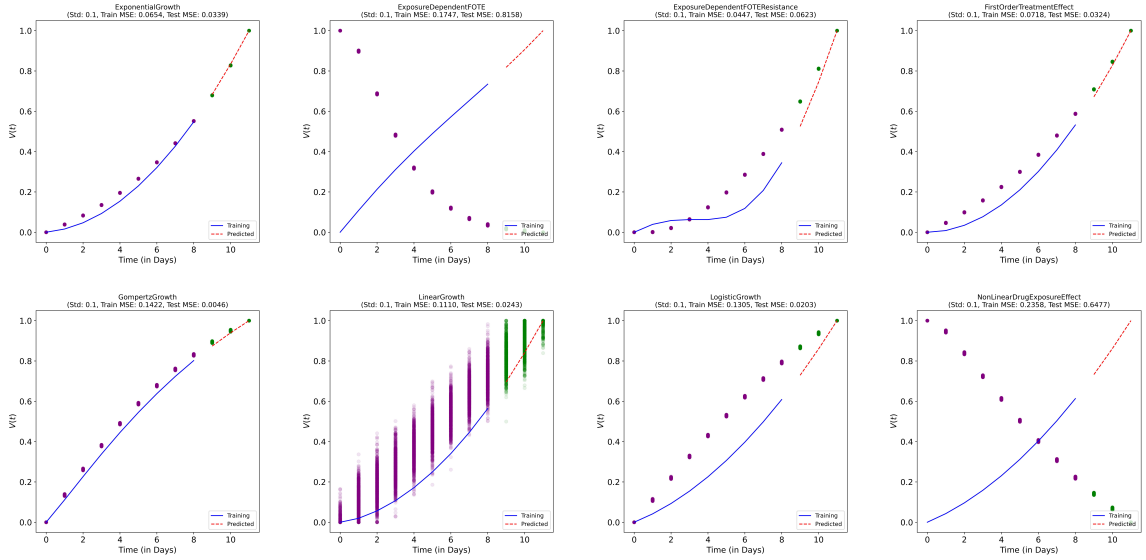


Figure 4: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.1$ using 8 qubits.

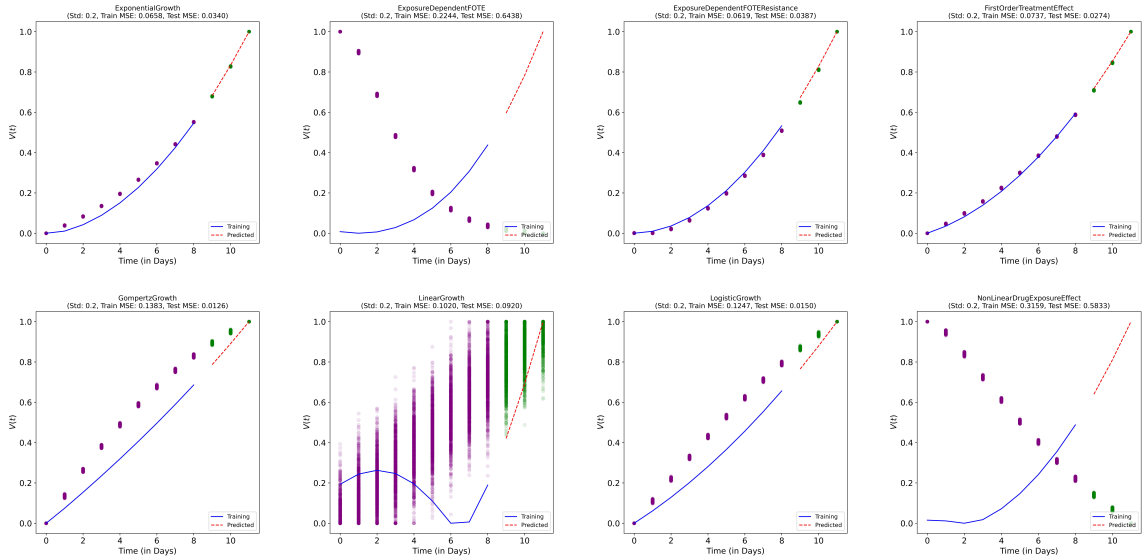


Figure 5: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.2$ using 8 qubits.

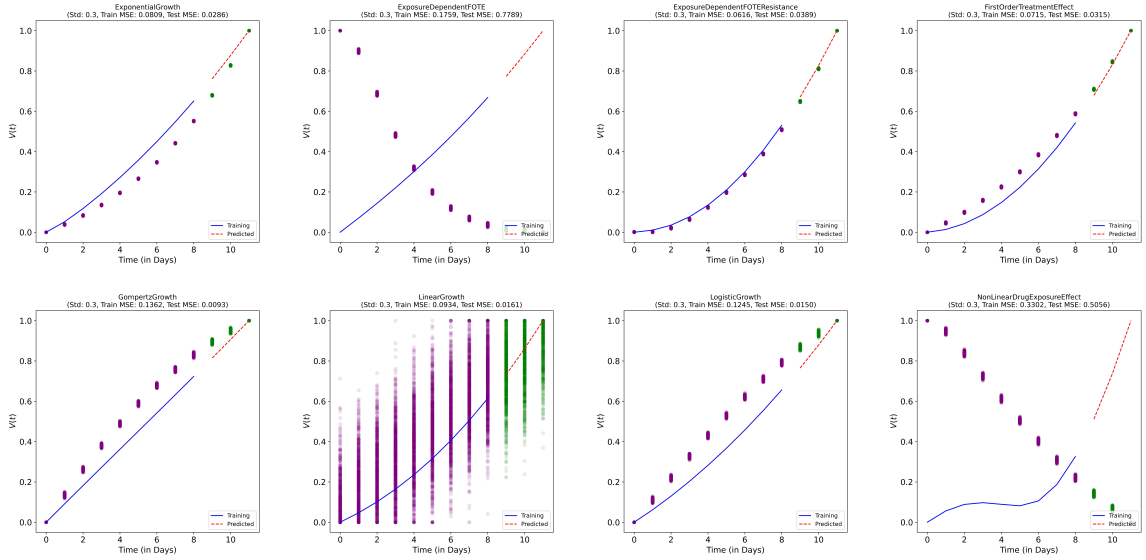


Figure 6: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.3$ using 8 qubits.

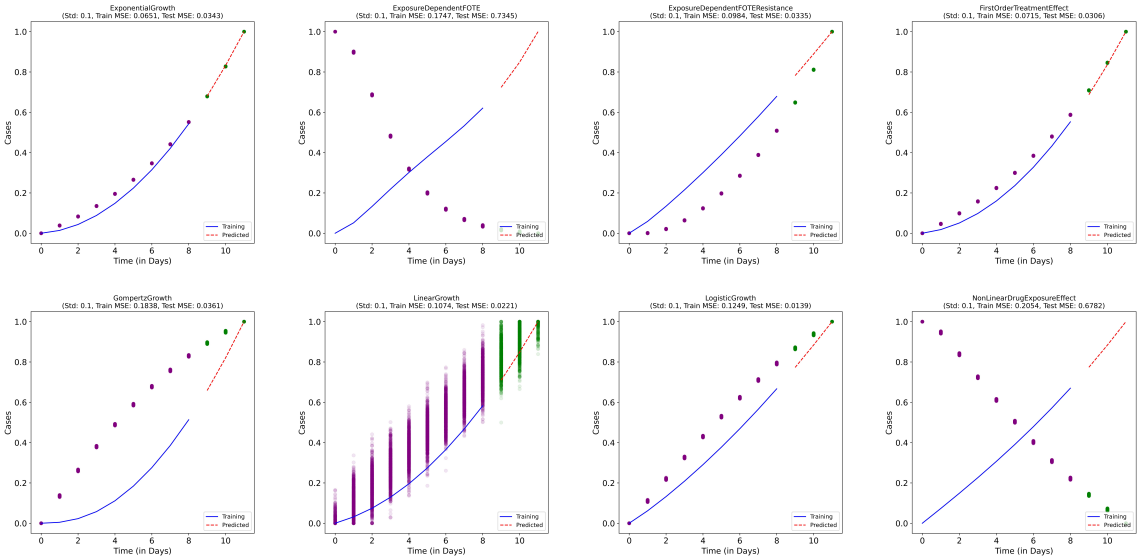


Figure 7: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.1$ using 12 qubits.

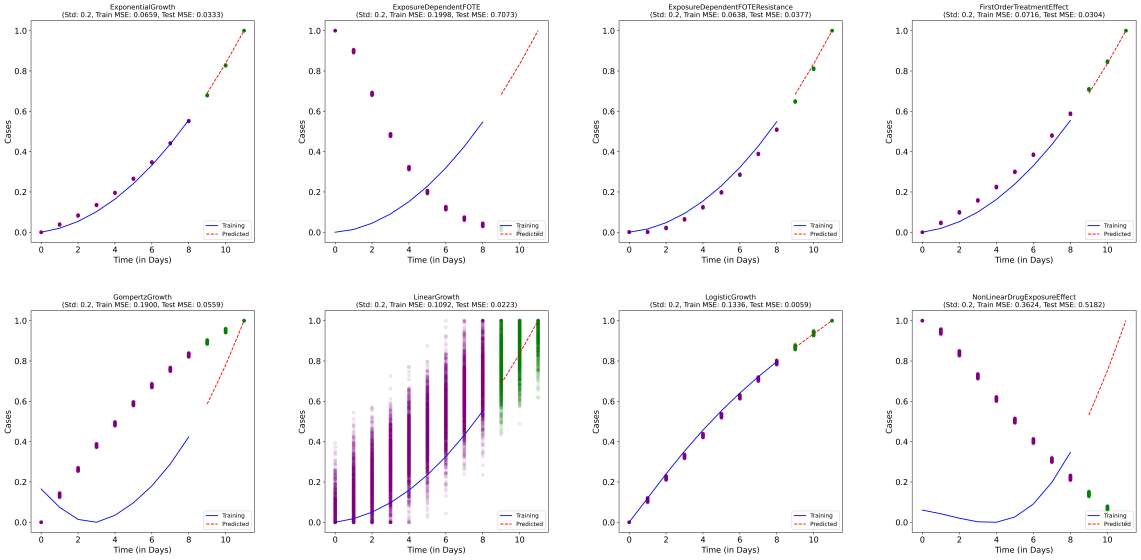


Figure 8: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.2$ using 12 qubits.

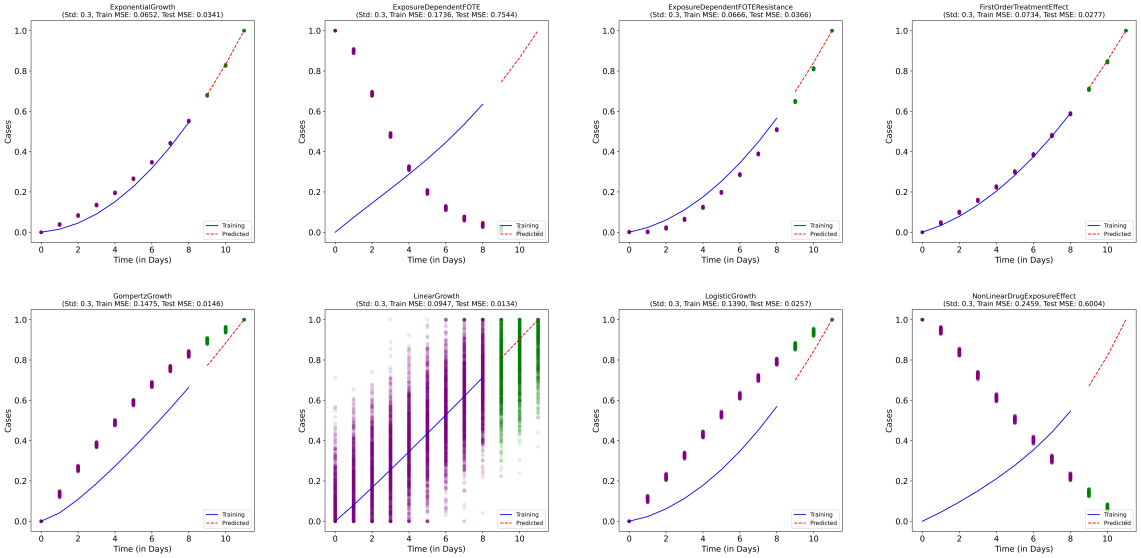


Figure 9: Numerical Results of the Proposed Models on Synthesized Dataset with Standard Deviation of Noise $\sigma = 0.3$ using 12 qubits.

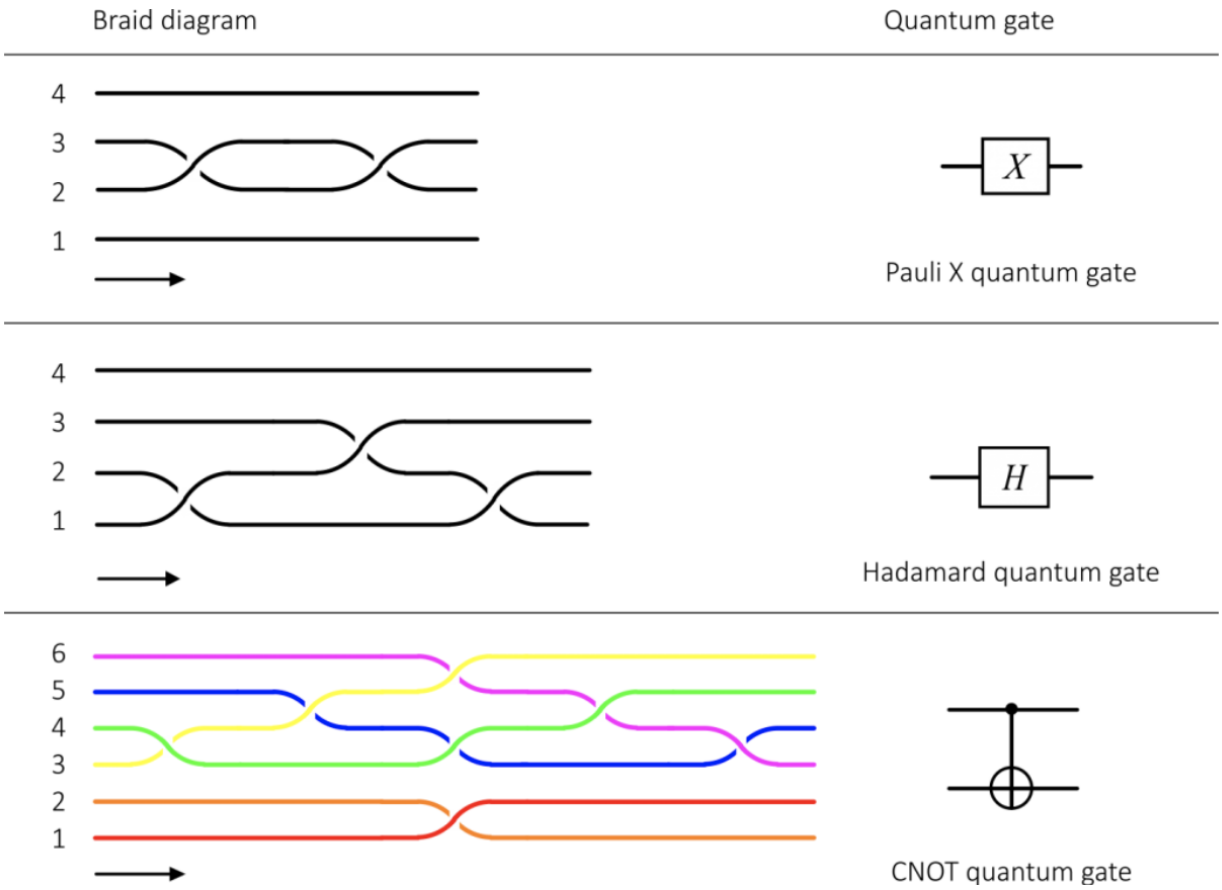


Figure 10: The Connection Between Braid Group and Quantum Gates [17]. Figures Courtesy: https://www.quantum-bits.org/?p=2226

D Connection Between Braid Group and Quantum Logical Gates