Pauli Error Propagation-Based Gate Rescheduling for Quantum Circuit Error Mitigation

Vedika Saravanan, Samah Mohamed Saeed
City College of New York, City University of New York, New York, USA
vsarava000@citymail.cuny.edu, ssaeed@ccny.cuny.edu

Abstract—Noisy Intermediate-Scale Quantum (NISQ) algorithms, which run on noisy quantum computers should be carefully designed to boost the output state fidelity. While several compilation approaches have been proposed to minimize circuit errors, they often omit the detailed circuit structure information that does not affect the circuit depth or the gate count. In the presence of spatial variation in the error rate of the quantum gates, adjusting the circuit structure can play a major role in mitigating errors. In this paper, we exploit the freedom of gate reordering based on the commutation rules to show the impact of gate error propagation paths on the output state fidelity of the quantum circuit, propose advanced predictive techniques to project the success rate of the circuit, and develop a new compilation phase post-quantum circuit mapping to improve its reliability. Our proposed approaches have been validated using a variety of quantum circuits with different success metrics, which are executed on IBM quantum computers. Our results show that rescheduling quantum gates based on their error propagation paths can significantly improve the fidelity of the quantum circuit in the presence of variable gate error rates.

Index Terms—Noisy Intermediate-Scale Quantum (NISQ) computer, quantum circuit, Pauli errors, error propagation, gate rescheduling, commutation rules, quantum circuit mapping, reliability.

I. INTRODUCTION

Noisy Intermediate-Scale Quantum (NISQ) computers, which have tens to hundreds of quantum bits (qubits) are anticipated to speed up the computational power of classical computers for certain classes of problems. However, the noisy nature of their qubits and operations degrades the reliability of NISQ systems. Thus, errors should be mitigated to improve the fidelity of NISQ algorithms, which necessitates error characterization and circuit design approaches to minimize the impact of errors on the quantum circuit output. Randomized benchmarking is typically used to characterize different sources of noise in the quantum hardware including gate, measurement, and decoherence errors [1]. The computed error rates are used by noise-aware quantum compilers to generate physical quantum circuits with high fidelity.

While quantum noise is very complex, simplified error models based on error rates collected during the calibration process are typically used for quantum circuit compilation [2]. Depending on the permissible access to the quantum computer, more advanced error characterization can be unavailable. A predictive analysis capable of delivering a projection with a reasonable accuracy about the circuit error rates, despite the lack of complete and accurate error model is also required. A predictive technique, which is computationally fast and applicable to many quantum computing technologies, and takes into account the circuit structure can enhance the compilation process [3].

In this paper, we propose error propagation-based predictive techniques to project the quantum circuit success rate primarily based on gate errors. We take into account the propagation path of different gate errors in the circuit. While the existing estimated success probability typically used by noise-aware quantum compilers to generate physical quantum circuit incorporates the gate error rates into its model, it ignores the order of the gates that don’t affect the circuit depth or the gate count. Our proposed predictive techniques can differentiate between different physical implementations of the same quantum circuit, which use the same physical qubits, share the same set of quantum gates, and have the same depth, but vary in the order of quantum gates in the circuit and their corresponding error rates. We experimentally show the advantage of considering the order of the physical quantum circuit gates in the presence of variable gate errors. Accordingly, we also propose gate rescheduling algorithms at different quantum circuit abstractions to maximize the fidelity of the quantum circuit. Our proposed approaches can be integrated with other gate rescheduling algorithms that reduce the gate count and the circuit depth. To the best of our knowledge, no such analysis has been proposed to incorporate the impact of the error propagation paths to the quantum gate scheduling procedure. The main contributions of our paper are as follows:

- We propose a Weighted Estimated Success Probability (WESP) metric based on the quantum circuit structure.
- We propose new quantum circuit gate rescheduling algorithms at different design levels including complex and elementary gate-level based on our proposed WESP metric to boost the fidelity of the quantum circuit output.
- We experimentally validate the effectiveness of our proposed metrics and gate rescheduling algorithms using a variety of quantum circuits executed on different IBM quantum computers.

The remainder of paper is organized as follows. Section II provides a background on the quantum circuit compilation, quantum hardware errors, an application of NISQ computers, and different success criteria for quantum circuit evaluation. Section III discusses related works on quantum circuit mapping and gate reordering approaches. Section IV shows the
implication of the gate error propagation path on the output state fidelity of the quantum circuit. Section V provides our proposed reliability metric and rescheduling algorithms. Section VI validates the effectiveness of our approaches through several experiments. We conclude the paper in Section VII.

II. BACKGROUND

A. Quantum circuit compilation

Quantum algorithms are described using quantum circuits, which are executed on the quantum computer. A quantum circuit comprises of quantum gates, which change the state of the qubits. Single-qubit gates operate on a single qubit such as the Hadamard (H) gate, which creates a superposition state, and the $R_z$ gate, which rotates the qubit around the z-axis. A multi-qubit gate such as Controlled Not (CNOT) entangles two qubits [4]. To enable a quantum circuit execution on the targeted NISQ computer, the circuit has to go through several compilation steps. Complex gates should be constructed using elementary gates supported by the NISQ architecture. To reduce the depth and the gate count of the circuit, several gate-level optimization techniques are applied including template matching [6] and gate reordering [7] based techniques. In the former one, a cascade of quantum gates is substituted with a sub-circuit with a lower gate count, while in the later one, consecutive gates cancel each other based on the commutation rules of the quantum gates [4], which are defined as follows:

Definition 1: Let $U_1$ and $U_2$ be two unitary matrices. $U_1$ and $U_2$ are said to be commutative if $U_1U_2 = U_2U_1$ for any input state. Commutation rules of different quantum gates are shown in Figure 1.

Next, the physical qubits of the quantum circuit are allocated and their gates are scheduled to meet the constraints of the quantum architecture. This process is referred to as quantum circuit mapping. A main challenge in the mapping process is the restricted qubits connectivity in superconducting quantum architectures. Two-qubit gates can be applied to certain pairs of physical qubits described using the coupling graph of the quantum architecture, in which nodes represent qubits and edges show the connectivity between different qubits. To enable arbitrary multi-qubit gates in the presence of the coupling constraint, SWAP operations consisting of three CNOT gates are used for gate scheduling, resulting in an excessive gate count and a large circuit depth. Efficient mapping approaches are used to generate physical quantum circuits with as a minimum gate count as possible to be executed on the quantum computer. These approaches often represent the quantum circuit as a directed acyclic graph, referred to as a gate dependency graph, where each node represents a gate and each edge represents a direct dependency between two gates. The dependency graph is divided into levels, where each level consists of quantum gates applied simultaneously in the same circuit layer. All the nodes reachable from a given node $i$ in the graph are considered as the gates which are dependent on the $g_i$ gate ($g_i$ reachable gates).

Example 1: Figure 2 shows an example of a quantum circuit and its corresponding gate dependency graph in which the gates in each circuit layer are mapped to the graph nodes applied at the same level. For each quantum gate, the list reachable gates is extracted from the gate dependency graph. For example, there are 3 reachable gates from $g_3$ gate, which are $g_5$, $g_6$, $g_7$, and $g_8$ gates.

B. Reliability of quantum devices

Different sources of noise contribute to the noise complexity of NISQ systems. The elevated noise levels endanger the reliability of quantum circuits. Gate errors are a major source of quantum computing noise. They are modeled as Pauli gates that operate on a single qubit. They are represented as $X$ (a $\pi$ rotation around the x-axis), $Y$ (a $\pi$ rotation around the y-axis), $Z$ (a $\pi$ rotation around the z-axis), and $I$ (identity matrix) single-qubit gates [8]. The propagation of some of the Pauli errors through the CNOT gate is shown in Figure 3. The measurement/read-out error can occur while measuring the output state of the quantum circuit. A qubit can lose its state after a period of time, resulting in decoherence errors.
Quantum gates that act on different qubits simultaneously can cause crosstalk errors.

As NISQ computers do not support quantum error correction, they often rely on noise-aware quantum compilers to mitigate errors. The quantum compiler should maximize the success probability of the physical quantum circuit. An Estimated Success Probability (ESP) can be used by a noise-aware quantum compiler to select the physical qubits of the quantum circuit and schedule the quantum gates. Given the gate ($\gamma_i$) and measurement ($\epsilon_m$) errors for all the quantum circuit gates, and its qubits, ESP is computed as $\prod_{i=0}^{Q-1} (1 - \gamma_i) \times \prod_{i=0}^{Q-1} (1 - \epsilon_m)$ [2]. Randomized Benchmarking sequences are executed on the quantum hardware to compute single- and two-qubit gate errors based on the average sequence errors. The qubit measurement/readout error is computed as the average measurement error 0 and 1, in which the qubit is initially assigned to state 1 and 0, respectively.

C. Quantum approximation optimization algorithm

Quantum approximation optimization algorithm (QAOA) is a hybrid quantum-classical algorithm for solving optimization problems. It relies on not only the output of the quantum circuit but also the classical optimizer that updates the circuit parameters to improve the solution of the optimization problem [9]. The objective of the algorithm is to minimize/maximize a cost function ($C(X)$) of a given problem described using the following Hamiltonian: $H = \sum_{x \in \{0,1\}^n} C(x) |x\rangle \langle x|$. A QAOA quantum circuit contains two main components, namely, phase separation and mixing operations, which are applied $p$ times repeatedly with $(\gamma_1, \beta_1), \ldots, (\gamma_p, \beta_p)$ parameters. The phase separation can be represented using ZZ gates, while mixing operations are represented using single-qubit gate rotation around the x-axis ($R_x$). An example of a problem to be solved using the QAOA algorithm is the Maximum Cut (Max-Cut) problem. For a graph with $n$ vertices, the solution of the Max-Cut problem divides the graph into two subsets such that the total weight of the edges between those subsets is as maximum as possible. The graph vertices are described as a string of $n$ qubits in which each qubit can be measured as 0 or 1. The QAOA circuit output provides a candidate solution of the Max-Cut problem. Using an iterative procedure, we can find the $n$-qubit string that maximizes the total weight of the edges between the two subsets, referred to as the cost function.

D. Quantum circuit success metrics

Different success criteria/metrics have been proposed to show the impact of the quantum compilation approaches on the quantum circuit output. For a single-output quantum circuit, the Probability of Successful Trails (PST) measures the probability of the correct output of the quantum circuit as $\frac{\text{Number of Successful Trials}}{\text{Total Number of Trials}}$ [10]. For hybrid quantum algorithms, application-specific success criteria are used. For example for QAOA, which maximizes or minimizes a cost function, the Approximation Ratio (AR) of the cost function quantifies the success of the quantum circuit [11], [12]. It is defined as $\frac{\text{Mean of Cost Function Over all Sampled Output}}{\text{Maximum Cost Function Value}}$. Given AR, we can compute the Approximation Ratio Gap (ARG) to quantify how close is the output of the actual execution of QAOA circuit to the simulation result [11]. It is computed as $\frac{\text{AR of Simulation} - \text{AR of Execution}}{\text{AR of Execution}} \times 100$. The smaller the gap, the more reliable the QAOA circuit.

III. RELATED WORK

A. Quantum circuit mapping

Since quantum circuit mapping is an NP-Complete problem [13], traversing the search space is speculative for quantum architectures with a large number of qubits. Various heuristics-based techniques have been proposed to address this problem [14]. Some of these techniques target minimizing the number of circuit SWAP operations [15]–[19], while others focus on minimizing the circuit error rates [10], [20]. For example, quantum circuit mapping can be achieved by using BRIDGE and SWAP operations as transformation rules inserted based on a dynamic programming or a heuristic-based look-ahead schemes [15]. The BRIDGE gates are composed of a sequence of four CNOT gates. SAT solver minimizes the number of SWAP operations in the physical quantum circuit at the cost of extensive computation overhead (e.g. [16], [17]). The proposed technique in [16] divides the quantum circuit into sub-circuits that require no SWAP operations within each sub-circuit. Next, the SAT solver is used to construct the quantum circuit by inserting a minimum number of SWAP operations between the sub-circuits that share qubits. In [17], a logical quantum circuit is mapped to a physical quantum circuit using a SAT solver too. Constraints are added to the SAT solver based on the coupling graph of the quantum architecture to reduce its complexity at the expense of the gate count. An A* algorithm explores the search space in the coupling graph to find the shortest path between qubits for each two-qubit gate, and thus, reduce the number of SWAP operations in the quantum circuit [18]. To address the difference in the error rates of various qubits, a mapping technique has been proposed based on Dijkstra algorithm, which schedules quantum gates and allocates physical qubits with low error rates [10]. The success probability of the physical quantum circuit can be further improved by searching for isomorphic sub-graphs in the quantum architecture’s coupling constraint with the highest ESP [20]. Other mapping approaches address decoherence, correlated, and unexpected errors (e.g [21]–[23]). To mitigate crosstalk errors, they should be characterized first, followed by an efficient mapping approach that selectively serializes circuit gates, while satisfying the required coherence time [24].
quantum circuit using the previously proposed compilation to maintain the gate count and the circuit depth of the optimized compilation/optimization layers to further improve the output. In this work, we propose additional compilation layers which restrict the search space of our gate rescheduling algorithms to additional gates due to decoherence errors are injected into the quantum circuit to analyze their impact on the circuit output. According to [29], faults in the form of gate errors can be reordered based on commutation rules, and then logical instructions to control pulses [28]. In this approach, a two-step process is adopted to minimize the circuit depth using a breadth-first search algorithm, which finds a possible two-layer interchanges in each iteration until the desired depth is achieved. A layout synthesis algorithm has been proposed to enhance the performance of QAOA circuits by reordering the complex gates [11]. Due to the commutation relation, most of the dependencies are ignored. Instead, a two-level search process is adopted to minimize the circuit depth using a breadth-first search algorithm, which finds a possible two-layer interchanges in each iteration until the desired depth is achieved. A layout synthesis algorithm has been proposed to minimize the circuit depth by reordering complex gates of QAOA circuits using a Satisfiability Modulo Theories (SMT) solver [27]. An optimized pulse level compilation approach has been proposed to address the inefficiency of the standard gate compilation process that directly translates the logical instructions to control pulses [28]. In this approach, the gates are reordered based on commutation rules, and then a small set of gates are aggregated to a larger operation by finding the optimal control pulse based on the gradient descent method. Gate reordering has also been applied to minimize the impact of decoherence errors [29]. Faults in the form of additional gates due to decoherence errors are injected into the quantum circuit to analyze their impact on the circuit output. Accordingly, gates are reordered at the circuit locations, in which the decoherence errors significantly affect the circuit output. In this work, we propose additional compilation layers applied to the physical quantum circuit post-mapping. We restrict the search space of our gate rescheduling algorithms to maintain the gate count and the circuit depth of the optimized quantum circuit using the previously proposed compilation approaches. Thus, our work can be incorporated with the other compilation/optimization layers to further improve the output state fidelity of the quantum circuit.

IV. IMPLICATION OF THE GATE ERROR PROPAGATION PATH ON THE OUTPUT STATE FIDELITY

Stochastic Pauli noise is a primary factor to estimate the circuit error rates [2]. While tracking the precise impact of the error propagation on the circuit can be computationally expensive in the presence of universal gates [3], two-qubit gates enable error propagation across different qubits. Thus, the location of the quantum gates especially the two-qubit gates can affect the circuit error rates. For the sake of simplicity, we assume that the two-qubit gate (CNOT) error impacts the control and the target qubits, and thus, propagates to the gates applied next to the control and the target qubits.

Example 2: To illustrate the gate error propagation through the quantum circuit, let’s consider the example provided in Figure 4. Two equivalent sub-circuits using the commutation rules of quantum gates and their corresponding gate dependency graphs are provided. We use a dotted line to show the output qubits affected by the gates in the dependency graph. Every gate in the sub-circuit can produce Pauli errors that propagate to the qubit output through the gate dependency graph [3]. In Figure 4(a), $g_0$, $q_1$, and $q_2$ output qubits are affected by $\{g_0\}$, $\{g_0, g_1, g_2, g_3\}$, and $\{g_0, g_1, g_2\}$ gate errors, respectively. On the other hand, in Figure 4(b), $q_0$, $q_1$, and $q_2$ output qubits are affected by $\{g_0, g_1, g_2, g_3\}$, $\{g_0, g_1, g_2, g_3\}$, and $\{g_1, g_2\}$ gate errors, respectively. Thus, while the two quantum sub-circuits in Figure 4 are equivalent, their output qubits $g_0$ and $q_2$ are susceptible to different sets of gate errors.

Due to the spatial variation in error rates of the quantum hardware, gates applied to different qubits result in different error rates. Although all gate errors affect the fidelity of the quantum circuit, the two-qubit gate errors has lower fidelity than single-qubit gate errors. Thus, reordering quantum gates while maintaining the circuit functionality changes not only the error propagation but also the gate error impact on the output state of the circuit. In the previous example, if $g_0$ gate has the highest error rate in the circuit, which significantly deviates from other gate errors in the circuit, applying $g_0$ gate at the first layer of the sub-circuit as in Figure 4(a) will spread $g_0$ error to all the output qubits unlike the equivalent sub-circuit in Figure 4(b) in which $g_0$ error affects only $q_0$ and $q_1$ output qubits.

Example 3: Figure 5 shows the impact of gate reordering on the quantum circuit output fidelity. A 4-qubit Bernstein–Vazirani (BV) circuit is mapped to satisfy the coupling constraint of IBM Q Santiago architecture. Figure 5(a) represents the coupling graph and the error rates of IBM Q Santiago architecture, in which the values in the upper and lower part of each node represent the read-out and single-qubit errors, respectively, while the edge label is the two-qubit gate error applied to the corresponding pair of qubits. All error rates are multiplied by $10^{-5}$. Two physical implementations of the BV quantum circuit, namely M1 and M2, are provided in Figure 5(b) and 5(c), respectively. The difference between the two physical implementations of the circuit is the location of the first two CNOT gates of the quantum circuit, which can be reordered based on the commutation rules. The CNOT gate highlighted in red color has the highest two-qubit gate

![Fig. 4: (a) A quantum sub-circuit and its gate dependency graph, and (b) its corresponding equivalent sub-circuit and its gate dependency graph.](image-url)
error rate. According to the order of the quantum gates, the error propagation path also changes as shown in Figure 5(b) and 5(c). We demonstrate the impact of the gate reordering on the quantum circuit output by executing both M1 and M2 quantum circuits on IBM Q Santiago architecture, and reporting their PST as shown in Figure 5(d). To eliminate the impact of gate reordering on decoherence errors, we insert barriers before and after applying all the two-qubit gates in the two circuit implementations in addition to a barrier after the second CNOT gate of M2 quantum circuit in Figure 5(c), which ensure that the qubits in the two implementations are used for the same period of time. The PST values are 0.784 and 0.861 for the BV quantum circuit generated using the M1 and M2 mapping, respectively, and executed for 8192 shots/trials. Our observation indicates that reordering gates with higher error rates to earlier circuit layers will result in error propagation to larger number of circuit gates, and thus, degrade the output state fidelity. While PST is different for both mapping, the ESP of both M1 and M2 mapping is 0.7757, which implies the need to consider the gate error propagation path for estimating the success rate of the circuit, and thus, refining the quantum circuit mapping policies.

V. ERROR PROPAGATION-BASED GATE RESCHEDULING

A. Weighted estimated success probability

ESP enables quick and easy estimation of the circuit success probability. Yet, it overlooks the circuit structure since different orders of the quantum gates yield the same ESP. To account for the penalty of rescheduling the same set of physical quantum gates with variable error rates on the output state fidelity of the quantum circuit, we modify the ESP metric and propose a new metric referred to as Weighted Estimated Success Probability (WESP) defined as

\[
WESP = \prod_{i=0}^{G-1} (1 - (\lambda_i + e_{g_i})) \times \prod_{i=0}^{Q-1} (1 - e_{m_i})
\]

\(\lambda_i\) tunes the error rate of the corresponding \(g_i\) gate to account for the gate position as

\[
\lambda_i = w_i(e_{g_i} - \min(E_G))
\]

where \(w_i\) and \(\min(E_G)\) are the weight of the \(i^{th}\) gate and the minimum gate error rate in set of gate errors \(E_G\) of the quantum circuit, respectively. Instead of multiplying the weight of each gate by its error rate, which can significantly reduce the gate error of very noisy gates applied in later layers of the circuit, we multiply the weight with the difference between the current gate error and the minimum gate error in the circuit to account for the variation in the gate errors. The resulting number is added to the gate error. To compute \(w_i\), the gate dependency graph of the circuit is constructed first. For each gate \(g_i\) (node) in the circuit (graph), we identify the number of all subsequent dependent gates (reachable gates) that are applied to the same qubit or other qubits that are directly or indirectly connected to the current gate through two-qubit gates \(S_i\). For a quantum circuit with \(G\) gates, the weight of the \(i^{th}\) gate \(w_i\) is computed as \(w_i = \frac{S_i}{G}\). We exclude the \(R_z\) gate since the error rate is zero. Thus, the term \(\lambda\) is used to approximate the impact of noisy quantum
gates on the output state fidelity of the quantum circuit given the number of subsequent dependent (reachable) gates. To simplify our analysis, we omit how Pauli errors propagate through different quantum gates in the circuit. Instead, we assume that for each gate, Pauli errors will propagate to its corresponding qubits.

**Example 4:** Figure 6(a) and (b) show the physical quantum circuits implemented using elementary gates supported by IBM quantum computers for M1 and M2 quantum circuits in Figure 5 respectively. The $S_i$ values for each single- and two-qubit gates are shown in red color. For example, in the physical circuit provided in Figure 6(a), the $\sqrt{X}$ gate applied to qubit $q_4$ in layer 2 is connected to the next $\sqrt{X}$ gate applied to the same qubit at a later layer in the gate dependency graph. Hence, the value of $S_i$ for the first $\sqrt{X}$ gate is 1. Similarly, the value of $S_i$ for all the other gates except $R_y$ gates are computed. The number of erroneous gates ($G$) is 16. Based on the device error rates as shown in Figure 5(a), the value of WESP for M1 mapping is 0.7748. On the other hand, WESP for M2 mapping is 0.7764, which indicates that M2 circuit provides a higher PST. This observation is aligned with the PST value of M1 and M2 circuits provided in Figure 5(d).

**B. Elementary gate rescheduling based on WESP**

We exploit the gate commutation rules to reduce the impact of Pauli errors while maintaining the depth of the circuit. We propose a gate rescheduling algorithm post-quantum circuit mapping guided by the proposed WESP metric as a new optimization layer to reduce the circuit gate errors. Since the gates of the physical quantum circuit satisfy the coupling constraint of the quantum hardware, no additional SWAP operation is required. We develop a greedy approach for gate rescheduling based on the circuit dependency graph ($D(V, E)$) as described in Algorithm 1. The objective of our rescheduling algorithm is to minimize WESP while maintaining the circuit depth ($\text{Depth}_C$) to avoid any additional decoherence errors. An important condition is added to ensure that the updated gate scheduling does not increase the depth of the circuit ($\text{No}_{\text{Inc}}\cdot \text{Depth}_C$). For $g_i$ gate operating on $q_z$ and $q_y$ qubits at layer $t$ and $g_j$ gate operating on $q_z$ and $q_y$ qubits at layer $t+k$, $g_i$ and $g_j$ can be reordered if they are commuting and $q_y$ and $q_z$ are idle at layers $t+k$ and $t$, respectively. Our evaluation is repeated for every pair of immediate dependent gates in the gate dependency graph. For a quantum circuit with $D(V, E)$ gate dependency graph, in which $V$ (nodes) represents the set of quantum circuit gates and $|E|$ (edges) shows the dependency between different gates, the time complexity of calculating WESP is $O(|V| + |E|)$. Thus, the time complexity of our greedy heuristic, which is provided in Algorithm 1 is $O(|V|^2 \cdot (|V| + |E|))$. Our proposed gate rescheduling approach is computationally fast compared to the exhaustive approach that requires checking the entire search space, which can be very large for quantum circuits with large number of commuting gates.

**Example 5:** To motivate the need and the effectiveness of our greedy approach, we exhaustively generate all possible ways of gate scheduling for our BV circuit provided in Figure 7. Two quantum sub-circuits with different ZZ gate scheduling.

Figure 5 and 6 We pick the circuit with the highest WESP and compare it with the generated circuit using our greedy approach in terms of WESP and PST. Specifically, we label each CNOT gate in Figure 6 to list all possible combinations of gate scheduling. The total number of all possible ways of gate scheduling, which maintain the depth of the circuit is 14. Among all the gate scheduling combinations, B-C-D-E-F-A has the highest WESP of 0.7768, while the resulting circuit based on our greedy algorithm has a CNOT combination of B-A-C-D-E-F with a WESP of 0.7764. The PST value of the quantum circuit based on the exhaustive and our greedy approach are 0.863 and 0.861, respectively. Accordingly, the gap between the exhaustive and the greedy solutions is very small.

**Algorithm 1:** Gate rescheduling

```
Input: C = The physical quantum circuit
Output: C' = The updated physical quantum circuit
1 Construct the $D$ graph of $C$;
2 Initialize $C'$ = $C$;
3 Initialize $L(i) =$ gates applied at layer $i$ | $0 < i < \text{Depth}_C$;
4 for each each layer $i$ do
5   for each each gate $g_j$ in $L(i)$ do
6     for each immediate dependent gate $g_k$ do
7       if $g_k$ and $g_j$ are commutative then
8         if $\text{No}_{\text{Inc}}\cdot \text{Depth}(D, g_j, g_k)$ then
9           Compute WESP of $C'$ under rescheduled gates;
10       end
11     end
12   end
13 Select rescheduling of $g_j$ with max. WESP;
14 Update $C'$;
15 end
16 Return $C'$;
```

**C. Complex gate rescheduling based on WESP for QAOA**

QAOA quantum circuits consist of single-qubit rotations and two-qubit phase gates implemented using ZZ complex gate. Since ZZ gates are commutative [30], [31], we exploit the freedom of the ZZ gate placement at higher quantum circuit design level to further reduce the quantum circuit errors. ZZ gates are reordered without violating the circuit depth and the gate count.

**Example 6:** An example of ZZ gate reordering is provided in Figure 7 in which both quantum sub-circuits share the same gate count and circuit depth but vary in the ZZ gate scheduling.
We propose a look-ahead approach that approximates the error rate of each complex gate in the circuit. The approximation entails the generation of physical quantum circuit first, using an efficient quantum mapping approach that allocates physical qubits with minimum error rates and minimizes the number of SWAP operations. We scan the physical quantum circuit next to identify the corresponding two-qubit gates used for building each ZZ complex gate. Given the decomposed sub-circuit of each complex gate, we compute the error rate of each complex gate.

We adjust the WESP metric to operate on the intermediate representation of the quantum circuit prior to complex gate decomposition. In the updated WESP, $G$ is the number of ZZ gates, $e_g$, is the ZZ gate error rate computed as the product of the error rates of its elementary gates, $w_i$ is the ratio of dependent (reachable) ZZ gates, and $\min(E_G)$ is the minimum error rate of the ZZ complex gate in the circuit. The ZZ gates are reordered based on the updated WESP using a similar rescheduling algorithm as the one provided in Algorithm 1 in which $C$ is updated to be the intermediate representation of the quantum circuit and every pair of quantum gates $(g_i, g_j)$ corresponds to a pair of ZZ gates in the quantum circuit. The look-ahead approach requires scanning the entire physical circuit including all its gates. The time complexity of the look-ahead approach is linear in the number of gates \(O(|V|)\). The time complexity of our ZZ gate rescheduling algorithm is \(O(|V'|^2 \cdot (|V'| + |E'|))\) for \(D(V', E')\) complex gate dependency graph, in which \(V'\) represents the set of ZZ complex gates and \(E'\) is the set of edges that show the dependency between different ZZ gates.

We emphasize that our rescheduling approaches, which are applied post-mapping can be integrated with other quantum compilation approaches that exploit the gate commutation rules to reduce the circuit depth and the gate count since our algorithms do not alter the circuit depth nor the gate count. Furthermore, our proposed approaches can be easily applied to quantum circuits executed on different quantum hardware in the presence of significantly variable hardware gate errors.

### VI. Experimental Evaluation

#### A. Experimental setup

We show the effectiveness of our proposed rescheduling algorithms based on WESP in improving the output state fidelity of different quantum circuits executed on NISQ computers. We conduct two experiments. In the first experiment, we use PST to evaluate our proposed elementary gate rescheduling approach applied to different quantum circuits, while in the second experiment, we show the impact of our proposed complex ZZ gate rescheduling on the AR of QAOA quantum circuits.

We use a total of 12 benchmark circuits with different number of qubits and gate count. The Bernstein-Vazirani (BV) algorithm identifies a hidden string, which is encoded in the circuit \[32\]. The discrete fourier transform is represented using Quantum Fourier Transform (QFT) circuit \[5\]. A shift operation is carried out for a given input Boolean function using Hidden Shift (HS) algorithm \[33\]. The Grover Search (Grover) algorithm provides a solution \(x\) of a function \(f(x)\) equal to 1. Given a unitary operator, the eigenvalue of the eigenvector is estimated using Quantum Phase Estimation (QPE) algorithm \[5\]. Quantum Approximate Optimization Algorithm (QAOA) solves optimization problems \[9\]. Other reversible quantum circuits such as Adder, Toffoli gate, and Decoder are also used. We obtained QFT and QPE circuits from Qiskit \[34\], HS and QAOA circuits from Cirq \[35\], adder from \[36\], decoder from RevLib \[37\], and BV, Toffoli, and Grover Search by manual construction. The properties of all the benchmark circuits are provided in Table I prior to the circuit mapping, which includes the number of qubits (# Qubits), single-qubit gates (U), CNOT gates (CNOT), and measurement operations (M), the circuit depth (Depth), and the expected output. We use QAOA to solve the Max-Cut problem of three different graphs, which consist of 15 nodes and different number of edges of random weight per node. The QAOA quantum circuits are QAOA_1, QAOA_2, and QAOA_3, which target graphs of 15 nodes with 5, 6, and 7 edges per node, respectively. To construct and test the QAOA circuits, we set \(p\) to 1 with default \(\gamma\) and \(\beta\) parameters provided by Cirq \[35\]. Unlike the other quantum circuits listed above, we consider the entire output distribution of QAOA quantum circuits to evaluate their cost function.

We run our circuits on a variety of quantum architectures with different number of qubits and Quantum Volume (QV), as shown in Table II. The QV determines the largest size of random quantum circuits in terms of the number of qubits and the circuit depth, which can be successfully executed on the quantum computer with high fidelity \[38\]. Each execution of the quantum circuit consists of 8192 shots/trials.

### Table I: Properties of quantum circuit benchmarks.

| Benchmark | # Qubits | U | CNOT | M | Depth | Expected Output |
|-----------|-----------|---|------|---|-------|-----------------|
| BV_3      | 4         | 9 | 2    | 3 | 6     | 110             |
| Adder     | 3         | 30| 17   | 2 | 28    | 0100            |
| Grover    | 3         | 24| 7    | 2 | 22    | 10              |
| Toffoli   | 3         | 11| 6    | 3 | 12    | 111             |
| QFT_3     | 3         | 10| 6    | 3 | 14    | 0000            |
| QFT_4     | 4         | 26| 18   | 4 | 28    | 0000            |
| Decoder   | 5         | 28| 21   | 4 | 29    | 00100           |
| QPE_3     | 3         | 19| 7    | 2 | 17    | 01              |
| QPE_4     | 4         | 24| 14   | 3 | 31    | 001             |
| QAOA_1    | 15        | 437| 111  | 15| 176   | NA              |
| QAOA_2    | 15        | 525| 135  | 15| 214   | NA              |
| QAOA_3    | 15        | 602| 156  | 15| 224   | NA              |

### Table II: Properties of different IBM quantum computers.

| Quantum Computer | # Qubits | QV |
|------------------|----------|----|
| IBM Q16 Melbourne| 15       | 8  |
| IBM Q Casablanca | 7        | 32 |
| IBM Q Santiago   | 5        | 32 |
| IBM Q Rome       | 5        | 32 |
We execute 150, 153, 121, and 117 pairs of base and reordered circuits on IBM Q Melbourne, IBM Q Santiago, IBM Q Casablanca, and IBM Q Rome, respectively, based on the availability of these devices. The results of all these circuits show that the reordered quantum circuits according to WESP provided in Section V yield a better solution than the base circuits generated by Qiskit. We select a subset of these circuits to show the effectiveness of our approach.

**B. Elementary gate rescheduling results**

In the first experiment, we study the impact of our proposed elementary gate rescheduling on the PST of different benchmark circuits. For each quantum circuit to be executed on a given quantum computer, we first generate the corresponding physical quantum circuit using the highest optimization level of Qiskit based on the calibration data of the quantum hardware. We refer to this approach as a Base. Next, we apply our proposed elementary gate rescheduling algorithm based on WESP to the physical quantum circuit to generate a new optimized physical quantum circuit. We refer to this process as Prop. I. We run the two physical quantum circuits on the quantum computer and evaluate their PST.

Figure 8 provides a comparison of the PST of quantum circuits generated using Base and Prop. I and executed on different quantum computers. We also report the WESP of quantum circuits generated using Base and Prop. I, the number of gate reordering using Prop. I (R), and the depth of each physical quantum circuit (D) executed on different quantum computers in Table III. We emphasize that the depth of the physical quantum circuit prior to and post applying Prop. I is always the same to avoid additional decoherence errors.

Figure 8 shows that our proposed rescheduling algorithm improves the PST, and thus, the fidelity of the quantum circuit output. We observe significant improvement of PST for some quantum circuits more than others such as Adder, Toffoli, and Grover Search quantum circuits for most of the quantum computers, which implies that the circuit structure can impact the effectiveness of our approach. We also observe that our proposed approach can also perform well for quantum computers with a larger quantum volume, and thus, a better fidelity such as IBM Q Santiago, IBM Q Casablanca, and IBM Q Rome. Furthermore, as our approach exploits the variation in gate error rates, which can result in different error rates propagating throughout different circuit paths, it is expected to perform better in the presence of higher gate error standard deviation of the quantum hardware. This is further illustrated...
TABLE IV: Run time of Prop. I algorithm applied to different quantum circuits mapped to different IBM Q computers

| Benchmark | Execution Time (msec) |
|-----------|-----------------------|
| Base      | Melbourne             |
| BV_3      | 0.54 0.17 0.17 0.23   |
| Adder     | 2.17 1.37 1.37 1.52   |
| Grover    | 1.14 1.27 1.52 1.34   |
| Toffoli   | 0.20 0.53 0.85 0.90   |
| QFT_1     | 1.12 1.38 1.49 1.26   |
| QFT_4     | 1.47 2.59 2.92 1.57   |
| Decoder   | 2.52 3.27 3.11 3.28   |
| QPE_3     | 1.57 1.51 1.93 1.72   |
| QPE_4     | 2.51 1.86 2.32 1.89   |
| Casablanca| 1.86 1.66 1.47 1.86   |
| Santiago  | 1.25 1.37 1.89 2.57   |
| Rome      | 2.49 1.37 1.25 1.38   |
| WESP (complex) | 0.90   |
| WESP (elementary) | 0.20   |

in Figure 8 for which the standard deviation of the two-qubit gate error of IBM Q16 Melbourne, IBM Q Casablanca, IBM Q Santiago, and IBM Q Rome on average are 1.31E-02, 1.86E-02, 1.33E-02, 1.20E-02, respectively, which are considered relatively high.

As shown in [29], the output state fidelity of the quantum circuit can be improved by rescheduling quantum gates as late as possible in the circuit to reduce the usage time of the qubit, and thus, reduce the qubit decoherence errors. To ensure that the improvement in PST in Figure 8 is not due to a reduction in the qubit lifetime itself, we identify the first layer in which each physical qubit is being used in each pair of quantum circuits generated using the Base and the Prop. I approaches. For non-measured qubits we also identify the last layer being used in both physical quantum circuits. For measured qubits, all measurement operations will be at the end of the circuit according to IBM quantum computers. We observe that none of the circuits used in Figure 8 exhibits any reduction in their physical qubit lifetime after running the rescheduling algorithm, which confirms the effectiveness of our approach in reducing the impact of Pauli errors on the output state fidelity of the quantum circuit.

The run time of our proposed elementary gate rescheduling algorithm for each benchmark circuit to be executed on various quantum architectures is shown in Table IV. Each quantum circuit requires only a small fraction of a second for gate rescheduling time.

C. Complex gate rescheduling results

In the second experiment, we show the effectiveness of our proposed ZZ complex gate rescheduling algorithm with respect to AR, and the corresponding ARG of different QAOA quantum circuits, namely QAOA_1, QAOA_2, and QAOA_3, executed on IBM Q16 Melbourne quantum computer. We refer to the complex gate rescheduling algorithm as Prop. II. We also study the impact of applying Prop. II followed by Prop. I approach to generate the physical QAOA quantum circuit, which implies the use of ZZ complex gate rescheduling algorithm first followed by elementary gate rescheduling algorithm after complex gate decomposition based on the commutation rules in Figure 1. This process is referred to as Prop. I+II.

Figure 9(a) and (b) provide a comparison of the AR, and the corresponding ARG, respectively, of different QAOA circuits generated using Base, Prop. I, Prop. II, and Prop. I+II. The depth of physical QAOA_1, QAOA_2, and QAOA_3 quantum circuits are 315, 375, and 412, respectively, for all the proposed rescheduling approaches. The range of the number of reordering for QAOA_1, QAOA_2, and QAOA_3 quantum circuits are 3-35, 5-39, and 7-53, respectively.

The WESP computed at the complex and the elementary levels of the base and reordered QAOA quantum circuits and their corresponding ESP value.

TABLE V: WESP of QAOA quantum circuits generated using the Base, Prop.I, Prop. II, and Prop. I+II and their corresponding ESP value.

| Benchmark | ESP | WESP (complex) | WESP (elementary) |
|-----------|-----|----------------|--------------------|
| Base      |     |                |                    |
| QAOA_1    | 1.01E-04 | 5.01E-05 | 6.32E-05 |
| QAOA_2    | 6.81E-07 | 2.49E-07 | 4.58E-07 |
| QAOA_3    | 9.05E-10 | 3.78E-10 | 5.63E-10 |
| Prop. I   | 5.36E-07 | 2.17E-07 | 3.11E-07 |
| Prop. II  | 6.21E-07 | 4.98E-07 | 7.92E-10 |

I+II. The depth of physical QAOA_1, QAOA_2, and QAOA_3 quantum circuits are 315, 375, and 412, respectively, for all the proposed rescheduling approaches. The range of the number of reordering for QAOA_1, QAOA_2, and QAOA_3 quantum circuits are 3-35, 5-39, and 7-53, respectively.

The WESP computed at the complex and the elementary levels of the base and reordered QAOA quantum circuits and their corresponding ESP value.

Our results show that while ZZ gates rescheduling algorithm improves the AR of the QAOA circuits, and thus, reduces the ARG, applying elementary gate rescheduling can provide a better output state fidelity. Our results also show that using both rescheduling algorithms at the complex and the elementary gate-level can deliver the best reduction in the ARG, and therefore, the best output state fidelity. Thus, as both decoherence and gate errors contribute to the quantum circuit noise, by incorporating our gate rescheduling approaches that push very noisy gates to as later circuit layers as possible we can improve the output state fidelity even for QAOA circuits with large depth.

The run time of our proposed rescheduling approaches for
QAOA-1, QAOA-2, and QAOA-3 quantum circuits to be executed on IBM Q16 Melbourne quantum computer is shown in Table VII. Our proposed approaches consume only a small fraction of a second despite the large circuit depth and the number of qubits.

### VII. Conclusion

In this paper, we propose gate rescheduling algorithms based on the gate error propagation paths in the quantum circuits. Given the variation in the error rates of NISQ computers, we show that the location of the quantum gate can significantly affect the output state fidelity of the quantum circuit. Our proposed approaches can be easily integrated with other quantum compilation approaches, which ensure that the quantum circuit mapping process maximizes ESP. We also anticipate further improvements in the output state fidelity when applying our approaches with other compilation approaches that target coherence and correlated errors. Our future work will further investigate the quantum circuit structures that benefit the most from our proposed gate rescheduling algorithms.

### VIII. Acknowledgement

We thank IBM Quantum for the access to the IBM processors through the IBM Quantum Researchers Program.

### References

[1] E. Magesan, J. M. Gambetta, and J. Emerson, “Characterizing quantum gates via randomized benchmarking,” Phys. Rev. A, vol. 85, p. 042311, 2012.

[2] S. Nishio, Y. Pan, T. Sato, H. Amano, and R. V. Meter, “Extracting success from ibm’s 20-qubit machines using error-aware compilation,” CoRR, vol. abs/1903.10963, 2019. [Online]. Available: http://arxiv.org/abs/1903.10963

[3] S. T. Flammia and J. J. Wallman, “Efficient estimation of pauli channels,” ACM Transactions on Quantum Computing, vol. 1, no. 1, Dec 2020. [Online]. Available: https://doi.org/10.1145/3408039

[4] Y. Nam, N. J. Ross, Y. Su, A. M. Childs, and D. Maslov, “Automated optimization of large quantum circuits with continuous parameters,” npj Quantum Inf, vol. 4, no. 23, 2018.

[5] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information, Cambridge University Press, 2019.

[6] D. Maslov, G. Dueck, and D. Miller, “Toffoli network synthesis with templates,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 24, no. 6, pp. 807–817, 2005.

[7] K. Hietala, R. Rand, S.-H. Hung, X. Wu, and M. Hicks, “A verified optimizer for quantum circuits,” Proc. ACM Program. Lang., vol. 5, no. POPL, Jan 2021. [Online]. Available: https://doi.org/10.1145/3434318

[8] S. Janardan, Y. Tomita, M. Gutierrez, and K. R. Brown, “Analytical error analysis of Clifford gates by the fault-path tracer method,” Quantum Information Processing, vol. 15, no. 8, pp. 3065–3079, 2016.

[9] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithm,” 2014.

[10] S. S. Tannu and M. K. Qureshi, “Not all qubits are created equal: A case for variability-aware policies for nisq-era quantum computers,” in Proceedings of ASPLOS. ACM, 2019, pp. 987–999.

[11] M. Alam, A. Ash-Saki, and S. Ghosh, “Circuit compilation methodologies for quantum approximate optimization algorithm,” in 2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), 2020, pp. 215–228.

[12] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, “Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices,” Phys. Rev. X, vol. 10, p. 021067, Jun 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevX.10.021067

[13] M. Y. Siraichi, V. F. d. Santos, S. Collange, and F. M. Q. Pereira, “Qubit allocation,” in Proceedings of CGO. ACM, 2018, pp. 113–125.

[14] J. Kusyk, S. Saeed, and M. U. Uyar, “Survey on quantum circuit compilation for noisy intermediate-scale quantum computers - artificial intelligence to heuristics,” IEEE Transactions on Quantum Engineering, pp. 1–1, 2021.

[15] T. Itoko, R. Raymond, T. Imamichi, and A. Matsuo, “Optimization of quantum circuit mapping using gate transformation and commutation,” 2019.

[16] A. Matsuo, W. Hattori, and S. Yamashita, “Reducing the overhead of mapping quantum circuits to ibm q system,” in 2019 IEEE International Symposium on Circuits and Systems (ISCAS), 2019, pp. 1–5.

[17] R. Wille, L. Burgholzer, and A. Zulehner, “Mapping Quantum Circuits to IBM QX Architectures Using the Minimal Number of SWAP and H Operations,” in Proceedings of DAC, 2019, pp. 142–1.

[18] A. Zulehner, A. Paler, and R. Wille, “An efficient methodology for mapping quantum circuits to the ibm qx architectures,” 2018.

[19] 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, ser. ASPLOS ’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1001–1014.

[20] A. Ash-Saki, M. Alam, and S. Ghosh, “Qure: Qubit re-allocation in noisy intermediate-scale quantum computers,” in Proceedings of ACM/IEEE Design Automation Conference, 2019, pp. 141:1–141:6.

[21] M. Alam, A. Ash-Saki, and S. Ghosh, “Addressing temporal variations in qubit quality metrics for parameterized quantum circuits,” in 2019 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED), 2019, pp. 1–6.

[22] S. S. Tannu and M. Qureshi, “Ensemble of diverse mappings: Improving reliability of quantum computers by orchestrating dissimilar mistakes,” in Proceedings of MICRO, 2019, pp. 253–265.

[23] N. Acharya and S. M. Saeed, “A lightweight approach to detect malicious/unexpected changes in the error rates of nisq computers,” in Proceedings of the 39th International Conference on Computer-Aided Design, ser. ICCAD ‘20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/340302.3415684

[24] P. Murali, D. C. McKay, M. Martinosini, and A. Javadi-Abhari, “Software mitigation of crosstalk on noisy intermediate-scale quantum computers,” in Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems, 2020, p. 1001–1016.

[25] G. G. Guerreschi and J. Park, “Two-step approach to scheduling quantum circuits,” Quantum Science and Technology, vol. 3, no. 4, p. 045003, Jul 2018. [Online]. Available: http://dx.doi.org/10.1088/2058-9565/aacf0b

[26] M. Alam, A. A. Saki, and S. Ghosh, “An efficient circuit compilation flow for quantum approximate optimization algorithm,” in 2020 57th ACM/IEEE Design Automation Conference (DAC), 2020, pp. 1–6.

[27] B. Tan and J. Cong, “Optimal layout synthesis for quantum compilation,” in Proceedings of the 39th International Conference on Computer-Aided Design, ser. ICCAD ‘20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/340302.3415620

[28] Y. Shi, N. Leung, P. Gokhale, Z. Rossi, D. I. Schuster, H. Hoffmann, and R. Wille, “Optimized compilation of aggregated instructions for realistic quantum computers,” in Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems, Apr 2019. [Online]. Available: http://dx.doi.org/10.1145/3297858.3304018

[29] S. Resch, S. Tannu, U. R. Karpuczko, and M. Qureshi, “A day in the life of a quantum error,” IEEE Computer Architecture Letters, vol. 20, no. 1, pp. 13–16, 2021.

[30] G. E. Crooks, “Performance of the quantum approximate optimization algorithm on the maximum cut problem,” 2018.

[31] D. Venturelli, M. Do, E. Rieffel, and J. Frank, “Compiling quantum circuits to realistic hardware architectures using temporal planners,”
[32] E. Bernstein and U. Vazirani, “Quantum complexity theory,” *SIAM J. Comput.*, vol. 26, no. 5, p. 1411–1473, Oct. 1997.

[33] W. van Dam, S. Hallgren, and L. Ip, “Quantum algorithms for some hidden shift problems,” in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA ’03. USA: Society for Industrial and Applied Mathematics, 2003, p. 489–498.

[34] H. Abraham and et al., “Qiskit: An open-source framework for quantum computing,” 2019.

[35] Q. A. team and collaborators, “Cirq,” Oct. 2020. [Online]. Available: [https://doi.org/10.5281/zenodo.4062499](https://doi.org/10.5281/zenodo.4062499)

[36] S. A. Cuccaro, T. G. Draper, S. A. Kität, and D. Petrie Moulton, “A new quantum ripple-carry addition circuit,” *arXiv e-prints*, pp. quant-ph/0410184, Oct. 2004.

[37] R. Wille, D. Grosse, L. Teuber, G. W. Dueck, and R. Drechsler, “Revlib: An online resource for reversible functions and reversible circuits,” in *Proceedings of International Symposium on Multiple Valued Logic*, 2008, pp. 220–225.

[38] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, “Validating quantum computers using randomized model circuits,” *Physical Review A*, vol. 100, no. 3, Sep 2019. [Online]. Available: [http://dx.doi.org/10.1103/PhysRevA.100.032328](http://dx.doi.org/10.1103/PhysRevA.100.032328)