A New Routing Strategy to Improve Success Rates of Quantum Computers

Fang Qi
fqi2@tulane.edu
Tulane University
New Orleans, Louisiana, USA

Xin Fu
xfu8@central.uh.edu
University of Houston
Houston, Texas, USA

Xu Yuan
xyuan@udel.edu
University of Delaware
Newark, Delaware, USA

Nian-Feng Tzeng
tzeng@louisiana.edu
University of Louisiana at Lafayette
Lafayette, Louisiana, USA

Lu Peng
lpeng3@tulane.edu
Tulane University
New Orleans, Louisiana, USA

ABSTRACT

In the current noisy intermediate-scale quantum (NISQ) Era, Quantum Computing faces significant challenges due to noise, which severely restricts the application of computing complex algorithms. Superconducting quantum chips, one of the pioneer quantum computation technologies, introduce additional noise when moving qubits to adjacent locations for operation on designated two-qubit gates. The current compilers rely on decision models that either count the swap gates or multiply the gate errors when choosing swap paths at the routing stage. Our research has unveiled the overlooked situations for error propagations through the circuit, leading to accumulations that may affect the final output.

In this paper, we propose Error Propagation-Aware Routing (EPAR), designed to enhance the compilation performance by considering accumulated errors in routing. EPAR's effectiveness is validated through benchmarks on a 27-qubit machine and two simulated systems with different topologies. The results indicate an average success rate improvement of 10% on both real and simulated heavy hex lattice topologies, along with a 16% enhancement in a mesh topology simulation. These findings underscore the potential of EPAR to advance quantum computing in the NISQ era substantially.

CCS CONCEPTS

• **Computer systems organization** → *Quantum Computing*.

KEYWORDS

Quantum Computing, Routing Compilation, Performance

1 INTRODUCTION

Superconducting quantum computers, one promising platform in Quantum computing, are known for their scalability and experimental viability [3]. It offers advantages in qubit scalability, microwave operation control, and fast gate operations, but faces challenges with noise and high error rates [2, 3]. Quantum error correction (QEC) addresses these errors but is currently impractical for noisy intermediate-scale quantum (NISQ) era machines characterized by a few hundred qubits [9]. Noise in superconducting quantum computers arises from control parameter fluctuations, thermal effects, environmental interactions, etc. When compiling logic circuits, routing, which moves distance logical qubits adjacent, often triggers numerous swap operations, significantly impacting computational performance. Therefore, mitigating routing-induced errors is essential for harnessing the full potential of quantum computing in practical applications.

Traditional routing decision models have two main approaches. Finding paths with the fewest swap gates is a prevalent approach in NISQ. Newer NISQ-centric models prioritize noise variation, avoiding paths involving physical links with high error rates. These noise-aware models, guided by the Estimated Success Probability (ESP) method [8], estimate a circuit's Success Rate (SR) by multiplying individual gate success rates [8, 13], seek paths with the lowest total gate error rate along the route.

While the ESP-based method is effective in considering noise during quantum circuit compilation, it neglects error propagation, notably in 2-qubit operations, potentially amplifying errors and disrupting calculations. It also fails to consider that only accumulated errors in measured qubits affect circuit output, possibly leading to overestimated error impacts. These neglects can lead to inefficient routing decisions, increased error rates, and reduced circuit performance.

Addressing this gap, we introduce Error Propagation-Aware Routing (EPAR), a novel approach that integrates error propagation effects into routing decisions. EPAR, designed to minimize error propagation to measurement gates, includes Path Exploration & Selection and Path Cost modules for optimizing swap paths and calculating accumulative error rates with error propagation considerations. Our evaluation of EPAR's performance was conducted through benchmarks on a 27-qubit quantum machine and two simulated systems with varying topologies. The outcomes showed an approximation of a 10% increase in average success rates for both

actual and simulated heavy hex lattice topology and an impressive 16% enhancement in mesh topology simulations. These findings underscore EPAR's capacity to substantially propel quantum computing forward in the NISQ era. The contributions of our paper are listed below.

- Identification of a significant gap in existing noise-aware routing methods, which overlook the impact of error propagation and qubit error accumulation on path selection.
- We propose and develop a novel error propagation-aware routing methodology designed to integrate error propagation insights for optimizing circuit performance.

2 BACKGROUND

2.1 Compilation and Noise

For a logic circuit to run effectively on a specific quantum computer, the compilation process must create a tailored and optimized executable circuit, accounting for the machine's specifications and noise profile [4]. This process includes mapping logical qubits to physical ones, strategically inserting SWAP gates during routing, converting gates to machine-compatible types, and performing optimization steps.

Superconducting quantum computers encounter errors due to environmental conditions and control mechanism imprecisions. These errors are quantified through operational error rates estimated via randomized benchmarking and retention errors assessed through relaxation (T1) and decoherence (T2) times [4]. Randomized benchmarking analyzes sequences of random gates to provide an average error rate, distinguishing gate errors from state preparation and measurement errors.

2.2 Routing Related Work

In routing, minimizing swap gates is crucial to lower operational complexity, conserve computational resources, and enhance performance. Each swap gate typically consists of three CNOT gates, with an error rate about ten times higher than that of a single-qubit gate, underlining the importance of swap reduction [12, 15]. Currently, the noise-aware routing, guided by the Estimated Success Probability (ESP) method [8], optimizes decision models to favor paths with the lowest total error rates, considering the machine's noise profile [7, 8, 13]. Error propagation analysis in quantum circuits, crucial for performance assessment, has been studied previously [11], examining how errors from specific noise sources impact outcomes. However, such a study relies on individual noise probabilities, which are often absent in current quantum computer error reports. Conversely, our methodology utilizes error rates from randomized benchmarking to provide practical insights for error propagation analysis, addressing a gap left by earlier investigations.

$$g_i^s = (1 - g_i^e), \; m_i^s = (1 - m_i^e) \tag{1}$$

$$\text{ESP} = \prod_{i=1}^{N_{\text{gates}}} g_i^s \times \prod_{i=1}^{N_{\text{Meas.}}} m_i^s \tag{2}$$

3 ERROR PROPAGATION ANALYSIS

The ESP method is commonly used for quick success rate estimation in quantum circuits and guiding path selection in error mitigation

techniques. As in Equation 2, ESP predicts correct output trial probability by multiplying the success rate (or fidelity) of each gate (g_i^s) and measurement (m_i^s) operation, derived from one minus the respective gate (g_i^e) and measurement (m_i^e) error rates as shown in Equation 1. However, ESP simplifies its approach by uniformly treating all gates and qubits, overlooking crucial aspects such as error propagation and measurement effects on subsystems.

Error propagation in quantum circuits, as illustrated in Fig. 1, demonstrates how errors, quantified through randomized benchmarking, propagate through one-qubit and two-qubit gates, as well as measurement gates, encompassing various error types such as amplitude, phase, environmental, and retention errors. These errors, stored within individual qubits in quantum chips, can spread to others during multiqubit operations, affecting overall circuit performance. Analyzing this cumulative effect involves converting qubit error rates into success rates and accumulating them over successive cycles. This enhances our understanding of the quantum system's behavior and informs the development of more effective error mitigation strategies and optimization techniques. ESP fails to capture these propagation dynamics, which is particularly evident in the error propagation through 2-qubit gates. Moreover, ESP overlooks that only errors from measured qubits contribute to a circuit's final output, disregarding errors from unmeasured qubits. Incorporating these factors into noise models can lead to more accurate predictions of quantum circuit success rates, informing quantum compiler and circuit optimizations.
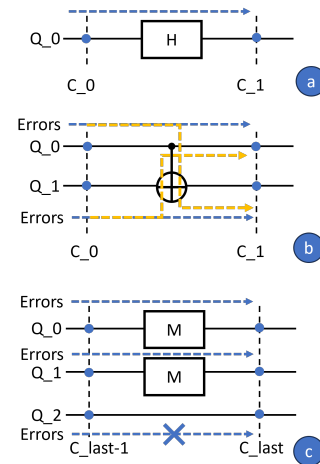
Figure 1: Error propagation cases in compiled circuits with 1-qubit, 2-qubit, and measurement gates. Yellow arrows and Blue cross highlight aspects overlooked by ESP.

4 EPAR-BASED COMPILATION

4.1 Impact of Error Propagation on Routing

We analyze how error propagation influences quantum circuit compilation routing, affecting circuit performance. A hexagonal topology quantum computer illustrates the impact of different routing models on path selection, as shown in Fig. 2. Using real IBM quantum hardware data, we examine a basic logic circuit with three

CNOT and three measurement gates. Assuming error-free measurement gates for simplicity, two paths are possible: a long path (three swap gates) and a short path (one swap gate), each swap gate comprising three CNOT gates. ESP and EPAR yield different success rate (SR) estimations for both paths and select the higher SR path. In ESP-based compilation, the short path is preferred due to its higher estimated success rate and shorter swap distance. Conversely, EPAR favors the long path, considering reduced error propagation: the CNOT between qubits q1 and q2 doesn't add noise to qubit q0, and ancilla qubits on the long path, not being measured, don't contribute errors to the final output. This emphasizes the importance of error propagation and subsystem measurements in routing. EPAR-based routing, guided by error propagation analysis, leads to superior performance.
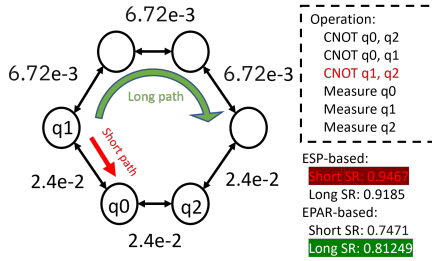

Figure 2: ESP- and EPAR-based compilation.

4.2 The Qubit Success Rate at The Cycle End

We present the process for calculating the success rate of a qubit q_n at the end of a cycle c_m. Notably, the success rate of a qubit or gate inversely relates to its error rate using Equation 1. Assuming that all qubits are ideally at the end of each cycle, having completed previous operations without starting new ones. The success rates at the end cycle are influenced by the combined success rate calculated using Equation 3, considering the probabilistic relationship between previous and current cycle errors.

$$SR_{q_n}^{c_m} = SR_{q_n}^{c_m-1} * SR_{errors\ in\ c_m} \qquad (3)$$

For 1-qubit gates (like rotations or phase gates), the $SR_{errors\ in\ c_m}$ is the gate fidelity. For 2-qubit gates (like CNOT) as seen in Fig. 1.b, the success rate at the cycle end is the combination of its previous cycle's SR, the gate fidelity (g_{cnot}^s), a portion w of its counterpart's error rate $1 - SR_{q_o}$ at the last cycle. The formula for this is given in Equation 4. Measurement gates, which typically have higher error rates, also influence the qubit success rate. The circuit's overall success rate depends on the success rates of the measured qubits, as these gates directly affect the circuit output.

$$SR_{q_n}^{c_m} = SR_{q_n}^{c_m-1} * g_{cnot}^s * [1 - w(1 - SR_{q_o}^{c_m-1})] \qquad (4)$$

This approach helps to precisely estimate the error propagation and the error accumulation effects leading to a better routing decision model. It can also guide the optimization of quantum algorithms and the design of reliable quantum circuits, enhancing computation efficiency and fidelity.

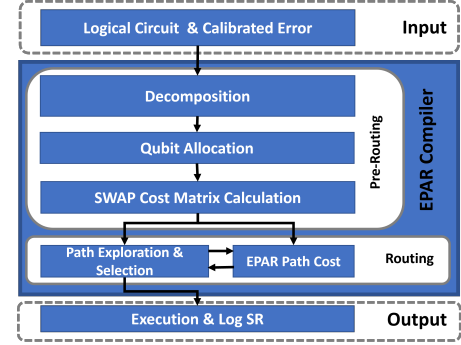## 4.3 Error Propagation-Aware Routing Overview

Figure 3: EPAR-based Compiler Flow.

The Error Propagation-Aware Routing (EPAR) is a novel approach designed to enhance quantum circuit performance by integrating insights from error propagation in routing decisions. The EPAR-based compiler, as illustrated in Figure 3, inputs a logical circuit and calibrated error rates from a target quantum machine.

The compilation begins with pre-routing stages including decomposing multi-qubit gates into 2-qubit gates and transforming logic gates into the target machine's native gate set. Such prepares the circuit with the machine's noise profile for error propagation-aware routing. The next qubit allocation module assigns logic qubits to physical qubits and could integrate other optimization passes at this stage for further enhancement.

In the Swap Cost Matrix Calculation Module, we employ the Floyd-Warshall algorithm to generate two matrices containing the shortest path for all node pairs, considering the machine error profile. This step generates alternative shortest paths, expanding path possibilities for swap candidate evaluation.

The routing process comprises the Path Exploration and Selection module and the EPAR Path Cost module. The former updates qubit success rates at each cycle and explores routing paths for 2-qubit gates needing adjacency adjustments. The latter quantifies the expected performance for each routing option.

The final routing decision aims to maximize total success rates on measure qubits, thus picking the routing that has the lowest error rate towards the final output. The methodology incorporates constraints and optimization objectives to align the routing with specific requirements.

4.4 Path Exploration and Selection

The Path Exploration & Selection module (Algorithm 1) is essential for achieving high-performance compiled quantum circuits. It iteratively processes the input circuit, updating qubit success rates based on gate operations and selecting optimal swap paths where necessary. This process ensures that each operation is executed effectively while considering the impact on qubit success rates.

The algorithm begins by initializing qubit success rates and mapping, then continuously updates the circuit with executable gates. For gates requiring adjacency adjustments, swap paths are evaluated based on their estimated success rates. The module selects

the most effective swap, applies it to the circuit, and adjusts qubit success rates accordingly.

Algorithm 1's complexity, mainly in the while loop evaluating swap candidates, scales as $O(G)$, where G represents the circuit's gate count. This module significantly boosts quantum circuit compilation efficiency and reliability through its emphasis on qubit success rates and optimal path selection.

Algorithm 1 Path Exploration & Selection
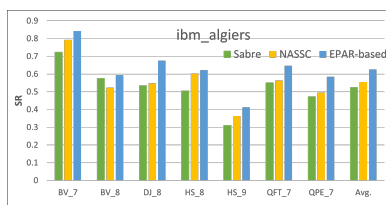

Require: $Circuit\ C; ErrorInfo.\ E$
Ensure: $Compiled\ Circuit\ C_{new}$
 1: Initialize SR_Q and qubit mapping
 2: Update circuit with unoperated gates
 3: **while** Circuit has unexecuted gates **do**
 4:     Execute operations if possible and update SR_Q
 5:     If no gates are executable, find swap candidates
 6:     Calculate cost of each swap using $Algorithm\ 2$
 7:     Add the most effective swap to the circuit
 8:     Update qubit success rates and mapping
 9:     Update circuit for next iteration
10: **end while**
11: Generate compiled circuit C_{new}
12: **return** C_{new}

4.5 EPAR Path Cost Module

The EPAR Path Cost Module is designed to estimate quantum circuit performance based on the given swap candidates. It starts by duplicating the circuit to a fixed number of future gates. After initializing the swap candidates, the algorithm iteratively updates qubit success rates for each gate operation within the FutureGates (lines 1-10). For every future 2-qubit gate encountered, the qubit success rate for both paths was calculated at $P1$ and $P2$, and the ones with the highest total success rate on measured qubits were chosen. The module calculates the cost by assessing gate interactions and their impact on qubits, which reflects the estimated success rate after executing the future gates(lines 11-15). It also penalizes circuit depth exceeding a preset limit to maintain estimation accuracy (lines 16-19).

Algorithm 2 evaluates a given swap candidate for its quantum circuit's success rate at a future point, considering gate execution, error information, and circuit depth restrictions. The Alg. 2's complexity depends on the number of future gates which is set as a constant. Therefore the complexity of the EPAR routing should be $O(G_f)$, where G_f represents the total future gate count.

Figure 4: Real success rate across all benchmarks on machine IBM_Algiers.

Algorithm 2 EPAR Path Cost Calculation

Require: $Circuit\ C; Swap; SR\_Q; Mapping; FutureGates;$
Require: $Error\ Info.\ E; Matrix\ P1, P2; Bonding\ Coef.\ b$
Ensure: $Cost$
 1: Initialize cost and duplicate circuit C with mapping for Swap
 2: Apply swap on SR_Q_copy and update mapping
 3: Trim C_copy based on FutureGates
 4: **for** operations in C_copy **do**
 5:     **if** operation not executable **then**
 6:         Determine optimal swap path based on P1 and P2
 7:         Update SR_Q_copy and mapping
 8:     **end if**
 9:     Mark operation as executed and update SR_Q_copy
10: **end for**
11: **for** qubit in SR_Q_copy **do**
12:     **if** qubit measured **then**
13:         Multiply $Cost$ by SR_Q_copy[qubit]
14:     **end if**
15: **end for**
16: Calculate depth_g using a greedy approach
17: **if** C_copy's depth exceeds limit **then**
18:     Adjust $Cost$ accordingly
19: **end if**
20: **return** $Cost$

---

5 EXPERIMENT SETUP

In our study, we utilized the Qiskit framework [4], an open-source platform for quantum computing, to implement and assess our EPAR methodology. Experiments were conducted using Qiskit-IBM-Runtime version 0.9.3 on IBM_Algiers, an actual 27-qubit IBM superconducting quantum machine featuring a heavy hex lattice topology. Additionally, EPAR routing was applied on two simulators, IBM_Montreal (mimicking heavy hex lattice) and IBM_Tokyo (mimicking mesh), employing real noise models derived from [10].

We evaluated EPAR's performance using benchmarks from [4, 5, 14], with half of the measurement gates randomly dropped to demonstrate performance for parietal measurement.

Default shot numbers were used for algorithm runs, facilitating a detailed assessment of the EPAR compiler's efficiency and circuit success rates. EPAR was compared with leading routing strategies SABRE [6] and NASSC [7]. Integration of other swap-reduction-based routing stages like [12, 15] into Qiskit-IBM-Runtime was not feasible due to compatibility issues. SABRE employs ESP for routing and considers future gates, while NASSC focuses on reducing CNOT gates by anticipating future optimizations.

Benchmark logic circuits were decomposed according to the target quantum machines' basis gates, and simulators were used to obtain reference outputs. Different routing strategies, including SABRE, NASSC, and EPAR, were applied to these circuits individually, followed by full compilation. The compiled circuits were then executed on the target quantum machine and simulators, with each experiment set to run 8192 shots. For EPAR, realistic settings involved choosing a bounding coefficient and weight value of 0.1. The performance of each routing strategy was assessed by calculating the real success rate (SR) based on the percentage of correct outputs logged in each experiment.

6 EVALUATION

This section details a thorough evaluation of our EPAR-based compiler's performance against the baseline SABRE and NASSC routing methods. As shown in Fig. 4, the results affirm EPAR's superior efficacy. In our tests, the EPAR-based compiler's success rates, illustrated in blue, present an average improvement of around 10% over the baseline methods represented in green (SABRE) and yellow (NASSC). The EPAR-based compiler's enhanced success rates stem from its strategic consideration of error propagation, enabling more informed routing choices. By evaluating the cumulative effect of errors during circuit execution, EPAR optimizes circuit performance, leading to more precise success rate predictions and improved outcomes.

In developing the EPAR, we encountered scenarios where ancilla qubits had to follow extended swap paths to avoid interaction with logic qubits measured at a circuit's end. This led to increased cycle times and retention errors due to prolonged T1 and T2 times, adversely affecting circuit performance. Although error characterization and randomized benchmarking focused on rotation errors, they are unable to accurately gauge the effects of these lengthy swap paths. To counter this, we introduced a bounding coefficient, b, in Alg. 2, to penalize routing choices leading to deep swap paths. Implementing this coefficient has effectively optimized routing decisions, enhancing compiled circuit success rates and striking a balance between circuit depth and performance, thus boosting reliability in quantum computing applications.

(a) IBM_Montreal
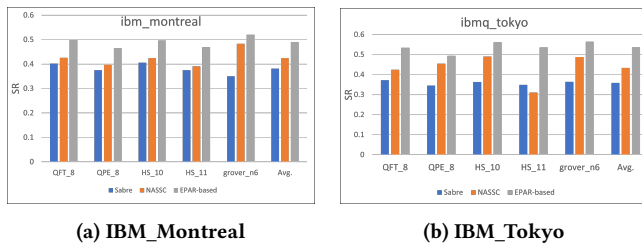
(b) IBM_Tokyo

Figure 5: Simulated results with real noise of IBM_Montreal and IBM_Tokyo quantum machine.

The EPAR excels in tracing error propagation during routing, favoring paths that minimize interaction with logic qubits, especially those leading to measurements. Theoretically, the EPAR compiler should perform better when fewer logic qubits have measurement gates and when the quantum machine offers more physical qubit connections, enhancing routing choices. The use of measurement gates, often a user's decision, is being studied to reduce their frequency and improve success rates [1].

To assess EPAR's effectiveness across different topologies, we tested it on IBM_Tokyo (mesh topology) and IBM_Montreal (hexagon topology), using real noise data from the Qiskit library. Benchmark results demonstrated that EPAR outperformed Sabre and NASSC, achieving a 20% and 12% increase in success rates on mesh topology, respectively, while maintaining approximately a 10% increase on hexagon topology, as illustrated in Fig. 5. This underscores how the higher connectivity in mesh topology provides more path options, reducing the reliance on deepening swaps and thereby enhancing success rates.

7 CONCLUSION

In this paper, we introduce EPAR, an advanced quantum circuit routing method that considers error propagation, outperforming traditional approaches. EPAR consistently achieves higher success rates, demonstrated on real machines and simulators with hexagon topology. It offers a comprehensive strategy, addressing swap operations, gate execution, error propagation, and circuit depth. Experiments on IBM_Algiers and IBM_Montreal reveal an impressive 10% average success rate improvement. Moreover, testing on IBM_Tokyo with mesh topology and real noise conditions shows EPAR surpassing the SABRE method by 16%. EPAR marks a significant leap in quantum circuit optimization, providing an effective and dependable solution for scalable quantum computing applications.

REFERENCES

[1] Poulami Das, Swamit Tannu, and Moinuddin Qureshi. 2021. Jigsaw: Boosting fidelity of nisq programs via measurement subsetting. In *MICRO-54: 54th Annual IEEE/ACM International Symposium on Microarchitecture*. 937–949.

[2] Yongshan Ding, Pranav Gokhale, Sophia Fuhui Lin, Richard Rines, Thomas Propson, and Frederic T Chong. 2020. Systematic Crosstalk Mitigation for Superconducting Qubits via Frequency-Aware Compilation. In *2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 201–214.

[3] Xiang Fu, Michiel Adriaan Rol, Cornelis Christiaan Bultink, J Van Someren, Nader Khammassi, Imran Ashraf, RFL Vermeulen, JC De Sterke, WJ Vlothuizen, RN Schouten, C. G. Almudever, L. DiCarlo, and K. Bertels. 2017. An experimental microarchitecture for a superconducting quantum processor. In *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture*. 813–825.

[4] IBM. [n. d.]. Open-Source Quantum Development. https://qiskit.org/. Retrieved on 04-16-2021.

[5] Ang Li, Samuel Stein, Sriram Krishnamoorthy, and James Ang. 2020. Qasmbench: A low-level qasm benchmark suite for nisq evaluation and simulation. *arXiv preprint arXiv:2005.13018* (2020).

[6] Gushu Li, Yufei Ding, and Yuan Xie. 2019. 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*. 1001–1014.

[7] Ji Liu, Peiyi Li, and Huiyang Zhou. 2022. Not All SWAPs Have the Same Cost: A Case for Optimization-Aware Qubit Routing. In *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. 709–725. https://doi.org/10.1109/HPCA53966.2022.00058

[8] Shin Nishio, Yulu Pan, Takahiko Satoh, Hideharu Amano, and Rodney Van Meter. 2020. Extracting success from ibm's 20-qubit machines using error-aware compilation. *ACM Journal on Emerging Technologies in Computing Systems (JETC)* 16, 3 (2020), 1–25.

[9] John Preskill. 2018. Quantum Computing in the NISQ era and beyond. *Quantum* 2 (2018), 79.

[10] Fang Qi, Kaitlin N Smith, Travis LeCompte, Nian-feng Tzeng, Xu Yuan, Frederic T Chong, and Lu Peng. Jan. 2024. Quantum Vulnerability Analysis to Guide Robust Quantum Computing System Design. *IEEE Transactions on Quantum Engineering* (Jan. 2024).

[11] Salonik Resch, Swamit Tannu, Ulya R Karpuzcu, and Moinuddin Qureshi. 2020. A day in the life of a quantum error. *IEEE Computer Architecture Letters* 20, 1 (2020), 13–16.

[12] Bochen Tan and Jason Cong. 2020. Optimal layout synthesis for quantum computing. In *Proceedings of the 39th International Conference on Computer-Aided Design*. 1–9.

[13] Swamit S Tannu and Moinuddin Qureshi. 2019. Ensemble of diverse mappings: Improving reliability of quantum computers by orchestrating dissimilar mistakes. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*. 253–265.

[14] Robert Wille, Lukas Burgholzer, and Alwin Zulehner. 2019. Mapping quantum circuits to IBM QX architectures using the minimal number of SWAP and H operations. In *Proceedings of the 56th Annual Design Automation Conference 2019*. 1–6.

[15] Chi Zhang, Ari B Hayes, Longfei Qiu, Yuwei Jin, Yanhao Chen, and Eddy Z Zhang. 2021. Time-optimal qubit mapping. In *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*. 360–374.