

Improving Quantum Circuit Synthesis with Machine Learning

Mathias Weiden¹, Ed Younis², Justin Kalloor¹, John Kubitowicz¹, and Costin Iancu²

¹Department of Electrical Engineering and Computer Science, University of California, Berkeley
 {mtweiden, jkallor3, kubitron}@cs.berkeley.edu

²Computational Research Division, Lawrence Berkeley National Laboratory
 {edyounis, cciancu}@lbl.gov

Abstract—In the Noisy Intermediate Scale Quantum (NISQ) era, finding implementations of quantum algorithms that minimize the number of expensive and error prone multi-qubit gates is vital to ensure computations produce meaningful outputs. Unitary synthesis, the process of finding a quantum circuit that implements some target unitary matrix, is able to solve this problem optimally in many cases. However, current bottom-up unitary synthesis algorithms are limited by their exponentially growing run times. We show how applying machine learning to unitary datasets permits drastic speedups for synthesis algorithms. This paper presents QSeed, a seeded synthesis algorithm that employs a learned model to quickly propose resource efficient circuit implementations of unitaries. QSeed maintains low gate counts and offers a speedup of $3.7\times$ in synthesis time over the state of the art for a 64 qubit modular exponentiation circuit, a core component in Shor’s factoring algorithm. QSeed’s performance improvements also generalize to families of circuits not seen during the training process.

I. INTRODUCTION

Quantum computers are capable of solving many problems that are thought to be intractable on classical computers [17]. In the Noisy Intermediate Scale Quantum (NISQ) [21] era, where full error correction is not yet realizable, careful consideration of error mitigation techniques in compilation are vitally important. To this end, techniques such as random compilation [30], noise-adaptive mapping [16], [28], and zero noise extrapolation [6] have been proposed. Google’s quantum supremacy experiments demonstrated empirically that simply reducing the number of operations remains an impactful noise mitigation strategy [2]. The need for optimization algorithms that reduce gate counts is therefore critical.

Rule based circuit optimization algorithms are widely used because they are quick and effective for small circuits [1], [9], [13], [20]. These techniques rely on recognizing small gate patterns and replacing them with more efficient implementations, and by making small, heuristic-guided transformations. Recently, deep Reinforcement Learning has emerged as a technique to automate the discovery and application of circuit transformation rules [5], [18], [23]. However, these techniques remain hindered by a narrow, local view of the circuits they optimize or by the limitations of human intuited heuristics.

Unitary synthesis is a bottom-up method for constructing circuit implementations of unitary matrices. It has become popular as an optimization technique as it eschews a gate

level view of the program for a more global perspective. This broader frame of reference allows for optimization that is otherwise impossible with rule based methods. State-of-the-art synthesis algorithms have demonstrated that they are capable of producing the best known implementations of a variety of circuits [3], [27], [35]. Synthesis searches over a tree of increasingly more complex candidate circuits and evaluates each one’s ability to implement a unitary target. Yet, these techniques remain limited in their applicability due to their exponentially growing run times. Circuits wider than five qubits must be partitioned into smaller subcircuits, which can each be individually synthesized. Importantly, even in these cases, where synthesis is not given a full global view of the circuit being optimized, it still provides unmatched optimization results [31], [32].

The core contribution of this paper is a method for replacing the expensive circuit tree search of unitary synthesis with Machine Learning (ML) inference. Given a unitary matrix, we immediately propose resource efficient circuit templates that, with high probability, are able to directly implement the unitary. We present our findings relating to the learnability of certain sets of unitary matrices. To our knowledge, we are the first to apply ML to unitary datasets for optimization. We believe that our approach is generic and will apply to the field of quantum compilation no matter how quantum programming evolves. This project focuses on demonstrating the advantages of leveraging ML for the synthesis of three qubit unitaries, but remains both applicable and promising for accelerating synthesis of larger matrices. With the use of circuit partitioning, we demonstrate how QSeed is capable of scaling to circuits of 100 qubits and beyond. Our technique provides substantial speedups compared to state-of-the-art synthesis algorithms, ensuring more practical applicability of this style of optimization.

The paper is divided as follows: Section II describes the unitary synthesis process and illustrates its scalability limitations. Section III discusses our findings on applying ML datasets of unitaries. Section IV introduces QSeed, a seeded synthesis algorithm that uses ML to accelerate the synthesis tree search process. Section V evaluates QSeed’s performance. Section VI provides further discussion related to QSeed’s performance and future research directions.

II. BACKGROUND

In this section we present necessary context and provide definitions for notable concepts.

A. Unitary Synthesis and Parameterized Quantum Circuits

Unitary synthesis algorithms are a class of algorithms that take as input a unitary matrix and iteratively build up a parameterizable quantum circuit implementing that unitary.

Definition II.1 (Unitary Synthesis). *Given a target unitary matrix $U_T \in \mathbb{U}(N)$, a unitary synthesis algorithm constructs a quantum circuit that implements a unitary U_S satisfying*

$$\|U_T - U_S\|_{HS} \leq \epsilon$$

where $\|\cdot\|_{HS}$ is the Hilbert-Schmidt norm defined as

$$\|A\|_{HS} = \text{Tr}(A^\dagger A).$$

Definition II.2 (Parameterized Quantum Circuit). *A Parameterized Quantum Circuit (PQC) is a quantum circuit that contains gates whose operations are a function of real valued parameters.*

Unitary synthesis algorithms such as QSearch [3] and LEAP [27] are capable of finding the best known implementations of a variety of quantum algorithms. In these synthesis algorithms, sequences of parameterizable multi-qubit gates are appended to a simple base PQC. This style of synthesis is referred to as bottom-up synthesis and is depicted in Fig. 1. The allowed placement of these gates is determined by a *qubit topology*.

Definition II.3 (Qubit Topology). *A qubit topology is a graph $G = (V, E)$. Vertices $v \in V$ represent qubits, while edges $(u, v) \in E$ represent (hardware) allowed interactions between qubits.*

The bottom-up synthesis process can be viewed as a search through a tree of parameterizable circuits. Traversal begins at the tree's root node, and child nodes are explored until a circuit capable of implementing the target unitary is found. When visiting a node in this traversal, the parameters in the associated circuit must be solved for in order to minimize the

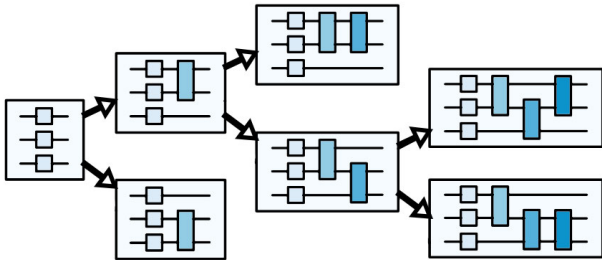


Figure 1: In bottom-up synthesis, parameterizable sequences of gates are appended to a base PQC. Circuits grow according to a circuit tree where parent circuits serve as prefixes to child circuits.

distance described in Definition II.1. This process of solving for parameters is called *instantiation*. This computationally expensive process can be thought of as deciding whether or not a PQC is suitable for the implementation of a unitary.

Definition II.4 (Instantiation). *Given a target unitary U_T and a PQC represented by a unitary operator $U(\theta)$ parameterized by $\theta \in \mathbb{R}^m$, instantiation solves the problem*

$$\arg \min_{\theta} \|U_T - U(\theta)\|_{HS}.$$

B. Run Time Expense

The run time of unitary synthesis algorithms such as QSearch and LEAP grows exponentially in the number of qubits. This is due in part to the fact that the size of the unitaries that must be evaluated during instantiation grows as $O(2^n)$. Furthermore, the maximum number of CNOT gates needed to implement an arbitrary unitary is $\frac{1}{4}[4^n - 3n - 1]$ [25]. This means the synthesis circuit tree's depth grows exponentially. As the size of the tree grows, synthesis algorithms also tend to explore regions that do not contain solution circuits. In these scenarios, synthesis must backtrack up the tree and probe new regions. These things combined mean the amount of time spent searching the synthesis circuit tree quickly becomes untenable.

C. Circuit Partitioning

The exponentially scaling run time of bottom-up synthesis algorithms limits the width of, or number of qubits in, circuits that can be synthesized. Wide quantum circuits must be split up, or *partitioned*, into subcircuits that can be synthesized independently. Doing so enables synthesis to tackle optimization problems up to 100 qubits and beyond. An example of a partitioned circuit can be seen in Fig. 2. The partitioning strategy used is described in the QUEST and TopAS projects [19], [31].

By partitioning, the synthesis of a single large quantum circuit is broken into many smaller pieces that can be handled individually and in parallel. The run time of synthesis algorithms applied to partitioned circuits is no longer limited by the exponentially growing cost of unitary function evaluation, but by the number of partitions formed. This number scales polynomially with a circuit's depth and width. Even so, with large circuits, the number of partitions formed can be high enough to cause intractability, even without this exponentially scaling cost.

III. LEARNING ON UNITARY MATRICES

Here, we present our findings relating to the learnability of datasets which consist of unitary matrices. Say we have n qubits, and a set of unitaries $\mathcal{U}(2^n)$ that represent operations we can perform on our qubits. For our n qubits, we can also construct a set that contains all parameterized quantum circuits with K or fewer multi-qubit gates. We call this set $\mathcal{PQC}(n, K)$.

What we wish to do is find a mapping

$$\phi : \mathcal{U}(2^n) \longrightarrow \mathcal{PQC}(n, K)$$

such that each unitary $U \in \mathcal{U}(2^n)$ is mapped to the circuit $PQC \in \mathcal{PQC}(n, K)$ that implements U with the fewest number of multi-qubit gates. We assume that K is large enough so that some circuit in $\mathcal{PQC}(n, K)$ implements U to within some error ϵ . In this section, we describe the efforts we have made to solve this problem by training a deep neural network to learn this mapping ϕ .

A. Canonical Representation of Unitaries

It is a well known fact of quantum mechanics that two unitaries that differ only by a global phase will have identical measurement statistics [17]. For the sake of easing the training of a neural network, it is desirable to have a representation of unitary matrices such that any two unitaries that differ only by a global phase are represented the same way. This way unitaries that are, for all intents and purposes, the same, are represented identically. Formally, we want a function f such that for unitaries U, U^* ,

$$\forall \theta \in \mathbb{R} \quad U = e^{i\theta} U^* \implies f(U) = U^*.$$

If $f(U) = U^*$ we call the matrix U^* the canonical representation of U .

The process of computing $f(U) = U^*$ happens in two steps:

- 1) Convert U to a special unitary $V \in \text{SU}(N)$ so that

$$V = e^{i2\pi k/N} U^*$$

for some integer k .

- 2) Find the integer k and multiply

$$U^* = e^{-i2\pi k/N} V$$

so that the complex phase of the first nonzero element of V is cancelled.

Performing these two steps ensures that each unitary is mapped to a canonical special unitary.

Proof. The first step ensures that the determinant $|V| = 1$. However, each $U \in \mathcal{U}(N)$ could be mapped to any one of the N possible special unitaries in the form $V = e^{i2\pi k/N} U^*$ for $k \in \mathbb{Z}$. To ensure that $f(\cdot)$ yields the same special unitary no matter the global phase, we perform the second step. Even with different global phases, two special unitaries $V_1 = e^{i2\pi k_1/N} U^*$ and $V_2 = e^{i2\pi k_2/N} U^*$ that differ only by some root of unity global phase will have nonzero elements in the same locations. Because V_1, V_2 are unitary, there must be at least one nonzero

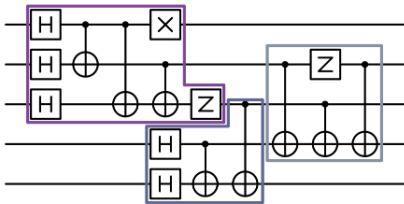


Figure 2: A width five quantum circuit that has been partitioned into subcircuits of three qubits.

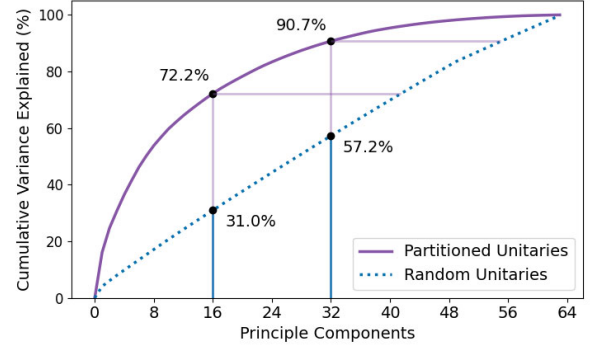


Figure 3: Cumulative variance explained by Principle Components of two datasets of three qubit unitaries. One is randomly generated, the other is made of partitioned unitaries from circuits in Table I. The PCA of the partitioned unitaries implies these unitaries lie on a low dimensional manifold; there are patterns that can be identified in this dataset. The random unitaries do not exhibit identifiable low dimensional patterns. Learning on random unitaries is difficult, learning on partitioned unitaries is easier.

element in each of their rows and columns. Ensuring that the complex phase of the first elements of V_1, V_2 are the both zero will yield the same total global phase. Thus,

$$f(V_1) = f(V_2) = U^*$$

□

B. Dimensionality of Unitary Datasets

Various works in the fields of Variational Quantum Algorithms (VQAs) and Quantum Machine Learning (QML) have demonstrated the high expressive power of parameterized circuits [4], [7], [11], [26]. A single PQC can express many unitaries, and one unitary is likely to be implementable by many PQCs. This makes the problem of learning a mapping from unitaries to circuits challenging.

To better convey how this issue relates to the learnability of unitary datasets, we can examine a set of random unitaries generated by PQCs. Doing so allows us to make observations about the underlying dimensionality of the unitaries that a set of PQCs can implement. This is helpful because data that exhibit some form of low dimensionality are more easily learned by neural networks [24]. To quantify dimensionality, we can perform a Principle Component Analysis (PCA) [12] on this set of unitaries. Fig. 3 illustrates the variance explained by the principle components of a set of $2^3 \times 2^3$ complex unitaries randomly generated by twelve PQCs. As the variance in this dataset is spread evenly across each principle component, we have little expectation that we can learn a generalizable mapping from unitaries to PQCs for this dataset.

In practice, we need only learn a mapping from unitaries to PQCs for unitaries and PQCs of interest. Consider for example the task of optimizing wide quantum circuits taken from a diverse benchmark suite. If we partition the circuits outlined

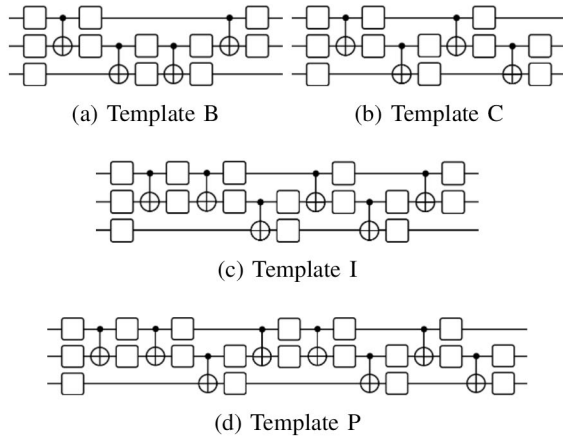


Figure 4: Labeled PQC templates that commonly appear when synthesizing unitaries. Labels in these figures correspond to the PQC Template Labels in Fig. 5. Squares represent parameterized U3 gates.

in Table I into unitaries of three qubits as mentioned in Section II, we obtain a dataset which is far lower dimensional than in the randomly generated unitary case. Fig. 3 illustrates how linear combinations of the first 16 principle components of this partitioned unitary dataset describe 72.2% of the data’s variance.

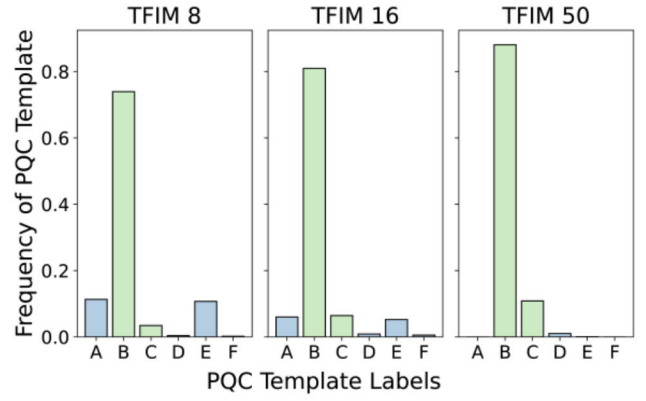
Because unitaries taken from partitioned quantum circuits exhibit patterns that lie on low dimensional manifolds, we expect that learning a mapping from these unitaries of interest to the PQCs that can express them will be possible.

C. Common Structures in Partitioned Quantum Circuits

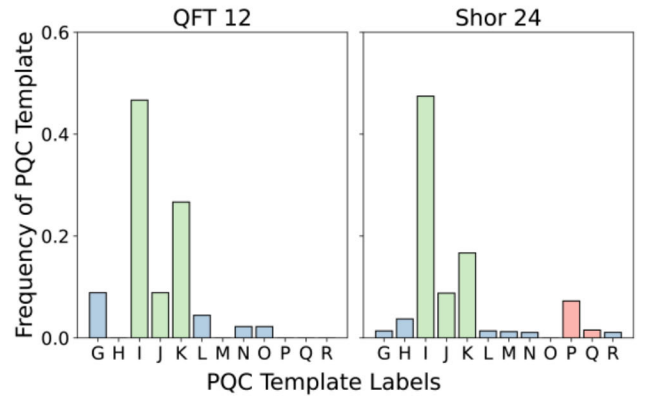
Across many runs of synthesis, we have observed two patterns. First, partitioned unitaries from different width circuits of the same family tend to have similar characteristics. Second, unitaries from circuits that are related but not from the same family also tend to share characteristics. Critically, both of these points remain true even when the unitaries from these circuits are not identical. We can begin to measure this by examining what PQCs commonly appear once unitaries from various circuits are synthesized. Fig. 4 illustrates four PQCs that often implement partitioned unitaries.

Fig. 5a illustrates the first point. Across multiple circuit widths, synthesized unitaries from the TFIM circuit family have shared structure. This figure shows the distribution of PQCs that implement unitaries from these TFIM circuits. Fig. 5b illustrates the second point. In this case, the modular exponentiation “Shor” and QFT circuits have replicated modes in their distributions of selected PQC templates. Unitaries taken from partitioned QFT and modular exponentiation circuits are similar, but not identical. We can also see from Fig. 5b that the modular exponentiation circuit has added complexity compared to the QFT, as seen by the extra mode around the PQC labeled “P” (Fig. 4d).

We expect that a learned mapping might be advantageous in this scenario as there are clearly shared characteristics across



(a) PQC templates that appear in synthesized TFIM circuits of various widths. Circuits of different widths share characteristic modes around the PQCs labeled B and C.



(b) PQC templates that appear in synthesized QFT and modular exponentiation circuits. These related circuits have similar modes around templates I, J, and K. The more complex Shor’s modular exponentiation circuit has extra templates not found in the QFT.

Figure 5: Histograms of PQC templates that appear in the optimized TFIM circuits of various widths (5a), and QFT and Shor modular exponentiation circuits (5b).

related circuits. Because the unitaries in these circuits are not the same, memorizing unitary-PQC pairs is not helpful in predicting what structures might appear for new, unseen, unitaries. Training a model on a variety of circuit families and across many circuit widths aids us in learning a mapping that is robust to unseen unitaries. Section IV-C further discusses the efforts we have made to highlight this phenomenon.

IV. QSEED

This section presents QSeed, a unitary synthesis tool which consists of two major components:

- 1) The recommender model: a deep neural network which uses information about a target unitary to propose PQCs that are likely to instantiate the target unitary.
- 2) Seeded synthesis: a unitary synthesis algorithm which begins the synthesis process at a provided seed circuit.

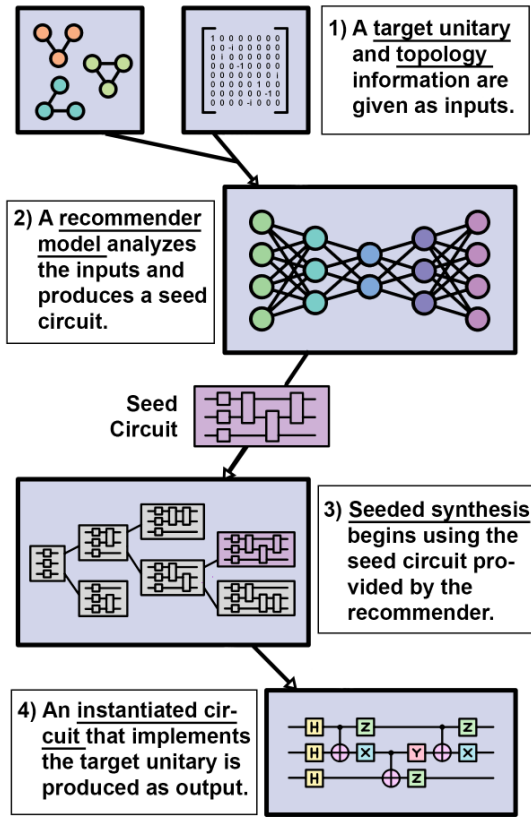


Figure 6: QSeed operational flow graph. QSeed takes as input a target unitary and topology information, and uses a recommender model to produce a seed PQC. This circuit is used as the starting point for seeded synthesis. The output is a circuit that implements the target unitary.

The programmatic flow of the QSeed algorithm is depicted in Fig. 6.

A. The Recommender Model

We have implemented a recommender model as a deep autoencoder neural network. This choice was made based off the observation that unitary matrices taken from partitioned quantum circuits exhibit low dimensionality. Section III-B empirically illustrates this observation by use of a PCA of unitaries taken from a wide variety of partitioned quantum circuits. Dimensionality reduction through the use of neural networks to learn low-dimensional manifolds is widely done in practice [10].

The three qubit recommender model takes as input a unitary matrix U , finds the canonical representation of the unitary (see Section III-A), then flattens the real and complex elements of the unitary into a vector $x_U \in \mathbb{R}^{128}$. The autoencoder output is a vector $y_U \in \mathbb{R}^{|\mathcal{PQC}|}$. Each element in the model's output corresponds to a PQC in the set $\mathcal{PQC}(n, K)$. For $n = 3$ qubit linear topologies and a maximum CNOT count of $K = 10$, $|\mathcal{PQC}(n, K)| = 1199$. The value of $\text{softmax}(y_U)_i \in [0, 1]$ represents the model's confidence that template i can implement unitary U .

Circuit Name	Training Widths	Test Widths
add	17, 65	41
grover	-	10
heisenberg	4, 6, 7, 8, 16, 32, 64	5
hhl	8	6
hubbard	4, 18, 50	8
mult	8, 32, 64	16
qae	11, 33, 101	65
qft	3, 4, 8, 16, 32	64
qml	4, 25, 60, 108	128
qpe	6, 10, 14	18
shor	16, 32	64
tfim	3, 4, 5, 6, 7, 8, 16, 32,	64
vqe	12, 14	18

Table I: Split of circuits withheld from recommender training for testing. No Grover's algorithm circuits are used in training.

During inference, the recommender model compresses the input unitary from an \mathbb{R}^{128} to an \mathbb{R}^{32} representation. Fig. 3 indicates that for our dataset of three qubit partitioned unitaries, this representation is capable of capturing at least 90.7% of the variance within these matrices.

B. PQC Templates

As mentioned in Section IV-A, the output of the recommender model is a vector of size 1199, each element corresponding to a possible output PQC template. These templates are generated by traversing the synthesis circuit tree in a breadth-first manner and enumerating each circuit. This recommender model was designed for the task of mapped circuit optimization, so only three qubit linear topologies are considered. Circuits with more than three consecutive two-qubit gates applied to the same qubits are not included, as three CNOTs are sufficient for implementing any two-qubit unitary [29]. All PQC templates conforming to these restrictions with zero to eight CNOT gates are considered. Example PQC templates are illustrated in Fig. 5. This paper only considers templates containing CNOT gates, but other multi-qubit gates can surely be used instead.

C. Training the Recommender

The recommender model is trained on a dataset of three qubit unitaries partitioned from twelve different circuit families. Each family of circuits contained multiple benchmarks of various widths. Each unitary in the dataset is paired with one PQC for each topology that is targeted. These "label" PQCs are found by running synthesis and ensuring the unitary can be implemented by that PQC.

To ensure generality across circuit widths, benchmarks from each circuit family were withheld from training. The test and training circuits are listed in Table I. Note that no Grover's algorithm circuits were included in the training dataset.

The recommender model was pretrained as a denoising autoencoder with the task of reconstructing noisy input unitaries. Afterwards, a PQC template selection head is appended to the encoder portion of the network, and the network is then fine tuned to learn the mapping from unitaries to PQCs. The network must be conditioned on qubit topology information

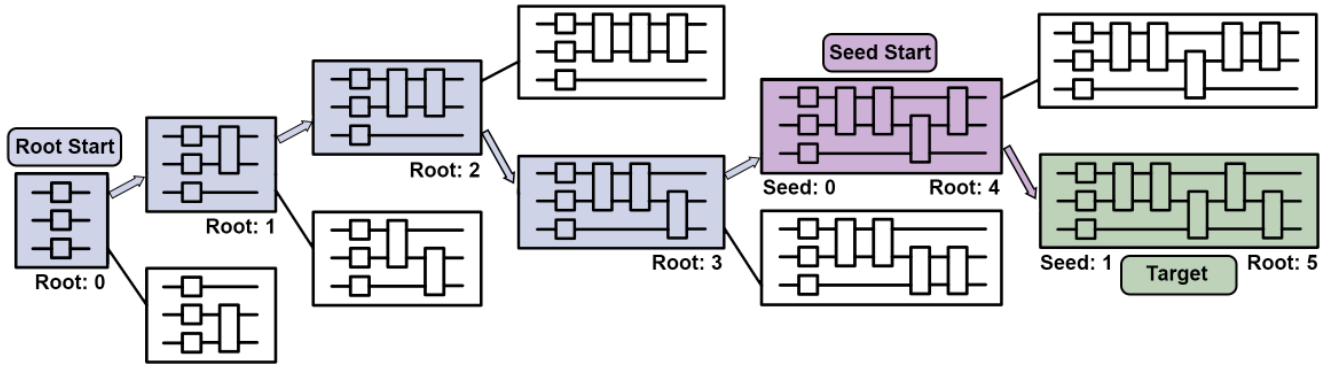


Figure 7: Seeded synthesis provides speed ups compared to synthesis which begins at the root because it allows for the circuit search process to start closer to the target. In this example, QSearch begins synthesis at the node labeled “Root Start”, and requires 5 hops (highlighted in blue) before reaching the target. Seeded synthesis instead begins at the node labeled “Seed Start”, and only requires 1 hop (highlighted in purple) before reaching the target.

in order to propose seeds conforming to different connectivities. After training, the network was able to reach a top-3 classification accuracy of 60%. This classification accuracy is deceptively low. As each unitary in the dataset is only associated with a single PQC label, even though multiple PQCs (often of the same depth) are capable of implementing a given unitary, good seed proposals may be classified as incorrect. Further tuning of the recommender network’s architecture and hyperparameters for higher classification accuracy is surely possible.

D. Seeded Synthesis

As described in Section II, there are two primary processes that happen during unitary synthesis: instantiation and tree traversal. This section describes how using seeded synthesis can alleviate the extensive run time of bottom-up synthesis by accelerating the tree traversal process.

Bottom-up synthesis algorithms, such as QSearch and LEAP, begin the circuit tree exploration process with a circuit that contains only single qubit rotation gates. This circuit is the root node of the circuit tree depicted in Fig. 1. Each node in the tree describes a different PQC that can be built by appending gates onto the root circuit.

It may be the case that some desired target circuit/node is situated very far down the tree. Fig. 7 provides an example of such a case. Here, the path from the “Root Start” node to the “Target” node has distance five. Each of the six circuits along this path must be instantiated before a solution is found.

Seeded synthesis begins the circuit tree traversal at a non-root node. In Fig. 7, the “Seed Start” node is highlighted in purple. The path from this node to the target has distance one, and only two instantiation calls are made, one for each circuit along this path. When good seed circuits are provided, seeded synthesis can perform synthesis far faster than bottom-up synthesis, as fewer parameterized circuits must be instantiated. In practice, many deep parameterized circuits are very likely to be able to implement some target unitary. This means that

proposing “bad” seed circuits, which correspond to high gate count solutions, will still provide speed ups. This phenomenon is further explored in Section V.

In order to account for scenarios where a given seed circuit contains too many gates, synthesis is also capable of traversing up the circuit tree. This means that gates can also be removed to examine less deep candidate circuits. The tree search process advances both towards the root and towards leaf nodes.

V. EVALUATION

To highlight the effectiveness and demonstrate the use of QSeed, we have evaluated it on the task of wide quantum circuit optimization. There are several proposed methods for using unitary synthesis to optimize wide quantum circuits, including QGo [32] and TopAS [31]. In both cases, wide quantum circuits are partitioned into smaller subcircuits which are synthesized independently. We demonstrate how using QSeed can accelerate QGo style optimization of mapped quantum circuits, while maintaining low multi-qubit gate counts. A partition width of three qubits was used. All benchmarks were run on a 64 core AMD EPYC 7702P processor. The synthesis tools that are evaluated here are available through the *BQSKit* toolkit [34].

A. Benchmark Generation

For our training and test benchmark circuits, we generated a group of circuits that represent a wide breadth of important high level algorithms. These benchmarks can generally be grouped into variational algorithms, Quantum Fourier Transform (QFT) based algorithms, Quantum Machine Learning (QML) algorithms, and algorithms that show promise in the field of Quantum Finance. For our variational algorithms, we used Qiskit Nature to generate an ansatz for the ground state energy of molecular interaction [22]. We then simulated VQE to generate solved parameters for LiH, BeH2, BH3, and CH4 molecules. For QML, we first have an ZZ feature map

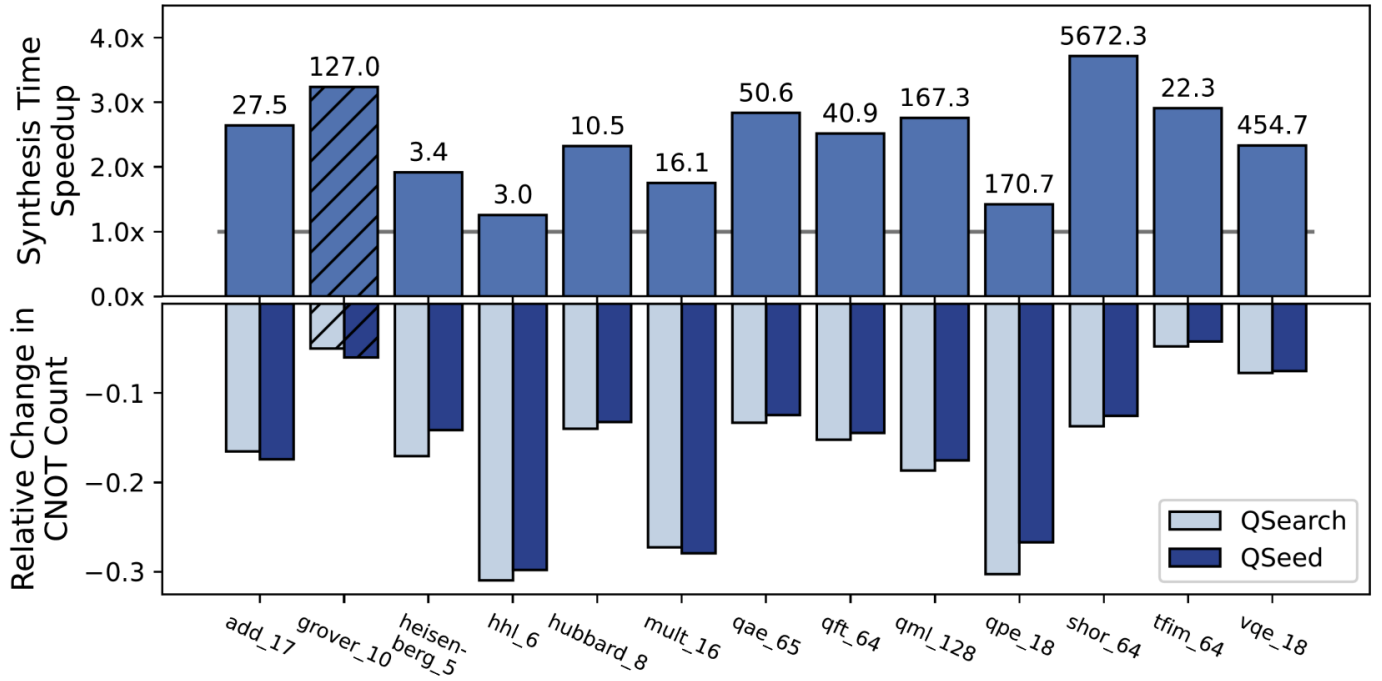


Figure 8: QSeed compared to QSearch on test circuits. The top graph shows speedup in wall-clock synthesis time and absolute time savings in seconds as data labels. Larger speedups are better. The bottom graph shows relative changes in CNOT gate count relative to Qiskit level 3 optimized circuits. Lower relative change in CNOT counts are better. Across all test circuits, QSeed has an average speedup of $2.4\times$, and an average of 1.0% more CNOT gates compared to QSearch. Synthesis time includes the time needed for QSeed’s recommender model do inference. These circuits were withheld from the training process (see Table I), and no Grover’s algorithm circuits were included in the training dataset. These values represent an average of three trials.

to encode classical bit data into quantum data, then a two local network layer is applied. Our QFT based algorithms include standalone QFT, Shor’s modular exponentiation, and general Quantum Phase Estimation (QPE) circuits. For quantum finance, we generated Quantum Amplitude Estimation (QAE) circuits which are commonly used for Monte Carlo Integration. Additionally, we generate Harrow-Hassidim-Lloyd (HHL) circuits applied to random square matrices as an example of a quantum linear solvers used in the field. [8]. Grover’s algorithm circuits were implemented according to [15]. These circuits are able to be scaled to various widths and partitioned to fit into our quantum synthesis process.

B. Compilation of Test Circuits

We evaluate the optimization of circuits based on two metrics: the time needed to perform synthesis and the number of multi-qubit gates in the optimized circuits. Each benchmark was optimized with Qiskit’s level 3 transpiler optimization [22] and mapped to a square mesh qubit topology using the SABRE Swap algorithm [14]. After synthesizing with QSearch or QSeed, each block is also quickly optimized with scanning gate removal, as described in [33].

Fig. 8 compares QSeed and QSearch using these two metrics on the suite of test circuits outlined in Table I. On average, QSeed offers a speedup of $2.4\times$ compared to QSearch across

all test circuits. The 64 qubit Shor’s modular exponentiation circuit sees a speed up of $3.7\times$. In wall-clock time, this is a savings of nearly 5700 seconds. The 18 qubit VQE also notably sees a speedup of $2.3\times$. We have highlighted the VQE circuit results because due to the classical-quantum hybrid style of computation in variational algorithms, savings in synthesis time is of particular importance. These circuits offer large speedups due in part to their large sizes. There is a strong correlation between the number of partitions in a circuit and the amount of speedup compared to QSearch. On average, QSeed suffers a 1.0% decrease in improvement in CNOT gate counts compared to QSearch. We feel that this difference is acceptable, and that with better recommender training techniques the gap between QSearch and QSeed optimized circuit gate counts can be closed.

The speedup of the 64 qubit Shor’s modular exponentiation circuit is particularly pronounced. The reason for this speedup can be understood by observing the distribution of PQC structures that appear after synthesis. As shown in Fig. 5, the 24 qubit modular exponentiation circuit’s PQC distribution contains a mode on the PQC template labeled “P”. As the width of the modular exponentiation circuit increases, this PQC becomes more prevalent. As it contains 8 CNOT gates, QSearch requires at minimum (and in practice, often more than) 9 instantiation calls to synthesize these partitions.

QSeed is often able to do this in a single instantiation call. Circuits such as the the 6 qubit *HHL*, and the 18 qubit *QPE* circuits see comparatively little improvement in synthesis time. This is due to the fact that synthesizing partitioned unitaries from these circuits requires very few CNOTs. The cost of doing inference before synthesizing these unitaries diminishes potential speedups. The number of partitions, as well as the number of CNOTs in each partitions, but influence the amount of speedup QSeed delivers.

C. Generalizing to Grover's

No unitaries from Grover's algorithm circuits of any width were included in the training dataset for the recommender model. The compilation of the 10 qubit Grover's algorithm benchmark still results in a speedup of $3.2\times$, which has been emphasized with hatches in Fig. 8. This is a clear demonstration that QSeed is capable of offering large speedups in synthesis time even for completely unfamiliar circuits.

D. Result Verification

As outlined in Definition II.1, unitary synthesis allows for some tunable degree of approximation through the setting of the threshold constant ϵ . It is therefore necessary that results be verified to ensure an acceptable amount of approximation error has accumulated in synthesized circuits. The results of [19] prove that the total accumulated error of a partitioned circuit is upper bounded by the sum of the individual approximation errors of each partition. To verify that the optimized circuits accurately implement the same functionality as the original circuits, we re-partition the optimized and original circuits into 12 qubit subcircuits, then sum the Hilbert-Schmidt distances of each corresponding 12 qubit block. The results of this experiment are summarized in Table II. Each compiled test circuit accurately implements the desired functionality, and approximation errors remain many orders of magnitude lower than any error that might be introduced by noise.

Circuit	Upper Bound on Approximation Error
add 41	1.4×10^{-9}
grover 10	2.4×10^{-13}
heisenberg 5	5.6×10^{-15}
hhl 6	2.2×10^{-16}
hubbard 8	3.1×10^{-14}
mult 16	2.5×10^{-14}
qae 65	3.6×10^{-9}
qft 64	2.4×10^{-9}
qml 128	4.2×10^{-14}
qpe 18	6.8×10^{-13}
shor 64	3.1×10^{-7}
tfim 64	1.8×10^{-15}
vqe 18	7.8×10^{-16}

Table II: Cumulative error due to synthesis approximation in test circuits. The Shor benchmark has the largest amount of error as it contains the greatest number of partitions. Each optimized circuit accurately recreates the function of the corresponding original test circuit.

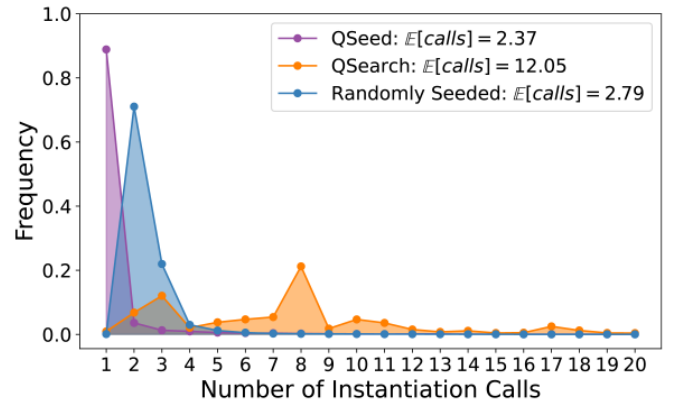


Figure 9: Distribution of instantiation function calls made. QSeed uses the recommender model while Randomly Seeded selects a random seed circuit for seeded synthesis. Both exhibit speed ups over QSearch because they decrease the expected number of instantiation calls. The X-axis is truncated at 20 instantiation calls, but continues up to 1704 for QSearch.

E. Evaluating the Recommender

Section V-B empirically demonstrates the advantages of using QSeed for wide quantum circuit optimization. However, one can imagine other seed recommendation schemes to accelerate synthesis. For example, instead of using a network to produce seeds, one could simply randomly select seeds from a set of parameterized quantum circuits.

Beyond measuring the time needed to perform synthesis, the speed of a seed selection strategy can also be compared by measuring the number of instantiation calls needed before a solution circuit is found. As mentioned in Section II, instantiation must be performed for each candidate circuit in the synthesis process, and its complexity grows exponentially with the number of qubits in the circuit being synthesized. Synthesis is performed faster when fewer instantiation calls are made.

Fig. 9 compares the number of instantiation calls needed to synthesize blocks in the suite of benchmark circuits shown in Table I for three recommendation schemes. The first strategy is a random seed strategy, where three random circuits are sampled to act as synthesis seeds for each block. The second is the static QSearch strategy, where synthesis of each block begins at the circuit tree root as described in Section II. The final strategy is the QSeed recommender model described in Section IV. Echoing the results of Fig. 8, QSearch tends to require more instantiation calls and takes more time to synthesize circuits. Similar to QSeed, the random initialization strategy tends to require very few instantiation calls. However, randomly selecting seeds degrades solution quality. Fig. 10 illustrates the distribution of growth/shrinkage across all synthesized partitions. QSeed is able to match the quality of QSearch very closely. We can expect that both will decrease CNOT counts on average. The randomly seeded strategy tends to result in growth, as signified by the expected relative CNOT

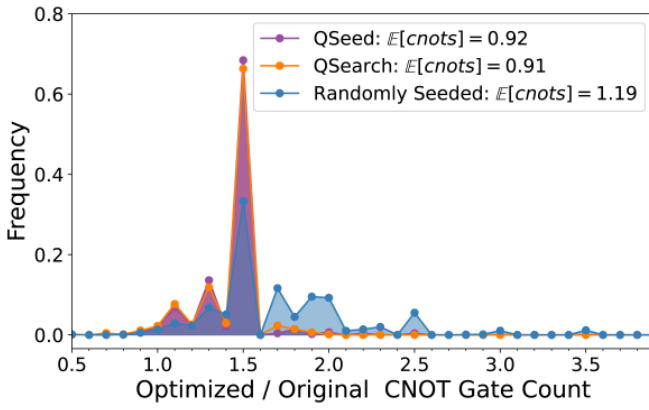


Figure 10: Distribution of relative CNOT gate counts. Values less than one correspond to fewer gates in the optimized compared to the original circuits (lower is better). The distribution of relative CNOT counts for QSeed closely matches that of QSearch. Randomly Seeded often results in circuit growth, which corresponds to values larger than 1.

count of 1.19. An added benefit of seeded synthesis is that multiple seed locations can be considered simultaneously. Both the randomly seeded and QSeed recommender use this to their advantage, while QSearch must always begin synthesis at the same, single root seed.

As mentioned in Section IV-C, the recommender model was trained until an accuracy of 60% was reached on a random subset of unitaries from the test circuits (see Table I). Even so, Fig. 9 demonstrates that most unitaries can be synthesized with a single instantiation call. This observation parallels the results of the random seed case, proposing any seed tends to lead to speedups compared to QSearch. The differences in CNOT gate counts highlighted in Fig. 8 can therefore likely be diminished with a more accurate recommender model.

Overall, the results of Fig. 9 and Fig. 10 illustrate how QSeed is able to maintain speed advantages over QSearch, with minimal compromise in gate count.

VI. DISCUSSION

In this section, we provide further discussion relating to recommender generalization, the performance of QSeed, and the future of machine learning in quantum circuit optimization.

A. Recommending General Unitaries

This paper demonstrates that learning can be done for a subset of unitaries, unitaries taken from partitioned quantum circuits. How can this approach be made more generic and applied beyond the domain specific case? Our observations outlined in Section III detail how partitioned circuits can be characterized by distributions of PQC templates. The similarity between circuits of different domains and sizes can be evaluated using these distributions. We conjecture that across all algorithms expressed as quantum circuits, the prevalence of PQCs is not uniform. Certain templates are inherently more common than others. It is sufficient that we learn to

associate unitaries with these important PQCs, rather than all PQCs. This explains why QSeed’s recommender model is capable of generalizing to unseen circuits. With a well trained recommender model, it is possible that a few-shot learning approach, where only a small set of unitaries from unseen circuits are given, could further improve speedups and gate reductions. The results of Section V, especially the compilation of the completely unfamiliar 10 qubit Grover’s circuit, demonstrate that learned mappings on this partitioned unitary dataset generalize to unseen partitioned unitaries.

The KAK decomposition demonstrates how to decompose arbitrary two-qubit unitaries [29]. Using ML to propose circuit templates differs in that it is not limited to two qubits, but it is not expected to work for general unitaries. A PCA of random unitaries shows that patterns are difficult to identify for the general (random) unitary case (Fig. 3).

B. Scaling Seeded Synthesis Width

The results for this paper demonstrate a seeded synthesis algorithm for three qubit unitaries. Partitioning quantum circuits with different strategies and with larger partition widths still support the results illustrated in Fig. 3. Unitaries of interest lie on low dimensional manifolds. Therefore, we expect that learning is still possible in these cases. The process of mapping unitaries to the PQCs that implement them becomes more challenging as the number of qubits increases. As discussed in Section II, the number of gates needed to implement an arbitrary unitary of n qubits grows with $O(4^n)$ complexity. If done in a topology-aware manner, more qubit topologies, and vertex labelings within these topologies must also be considered. Therefore, to apply QSeed beyond four or five qubits, more work must be done to narrow this space of possible output PQCs.

To this end, proposing circuit prefixes instead of PQCs that are expected to implement a unitary directly could still provide benefits. As illustrated in Fig. 7, this corresponds to proposing PQCs that are simply closer to the target circuit than the root node.

As previously stated, we conjecture that a relatively small set of PQC templates are needed to implement all unitaries of interest. In the three qubit case, although there are 1199 possible PQCs to choose from (see Section IV-B), only about 415 appear for more than 10 instances in our dataset of over 203,000 unitaries. Of these, 285 appear more than 50 times, and 220 more than 100 times. For wider unitaries, it may therefore still be possible to enumerate all PQC templates that appear frequently.

The greatest complication with scaling this style of seeded synthesis to larger width unitaries is the availability (or lack thereof) of a large enough *labeled* dataset of partitioned unitaries. The process of creating a labeled dataset for this style of optimization is outlined in Section IV-C. Complications arise because the process of labeling unitaries with the PQCs that implement them is synthesis. Data augmentations that exploit knowledge of pre-labeled unitaries should therefore be explored.

C. Future Directions

Reinforcement Learning for circuit optimization has been explored in the past, notably in [5]. We believe this approach remains a promising means of optimizing quantum circuits. The authors expect that training RL systems on datasets of non-random benchmarks can further advance real world performance of this style of peephole optimization. The use of RL as a method for searching the synthesis tree could also provide further speedups. Perhaps the use of the canonical representation of unitaries formulated in Section III-A could aid in this venture.

Our presentation of a canonical representation of unitaries is also widely applicable to any domain where numerical methods are applied to unitary matrices. This canonical representation allows for the use of non-phase resilient distance metrics on unitaries, an area that is not currently widely explored.

As these unitary datasets are expensive to generate and label, it is also worth exploring possible data augmentations that preserve unitarity. We have preliminarily explored adding small amounts of random Gaussian noise to parameters to generate new unitaries that are closely related to existing unitaries (and can be implemented by the same PQC), but these experiments have not aided the recommender model's training.

As the size of unitaries used in ML tasks grow, using fully connected layers, as is done in the recommender model, will quickly become nonviable. It is unclear what neural network architecture is best suited for learning on unitary datasets. Convolutional neural networks are popular in computer vision tasks, as they are translationally equivariant and particularly adept at extracting spatially local patterns. We see no reason to believe that these inductive biases would be helpful for unitary datasets. Therefore, this domain deserves further attention.

VII. CONCLUSION

In this paper, we have demonstrated how machine learning can be applied directly onto unitary matrices to improve quantum circuit compilation and optimization. To our knowledge, we are the first to explore employing machine learning on unitary datasets for this purpose. We highlight how patterns in a diverse set of optimized real-world benchmarks emerge, and how these patterns imply learnable mappings between unitaries of interest and resource efficient circuits. We have introduced QSeed, a seeded unitary synthesis algorithm that replaces the expensive synthesis tree search process with machine learning inference. By mapping a canonical representation of unitaries to parameterized quantum circuits that may have generated them, the time needed to perform synthesis can be substantially improved with little change in the quality of solutions found. Notably, QSeed is able to synthesize a 64 qubit modular exponentiation circuit from Shor's algorithm $3.7\times$ faster than the state-of-the-art. Generalization of QSeed's recommender model, a deep neural network which proposes seed circuits, is demonstrated through the compilation of

circuits not seen during training, including an entirely unfamiliar Grover's algorithm circuit. This project illustrates how simple machine learning techniques can make substantial improvements to quantum circuit compilation tools, and ensure their applicability to a wider scope of problems.

ACKNOWLEDGEMENTS

This work was supported by the DOE under contract DE-5AC02-05CH11231 through the Office of Advanced Scientific Computing Research (ASCR) Quantum Algorithms Team and Accelerated Research in Quantum Computing programs, and by the NSF Challenge Institute for Quantum Computation (CIQC) program under award OMA-2016245.

REFERENCES

- [1] M. Amy, D. Maslov, M. Mosca, and M. Roetteler, "A meet-in-the-middle algorithm for fast synthesis of depth-optimal quantum circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2013.
- [2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.
- [3] M. G. Davis, E. Smith, A. Tudor, K. Sen, I. Siddiqi, and C. Iancu, "Heuristics for quantum compiling with a continuous gate set." [Online]. Available: <http://arxiv.org/abs/1912.02727>
- [4] Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, "Expressive power of parametrized quantum circuits," *Phys. Rev. Res.*, vol. 2, p. 033125, Jul 2020. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevResearch.2.033125>
- [5] T. Fosel, M. Niu, F. Marquardt, and L. Li, "Quantum circuit optimization with deep reinforcement learning," *null*, 2021.
- [6] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, A. Mari, and W. J. Zeng, "Digital zero noise extrapolation for quantum error mitigation," *arXiv: Quantum Physics*, 2020.
- [7] A. W. Harrow and R. A. Low, "Random quantum circuits are approximate 2-designs," *Communications in Mathematical Physics*, vol. 291, no. 1, pp. 257–302, 2009. [Online]. Available: <https://doi.org/10.1007/s00220-009-0873-6>
- [8] D. Herman, C. Googin, X. Liu, A. Galda, I. Safro, Y. Sun, M. Pistoia, and Y. Alexeev, "A Survey of Quantum Computing for Finance," Jun. 2022, arXiv:2201.02773 [quant-ph, q-fin]. [Online]. Available: <http://arxiv.org/abs/2201.02773>
- [9] L. E. Heyfron and E. T. Campbell, "An efficient quantum compiler that reduces t count," *arXiv: Quantum Physics*, 2017.
- [10] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, pp. 504–7, 2006. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/16873662/>
- [11] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, "Connecting ansatz expressibility to gradient magnitudes and barren plateaus," *PRX Quantum*, vol. 3, p. 010313, Jan 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PRXQuantum.3.010313>
- [12] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, 2016. [Online]. Available: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2015.0202>
- [13] A. Kissinger and J. van de Wetering, "Reducing the number of non-clifford gates in quantum circuits," *arXiv: Quantum Physics*, 2019.
- [14] G. Li, Y. Ding, and Y. Xie, "Tackling the qubit mapping problem for nisq-era quantum devices," in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2019, pp. 1001–1014.
- [15] A. Mandviwalla, K. Ohshiro, and B. Ji, "Implementing grover's algorithm on the ibm quantum computers," in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 2531–2537.
- [16] P. Murali, J. M. Baker, A. Javadi-Abhari, F. T. Chong, and M. Martonosi, "Noise-adaptive compiler mappings for noisy intermediate-scale quantum computers," *International Conference on Architectural Support for Programming Languages and Operating Systems*, 2019.

- [17] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," 2002.
- [18] M. Ostaszewski, L. M. Trenkwalder, W. Masarczyk, E. Scerri, and V. Dunjko, "Reinforcement learning for optimization of variational quantum circuit architectures," *Neural Information Processing Systems*, 2021.
- [19] T. Patel, E. Younis, C. Iancu, W. de Jong, and D. Tiwari, *QUEST: Systematically Approximating Quantum Circuits for Higher Output Fidelity*. New York, NY, USA: Association for Computing Machinery, 2022, p. 514–528. [Online]. Available: <https://doi.org/10.1145/3503222.3507739>
- [20] A. K. Prasad, V. Shende, I. L. Markov, J. P. Hayes, and K. N. Patel, "Data structures and algorithms for simplifying reversible circuits," *ACM Journal on Emerging Technologies in Computing Systems*, 2006.
- [21] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, Aug 2018. [Online]. Available: <http://dx.doi.org/10.22331/q-2018-08-06-79>
- [22] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," 2023.
- [23] N. Quetschlich, L. Burgholzer, and R. Wille, "Compiler optimization for quantum computing using reinforcement learning," *arXiv.org*, 2022.
- [24] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*. Cambridge, MA, USA: MIT Press, 1986, p. 318–362.
- [25] V. V. Shende, S. S. Bullock, and I. L. Markov, "Synthesis of quantum-logic circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 6, pp. 1000–1010, 2006.
- [26] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," *Advanced Quantum Technologies*, vol. 2, no. 12, p. 1900070, 2019. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.201900070>
- [27] E. Smith, M. G. Davis, J. M. Larson, E. Younis, C. Iancu, and W. Lavrijsen, "LEAP: Scaling Numerical Optimization Based Synthesis Using an Incremental Approach," *arXiv preprint arXiv:2106.11246*, 2021.
- [28] S. S. Tannu and M. K. Qureshi, "A case for variability-aware policies for nisq-era quantum computers," *arXiv: Quantum Physics*, 2018.
- [29] R. R. Tucci, "An introduction to cartan's kak decomposition for qc programmers," *arXiv preprint quant-ph/0507171*, 2005.
- [30] J. J. Wallman and J. Emerson, "Noise tailoring for scalable quantum computation via randomized compiling," *Physical Review A*, 2016.
- [31] M. Weiden, J. Kalloor, J. Kubiawicz, E. Younis, and C. Iancu, "Wide quantum circuit optimization with topology aware synthesis," in *2022 IEEE/ACM Third International Workshop on Quantum Computing Software (QCS)*, 2022, pp. 1–11.
- [32] X.-C. Wu, M. G. Davis, F. T. Chong, and C. Iancu, "Qgo: Scalable quantum circuit optimization using automated synthesis," 2021.
- [33] E. Younis and C. Iancu, "Quantum circuit optimization and transpilation via parameterized circuit instantiation," in *IEEE International Conference on Quantum Computing and Engineering, QCE 2022, Broomfield, CO, USA, September 18-23, 2022*. IEEE, 2022, pp. 465–475. [Online]. Available: <https://doi.org/10.1109/QCE53715.2022.00068>
- [34] E. Younis, C. C. Iancu, W. Lavrijsen, M. Davis, and E. Smith, "Berkeley quantum synthesis toolkit (bqskit) v1," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Tech. Rep., 2021. [Online]. Available: <https://bqskit.lbl.gov>
- [35] E. Younis, K. Sen, K. Yelick, and C. Iancu, "QFAST: Quantum synthesis using a hierarchical continuous circuit space," *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2020. [Online]. Available: <http://arxiv.org/abs/2003.04462>