

Pinpointing the System Reliability Degradation in NISQ Machines

Betis Baheri^{*||}, Zixuan Xu^{†||}, Vipin Chaudhary[‡], Ying Mao[§], Bo Fang[¶], Shuai Xu[‡], Qiang Guan^{***}

^{*} Kent State University, Ohio, USA

[†] The Chinese University of Hong Kong, Shenzhen, China

[‡] Case Western Reserve University, Ohio, USA

[§] Fordham University, New York, USA

[¶] Pacific Northwest National Laboratory, Washington, USA

{bbaheri, qguan}@kent.edu, 119010368@link.cuhk.edu.cn,

{vxc204, sxx214}@case.edu, ymao41@fordham.edu, bo.fang@pnnl.gov

Abstract—The growth of the need for quantum computers in many domains such as machine learning, numerical scientific simulation, and finance has urged quantum computers to produce more stable and less error-prone results. However, mitigating the impact of the noise inside each quantum device remains a present challenge. This paper utilizes the system calibration data collected from the existing IBMQ machines, applying reliability degradation analysis to generate the reliability degradation matrix (RDM). We define multiple new evaluation metrics based on the reliability degradation matrix to compare the reliability between qubits, qubit topologies, and quantum machines. New evaluation metrics can be used for exploring the most error-robust quantum machine. This contribution increases the users' expectation of result accuracy. It opens the opportunities for studying the insight of correlation between qubits that may further motivate the quantum compiler design for the qubit mapping.

Index Terms—Quantum Computing, System, Reliability, Analysis

I. INTRODUCTION

Noisy Intermediate-Scale Quantum (NISQ) machines are increasingly used to demonstrate the benefits of quantum computing for the high-performance computing (HPC) domain [5], [14], [15], [21], [24], [25]. A significant concern for NISQ machines is that the noise experienced by such machines is tightly affecting the execution of a quantum algorithm, and future NISQ machines are anticipated to suffer from this challenge with the increasing number of the qubits. [3], [11], [14], [21]. With the current availability and increased programmability of NISQ devices, the quantum computing systems and architecture community is identifying new problems and solutions in the space of quantum algorithm execution on erroneous quantum computing architectures. Understanding the error behavior based on calibration data and the relation between qubits errors, frequency, and readout errors is the key to minimizing the unexpected result and making quantum computing more useful. Previous research focused on intelligently mapping a quantum algorithm on different parts of a NISQ machine, each with a different error rate for different kinds of operations, to minimize the probability of errors occurring during quantum algorithm execution [2], [9], [13], [16], [17], [23].

^{||} These authors contributed equally.

^{***} Corresponding author: Qiang Guan (qguan@kent.edu).

While there is much work studying the performance of quantum computers, defining metrics and developing benchmarks for evaluating the performance of quantum computers is demanding yet significantly challenging. Differences in the implementation of the quantum hardware make it challenging to propose performance metrics that may be adopted across the different technologies of the quantum computers, e.g. trapped Ion or superconducting. In order to fill the gap, IBM proposed quantum volume (QV) [8] as a metric to quantify the performance of the quantum computing while executing the computation of the quantum circuits. The Quantum Volume method quantifies the largest random circuit of equal width and depth that the computer successfully implements. Quantum computing systems with high-reliability operations, high connectivity, large calibrated gate sets, and circuit rewriting toolchains are expected to have higher Quantum Volumes [12] QV is a property of the performance of quantum computer hardware and a service level agreement that a quantum computer can guarantee.

However, QV can only be used to statically compare the overall performance of NISQ machines, which is similar to measuring the peak performance of a high-performance computing system. When users need to choose the best performing IBM Q machines from the available system list, QV is not a helpful metric. The study from [28] has shown that there exists variation in the error rates of different qubits and links, which can have an impact on the decisions for qubits movement and qubit allocation. Even though the quantum computer may have the same QVs, the quantum computers' error rates at run-time are dynamically distinct between machines.

Rather than understanding the characteristics of quantum noise and the immediate influence on quantum applications, existing work mainly focuses on mitigating the noise [4], [7], [27], [29] and re-adjusting the results based on the outcome distribution of the quantum algorithms, and quantum noise data [19], [20]. However, error mitigation and result readjustment may either require extra qubits for error detection and correction or heavily depend on the property of the quantum algorithm. This is not applicable in the current NISQ era considering the limitation on the number of qubits and possible

connectivity of qubits from the topology.

This research focuses on analyzing the reliability degradation in NISQ machines. We utilize the system calibration data from the existing IBMQ machines, applying reliability degradation detection to generate the reliability degradation matrix. Based on the reliability degradation matrix, we define multiple new evaluation metrics to compare the reliability between the qubit topology of the quantum machines. These evaluation metrics can help users search for the most error-robust machine and qubits to expect the most accurate results. The insight of correlation between qubits we explored can further motivate the quantum compiler design for the qubit mapping.

The major contributions of this paper are:

- We enhance the understanding of the quantum error characteristics on NISQ machines by introducing the reliability degradation detection approach.
- We propose a new set of metrics for analyzing the reliability of NISQ computers.
- We analyze the IBM quantum computers using the evaluation metrics defined to explore the reliability degradation patterns in different qubit topologies, individual quantum machines, and individual qubits. The results from the study can (i) help design an efficient scheduling system for users to submit the jobs to quantum computers, with the focus of minimizing the impact of quantum errors; (ii). help the quantum compilers map logic qubits to physical qubits to exploit noise-resilience qubits better.

II. OVERVIEW OF THE QUANTUM SYSTEMS AND DATASET

Our study is based on the original data from the IBMQ quantum computing website. There are dozens of machines online that can be used for real-world quantum machines. The raw data we collected are from seven IBM quantum computing machines. *Athens*, *Bogota*, *Rome*, and *Santiago* use linear topology; *Vigo* and *Ourense* use tree topology and *Melbourne* use mesh topology. For each quantum machine, four main attributes ($T1$, $T2$, Readout Error and CNOT Error) influence their performances. We observed data of 114 consecutive days from machine *Athens*, *Rome*, *Bogota*, *Melbourne*, and *Santiago*, 61 consecutive days from machine *Ourense*, and 108 consecutive days from machine *Vigo*. We collected all four significant categories attributes of each machine above.

We look into four major categories of calibration data collected from IBM quantum computers [18]:

- 1) **$T1$ coherence time:** we call it amplitude damping as well. It is the period for a qubit's natural decay from the excited state to the ground state. Higher $T1$ value means the qubit is more reliable to stay at its own state.
- 2) **$T2$ coherence time:** we call it phase damping as well. It's the period for a qubit's state change due to environmental interaction. Higher $T2$ value of a qubit is essential to the reliability of a qubit, because a more significant number of operations can be accomplished before the output becomes erroneous beyond a tolerance limit. Qubits are error-prone because of high volatility and susceptibility to environmental perturbations.

3) **The readout error:** is the probability of incorrect measurement of a qubit state (referred to as 1-qubit readout operation).

4) **The CNOT (gate) error:** is the probability of introducing an error during a gate operation, for example, rotating a state of a qubit by a slightly erroneous angle.

Although the detail of building superconducting quantum machines is beyond this paper's scope, it is noted that each time the quantum circuit gets executed on a NISQ machine, the outcome of that execution depends on the frequency of each qubit as well as the $T1$ and $T2$ error rates [10]. If the machine is under low reliability between qubits or the calibration process was performed a long time ago, the likelihood of an erroneous outcome distribution would increase significantly. We consider such situations the reliability degradation [6].

Performing operations on qubits can also affect their state due to errors. Quantum operation errors can be categorized into three groups: a) single-qubit gate errors (also known as $U3$ gate errors); b) single-qubit readout errors; c) two-qubit gate errors (referred to as CX gate errors). Single and two-qubit gate errors occur when there is noise in the system when applying a gate to a qubit state. Readout errors are related to the faulty reading of the final qubit state; in NISQ machines, these errors are related to readout resonators. For publicly available IBM-Q quantum computers, the single-qubit instruction error rates are of the order of 10^{-3} , whereas for two-qubit instructions, such as CNOT, it is 10^{-2} . Google Quantum machine [26] is reported to have about one order of magnitude lower error rates than the IBM machines. However, detailed characterization data for this machine is not publicly available.

III. METHODOLOGY: DEGRADATION ANALYSIS

In the NISQ system, we target durable reliability degradation (DRD) rather than transient reliability degradation (TRD). The TRDs showing in time series are mostly sharp peaks and troughs (e.g., spikes or pulses) and may only last for a short period, which only impacts the performance of the quantum computers in a short period. On the contrary, DRDs in a time-series event stream are gentle peaks and troughs (e.g., level change or trends). DRDs usually last much longer than transient cases and could lead to catastrophic events. NISQ machines under DRD may be rectified with system calibrations. However, DRD is not usually detectable. Therefore we propose the degradation detection algorithm. We use the python package called 'Anomaly Detection Toolkit' (ADTK) [1], which detects anomalies in a given time series. The function we chose in ADTK is called 'PersistAD', which compares the value of each time series with its adjacent previous values. The function is implemented based on Double Rolling Aggregate, which rolls two sliding windows side-by-side along a time series, aggregates using a selected operation, and tracks difference of the aggregated metrics between the two windows. The algorithm is shown in Alg 1. We choose $c = -0.5$ to minimize the degradation detection errors. In this way, we can convert our raw data (hard to analyze) to be more understandable and readable (easy to analyze).

Algorithm 1: The reliability degradation detection algorithm

Input : Current value v from T_1 , T_2 , Readout Errors or CNOT Errors
Input : 25th percentile P_{25}
Input : 75th percentile P_{75}
Input : Control parameter c
Output : 1: anomalous state; 0: normal state

```
1 if  $v$  is from  $T_1$  or  $T_2$  then
2   if  $(c + 1)P_{25} - cP_{75} - v < 0$  then
3     return 0
4   else
5     return 1
6   end
7 else
8   if  $(c + 1)P_{75} - cP_{25} - v < 0$  then
9     return 1
10  else
11    return 0
12  end
13 end
```

The output of Alg 1 forms a reliability degradation matrix (RDM). We denote the RDM as D , for each entry d_{ij} , the value of d_{ij} can be 1 or 0, which means there is or is not a reliability degradation event occurring on the i_{th} sampling time for the j_{th} attribute respectively.

Then we use multiple ways to conduct the reliability degradation analysis based on RDM:

- We check each machine's attribute's degradation rate (defined in Section IV) and compare and analyze those rates between different topologies.
- We analyze the frequency of degradation occurrence for each machine. We check the Cumulative Distribution Function (CDF) of each machine's qubit for different attributes, namely T_1 , T_2 , and *ReadoutError* to convert the matrices into more readable graphs. We calculate the mean time between reliability degradation events for each machine's qubit. In this way, we can capture and analyze the characteristics and stability of machines by observing and comparing those CDF graphs.
- we use Spearman correlation coefficients to find the correlation between different attributes (the columns of the RDM). To visually analyze the correlation, we utilize heatmaps, which can display the magnitude of the correlation using different colors in the form of two-dimension cells, in which there are correlation values. We set the centre of the colorbar on the right-hand side of the heatmap as 0.3 [22], which can represent two attributes that have a moderate positive correlation. We abandoned those cells with negative values since the number of such cells is too small to analyze, and focusing on positive correlation is more meaningful. Thus, we mainly look for those values bigger than or equal to 0.3 for more details and insights.

IV. SYSTEM RELIABILITY DEGRADATION ANALYSIS

Instead of using QV, we define multiple new evaluation metrics for system reliability degradation analysis.

A. Reliability Degradation Rate (RDR)

We define the *reliability degradation rate* using $RDR_i = x/y$, where x = the number of entries of a column whose value is 1, and y = the number of the rows of column i in RDM. Therefore, the average RDR of a calibration data attribute is calculated by computing the average value of all reliability rates for that attributes of T_1 , T_2 , readout error (*RO*) and CNOT error (*CX*). The RDR results are shown in Table I. The row of "Linear Topology", "Tree Topology" and "Mesh Topology" represent the average of the machines with same type of topologies. We can observe that machines with mesh topology have a higher average RDR for every calibration data attribute, while machines with tree topology have the lowest average RDR. At the machine level, Vigo has the lowest average RDR, representing that Vigo is generally more stable than other machines, while machine Ourense has the highest average RDR among all given machines. We found that both Vigo and Ourense have tree topology, but their average RDRs are at quite different levels, implying the diversity of the noise conditions among tree topology machines.

B. Perfect Day Ratio (PDR)

We define the 'perfect day' as the day when there is no reliability degradation on any attributes of a machine on that day. We compute the 'perfect-day ratio', by using the number of perfect days to divide the total observation days. As shown in Table II, machine Vigo has the highest PDR while Ourense has the lowest PDR. In comparison, the difference in linear machines' PDR is quite small, which implies linear machines are relative stable.

C. Mean Time Between Reliability Degradation (MTBRD)

We compute the mean time between degradation events of each machine's qubit by analyzing the total number of degradation events detected on T_1 , T_2 , Readout error rate (*RO*), and CNOT error rate (*CX*). For each qubit, we look at all its corresponding attributes. For example, as for Q_1 of Rome, the corresponding attributes are $T_1 - Q_1$, $T_2 - Q_1$, $RO - Q_1$, $CX0 - 1$, and $CX1 - 2$. We compute the total number of degradation events on CNOTs ($CX - Q_1$) by summing all Q_1 related *CX*s column in RDM. Finally we get four values, $T_1 - Q_1$, $T_2 - Q_1$, $RO - Q_1$ and $CX - Q_1$. We then compute the $MTBRD - Q_1 = days / (T_1 - Q_1 + T_2 - Q_1 + RO - Q_1 + CX - Q_1)$.

The MTBRD of each qubit on all 5-qubits machines is shown in Table III. Vigo has the longest MTBRD for Q_1 , Q_2 , and Q_4 . Also, Vigo has the longest MTBRD considering all the qubits, while Ourense has the shortest MTBRD. These results illustrate the stability of Vigo and the instability of Ourense even though they are all 5-qubits machines with the same topology.

TABLE I
THE AVERAGE RDR FOR EACH MACHINE

Machine \ Attr	T1	T2	R-O Error	CNOT-Error	Average
Athens	0.2333	0.2140	0.2018	0.2083	0.2147
Bogota	0.2053	0.2123	0.2053	0.1996	0.2059
Rome	0.1912	0.1877	0.2000	0.1996	0.1944
Santiago	0.2211	0.1895	0.1877	0.1623	0.1916
Linear Topology	0.2114	0.1996	0.1978	0.1924	0.2003
Vigo	0.1759	0.1889	0.1593	0.1713	0.1740
Ourense	0.2197	0.2262	0.2230	0.2254	0.2235
Tree Topology	0.1917	0.2024	0.1822	0.1908	0.1918
Melbourne (Mesh Topology)	0.2187	0.2222	0.2041	0.2193	0.2170

TABLE II
THE PERFECT DAY RATE FOR EACH MACHINE

Machine's Name	Perfect Day Ratio
Athens	0.333
Bogota	0.360
Rome	0.360
Santiago	0.342
Ourense	0.049
Vigo	0.556
Melbourne	0.342

TABLE III
THE MTBRD ACROSS QUBITS OF EACH MACHINE

Qubit \ Machine	Athens	Bogota	Rome	Santiago	Ourense	Vigo
Q0	4.515	4.606	5.561	4.56	4.604	5.539
Q1	4.851	4.515	4.8	5.124	4.604	6.0
Q2	4.957	4.851	5.124	5.429	4.519	6.085
Q3	4.653	5.494	5.124	6.08	4.281	5.468
Q4	4.851	4.957	5.182	5.365	4.519	5.610
Average	4.765	4.884	5.158	5.311	4.561	5.740

D. Reliability Degradation Depth (RDD)

To better describe the reliability degradation condition, we define a new concept called reliability degradation depth, computed by summing the total number of degradation events of all the machine attributes in a day. As shown in Figure 1, For linear machines, Santiago has more days of $RDD < 6$. In contrast, Athens and Bogota have more days of $RDD > 6$. The density degrees of those RD curves from linear machines are almost the same. However, the RD curve of Ourense is the most consecutive among all six machines. Rome's RD curve is the sparsest. All five qubits machines have most of the $RDD < 8$.

E. The Cumulative Distribution Function (CDF)

The CDF of T1, T2 and readout error reliability degradation occurrences on each qubit are shown in Figure 2, 3, and 4. The longer time span that the curve can keep flat, the more stable the system's condition is. For T1 shown in Figure 2, Bogota's Q_3 and Q_4 have longer stages of zero slope and is relatively more stable compared to other qubits and other machines. For T2 shown in Figure 3, Santiago's Q_2 and Q_4 have a long stage of zero slopes at day 65. Rome's Q_3 has the longest stage of zero slopes among all qubits and machines. For Readout Errors shown in Figure 4, Athen, Santiago and Rome's qubits

TABLE IV
THE RATIO OF HAVING A STRONG CORRELATION BETWEEN ATTRIBUTES OF T1 AND T2

Topology	Ratio of strong correlation on special diagonals	Rate of strong correlation for all pairs
Linear	18/20 (90%)	45/100 (45%)
Tree	8/10 (80%)	18/50 (36%)
Mesh	15/15 (100%)	66/225 (29.3%)
Total	41/45 (91.1%)	129/375 (34.4%)

are showing diverse reliability between qubits while all qubits of Bogota are sharing similar reliability properties.

F. Reliability Degradation Correlation (RDC)

Based on the reliability degradation matrix, we compute the correlation between each calibration attribute columns. The pair-wised correlation coefficients are visualized as a heatmap. An example is shown in Figure 5. We define two calibration attributes with a "strong correlation" if the corresponding correlation coefficient is larger than 0.3 [22]. If the strongly correlated attributes are from the same qubit, we call it 'same-qubit high correlation'. We focus on the correlation of T_1 and T_2 related attributes, as shown in Table IV, we find that machines of these three topologies all have a high ratio of same-qubit high correlation, while comparing the ratio of strong correlations in all attribute pairs. Therefore the coherence time (T_1 and T_2) of a qubit is highly correlated in all types of topologies.

G. Sensitive Qubits (SQ)

Apparently qubits on the same machine are not performing equally from our observation. We then define the 'sensitive qubits'. A qubit of a machine is a sensitive qubit if the ratio of the pairs of attributes of a qubit in RDM having a strong correlation is bigger than or equal to 0.3. For example, Athens's sensitive qubit is Q_4 , Bogota's sensitive qubits are Q_0 and Q_3 , Santiago's are Q_1 , Q_2 , Q_3 and Q_4 , Rome's are Q_0 and Q_1 , Ourense's is Q_2 , and Vigo's are Q_0 , Q_2 , Q_3 and Q_4 . Therefore, when the compiler maps the logical circuits to quantum computers, we should avoid sensitive qubits as much as possible or assign less important roles to the sensitive qubits, e.g., ancilla qubits.

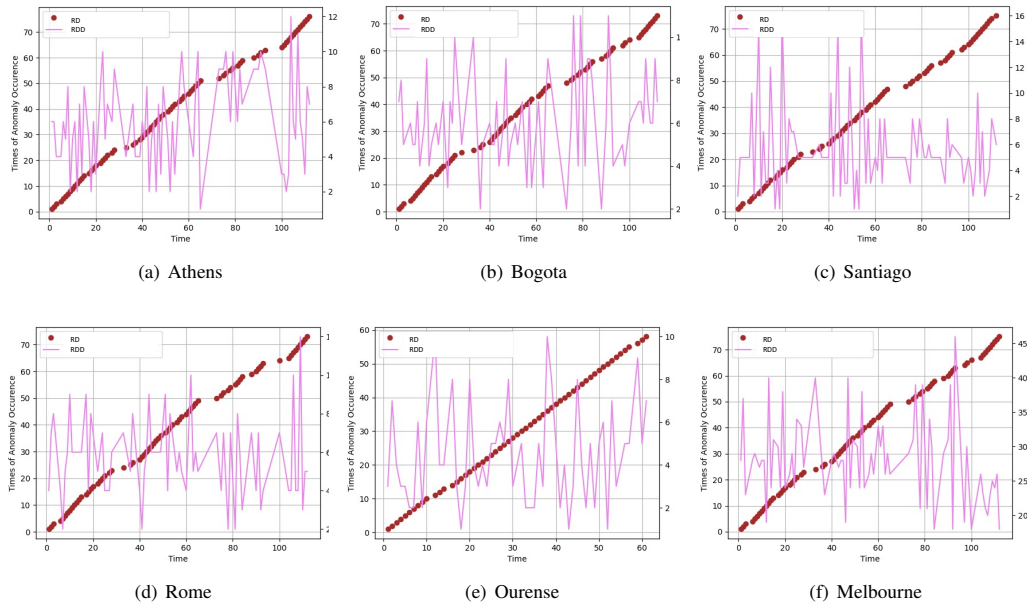


Fig. 1. Reliability degradation occurrence and depth. The left y-axis represents the number of reliability degradation occurrences, the right y-axis represents reliability degradation depth, and x-axis represents time (days).

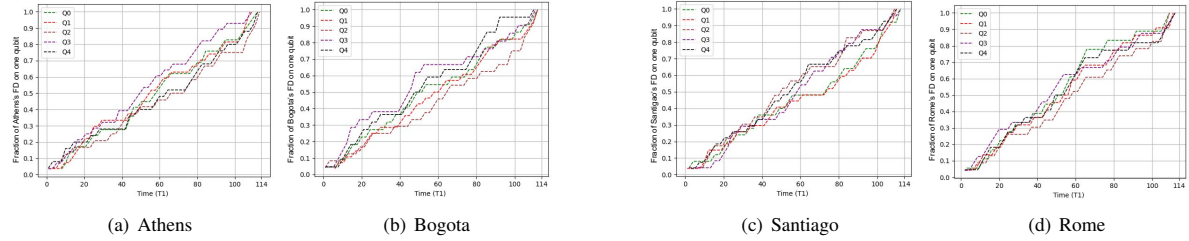


Fig. 2. CDF For Linear Machine's T1 of each qubit

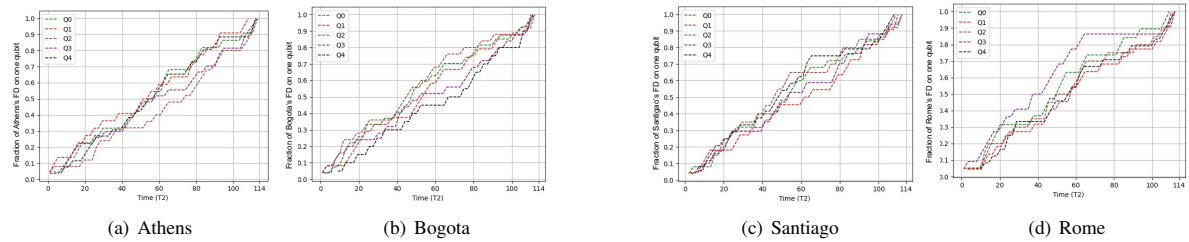


Fig. 3. CDF For Linear Machine's T2 of each qubit

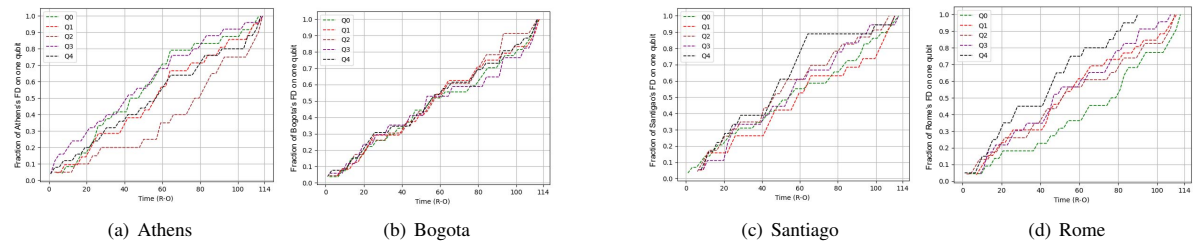


Fig. 4. CDF For Linear Machine's Readout Error of each qubit

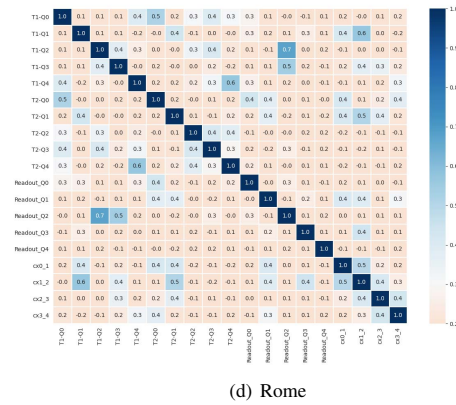
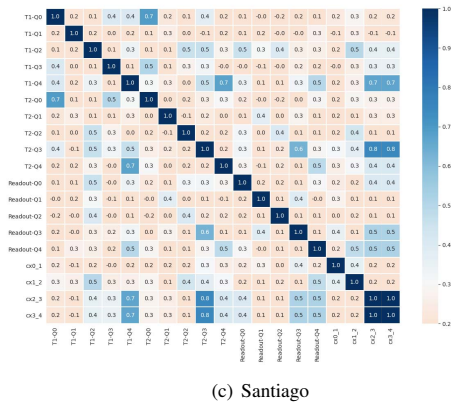
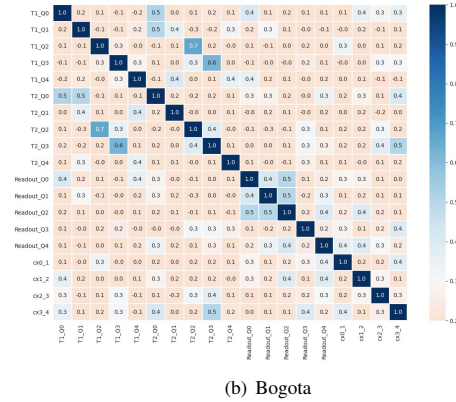
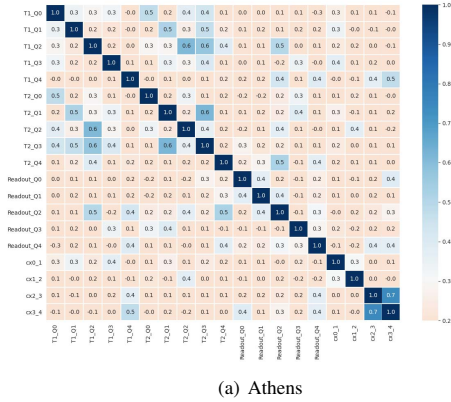


Fig. 5. Heatmaps For Linear Machine's Correlation Coefficients

H. Distance between Correlated Qubit Pairs(DQP)

When mapping logic qubits to physical qubits, it is essential to understand the degree of the qubits' reliability to avoid or reduce the use of less reliable qubits in swapping gates. Therefore, we define the distance between correlated qubit pairs for evaluating the routing algorithm for circuit transpilation. The distance between two qubits means the minimum steps we need to take to go from one qubit to another in the qubit connection graph. For example, the distance between Q_0 and Q_2 in Athens is 2, and the distance between Q_4 and Q_7 in Melbourne is 4. As shown in table V, all the pairs having a strong positive correlation (the absolute value of correlation coefficient > 0.3) have a distance smaller than or equal to 5. The qubit pairs that have a very weak correlation with each other (the absolute value of correlation coefficient < 0.05) distributes sparsely (from 1 to 8 distance). While considers building swapping gates, compiler should avoid using the strongly correlated pairs.

V. CONCLUSIONS

We present a novel reliability degradation analysis to study the reliability degradation in NISQ machines. We design a reliability degradation detection approach based on the collected

TABLE V
MELBOURNE'S STRONG CORRELATED QUBIT PAIRS AND WEAK CORRELATED QUBIT PAIRS

Distance Between Qubit Pair	1	2	3	4	5	6	7	8
Strong Correlation Pairs	3	2	4	3	2	0	0	0
Weak Correlation Pairs	3	6	2	6	2	2	0	1

system calibration data to generate the reliability degradation matrix that represents the system states regarding quantum errors. We define new evaluation metrics to help quantify the reliability of quantum computers, which can facilitate the design of compiler to minimize the impact from quantum errors by mapping the logical qubits to the most robust and independent qubits.

ACKNOWLEDGEMENT

This material is based upon work partially supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Science Center. The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830. This work was partially supported by NSF CCF #2216923 and CCF #2217021. Thanks to IBM Quantum Cloud and IBM Research Hub at NC State.

REFERENCES

- [1] Anomaly Detection Toolkit (ADTK), 2020.
- [2] Abdullah Ash-Saki, Mahabubul Alam, and Swaroop Ghosh. Qure: Qubit re-allocation in noisy intermediate-scale quantum computers. In *Proceedings of the 56th Annual Design Automation Conference 2019, DAC '19*, New York, NY, USA, 2019. Association for Computing Machinery.
- [3] Betis Baheri, Daniel Chen, Bo Fang, Samuel A. Stein, Vipin Chaudhary, Ying Mao, Shuai Xu, Ang Li, and Qiang Guan. TQEA: temporal quantum error analysis. In *51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2021, Taipei, Taiwan, June 21-24, 2021 - Supplemental Volume*, pages 65–67. IEEE, 2021.
- [4] Sergey Bravyi, Matthias Englbrecht, Robert König, and Nolan Peard. Correcting coherent errors with surface codes. *npj Quantum Information*, 4(1):55, Oct 2018.
- [5] Sergey Bravyi, Graeme Smith, and John A. Smolin. Trading classical and quantum computational resources. *Physical Review X*, 6(2), Jun 2016.
- [6] Michael Brooks. Beyond quantum supremacy: the hunt for useful quantum computers. *Nature*, 574(7776):19–21, October 2019.
- [7] Christopher Chamberland, Aleksander Kubica, Theodore J Yoder, and Guanyu Zhu. Triangular color codes on trivalent graphs with flag qubits. *New Journal of Physics*, 22(2):023019, Feb 2020.
- [8] Andrew W. Cross, Lev S. Bishop, Sarah Sheldon, Paul D. Nation, and Jay M. Gambetta. Validating quantum computers using randomized model circuits. *Phys. Rev. A*, 100:032328, Sep 2019.
- [9] Poulami Das, Swamit S. Tannu, Prashant J. Nair, and Moinuddin Qureshi. A case for multi-programming quantum computers. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture, MICRO '52*, page 291–303, New York, NY, USA, 2019. Association for Computing Machinery.
- [10] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications. *Phys. Rev. X*, 8:031027, Jul 2018.
- [11] Yipeng Huang and Margaret Martonosi. Statistical assertions for validating patterns and finding bugs in quantum programs. In *Proceedings of the 46th International Symposium on Computer Architecture, ISCA '19*, page 541–553, New York, NY, USA, 2019. Association for Computing Machinery.
- [12] IBM. Pushing quantum performance forward with our highest quantum volume yet.
- [13] Gushu Li, Yufei Ding, and Yuan 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, ASPLOS '19*, page 1001–1014, New York, NY, USA, 2019. Association for Computing Machinery.
- [14] Margaret Martonosi and Martin Roetteler. Next steps in quantum computing: Computer science's role, 2019.
- [15] Tiffany M Mintz, Alexander J Mccaskey, Eugene F Dumitrescu, Shirley V Moore, Sarah Powers, and Pavel Lougovski. Qcor: A language extension specification for the heterogeneous quantum-classical model of computation. *arXiv preprint arXiv:1909.02457*, 2019.
- [16] Prakash Murali, Norbert Matthias Linke, Margaret Martonosi, Ali Javadi Abhari, Nhung Hong Nguyen, and Cinthia Huerta Alderete. Full-stack, real-system quantum computer studies: Architectural comparisons and design insights, 2019.
- [17] Daniel Oliveira, Edoardo Giusto, Emanuele Dri, Nadir Casciola, Betis Baheri, Qiang Guan, Bartolomeo Montrucchio, and Paolo Rech. Qufi: a quantum fault injector to measure the reliability of qubits and quantum circuits. In *52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2022, Baltimore, MD, USA, June 27-30, 2022*, pages 137–149. IEEE, 2022.
- [18] Tirthak Patel, Abhay Potharaju, Baolin Li, Rohan Basu Roy, and Devesh Tiwari. Experimental evaluation of nisq quantum computers: Error measurement, characterization, and implications. In *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–15, 2020.
- [19] Tirthak Patel and Devesh Tiwari. Veritas: Accurately estimating the correct output on noisy intermediate-scale quantum computers. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '20*. IEEE Press, 2020.
- [20] Tirthak Patel and Devesh Tiwari. Qraft: Reverse your quantum circuit and know the correct program output. In *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS 2021*, page 443–455, New York, NY, USA, 2021. Association for Computing Machinery.
- [21] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.
- [22] Bruce Ratner. The correlation coefficient: Its values range between +1/-1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing*, 17, 06 2009.
- [23] Shaolun Ruan, Yong Wang, Ying Mao, Weiwen Jiang, Shuai Xu, and Qiang Guan. Vacsen: A visualization approach for noise awareness in quantum computing. In *2022 IEEE Visualization Conference (VIS)*, 2022.
- [24] Samuel A. Stein, Betis Baheri, Daniel Chen, Ying Mao, Qiang Guan, Ang Li, Shuai Xu, and Caiwen Ding. Quclassi: A hybrid deep neural network architecture based on quantum state fidelity. In Diana Marculescu, Yuejie Chi, and Carole-Jean Wu, editors, *Proceedings of Machine Learning and Systems 2022, MLSys 2022, Santa Clara, CA, USA, August 29 - September 1, 2022*. mlsys.org, 2022.
- [25] Martin Suchara, Yuri Alexeev, Frederic Chong, Hal Finkel, Henry Hoffmann, Jeffrey Larson, James Osborn, and Graeme Smith. Hybrid quantum-classical computing architectures. 11 2018.
- [26] Swamit Tannu and Moinuddin Qureshi. A case for variability-aware policies for nisq-era quantum computers. 05 2018.
- [27] Swamit S. Tannu and Moinuddin K. Qureshi. Mitigating measurement errors in quantum computers by exploiting state-dependent bias. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture, MICRO '52*, page 279–290, New York, NY, USA, 2019. Association for Computing Machinery.
- [28] Swamit S. Tannu and Moinuddin K. Qureshi. Not all qubits are created equal: A case for variability-aware policies for nisq-era quantum computers. In *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS '19*, page 987–999, New York, NY, USA, 2019. Association for Computing Machinery.
- [29] Theodore J. Yoder and Isaac H. Kim. The surface code with a twist. *Quantum*, 1:2, Apr 2017.