Stability of Quantum Computers

A Dissertation Presented for the

Doctor of Philosophy

The University of Tennessee, Knoxville

Samudra Dasgupta

May 2024

© by Samudra Dasgupta, 2024
All Rights Reserved.

Dedicated to my parents.

Acknowledgments

I am deeply thankful to Travis for his invaluable guidance, support, and expertise during my entire PhD journey. His steadfast dedication and insightful mentorship has significantly influenced my research and personal development. I also want to express my gratitude for the opportunity to utilize the resources of the Oak Ridge Leadership Computing Facility (a United States Department of Energy, Office of Science User Facility).

Abstract

Quantum computing's potential is immense, promising super-polynomial reductions in execution time, energy use, and memory requirements compared to classical computers. This technology has the power to revolutionize scientific applications such as simulating many-body quantum systems for molecular structure understanding, factorization of large integers, enhance machine learning, and in the process, disrupt industries like telecommunications, material science, pharmaceuticals and artificial intelligence. However, quantum computing's potential is curtailed by noise, further complicated by non-stationary noise parameter distributions across time and qubits. This dissertation focuses on the persistent issue of noise in quantum computing, particularly non-stationarity of noise parameters in transmon processors. It establishes a framework comprising computational accuracy, device reliability, outcome stability, and result reproducibility for assessing noisy outcomes amidst time-varying quantum noise. It further aims to determine the upper and lower bounds for this framework using available noise characterization data, in terms of the distance between time-varying noise densities. Using real data from a transmon processor, it validates the bounds on a test quantum circuit. It also demonstrates that if the physical platform's noise stays within the bounds determined by the analysis, experimental reproducibility can be guaranteed with a high degree of certainty. Furthermore, it develops a Bayesian algorithm to enhance outcome stability and accuracy for probabilistic error cancellation (PEC) in presence of time-varying quantum noise. The results obtained from experiments using a 5-qubit implementation of the Bernstein-Vazirani algorithm conducted on the ibm_kolkata device, underscore the effectiveness of the adaptive algorithm, showing a 42% improvement in accuracy over non-adaptive PEC and a 60% improvement in stability. Considering the time-varying stochastic nature of quantum noise, integrating adaptive estimation in error mitigation is crucial. In summary, by delving into the complexities of non-stationary noise in quantum computing, this dissertation provides valuable insights into quantifying and enhancing stability of outcomes from noisy quantum computers.

Table of Contents

1 Introduction **1**

1.1 Background 1

1.2 Research focus 9

1.3 Organization and notation 12

2 Noise in quantum computing **18**

2.1 Physical sources of noise 18

2.2 Cause of non-stationarity 19

2.3 Decoherence characterization 20

2.4 Modeling quantum noise channels 29

3 Performance evaluation framework **50**

3.1 Distance measures 51

3.2 Evaluation framework 57

3.3 Test circuit used 65

4 Reliability of device characterization **72**

4.1 Experimental data 72

4.2 Component reliability 73

4.3 Circuit reliability 76

5 Bounds on stability of program outcomes **97**

5.1 Sample bounds on accuracy 97

5.2 Sample bounds on reproducibility 104

5.3 General bounds on stability 105

5.4 General bounds on reliability 114

6 Enhancing histogram accuracy **128**

6.1 Uncorrelated noise 128

6.2 Correlated noise 133

7 Adaptive probabilistic error cancellation **151**

7.1 Background 153

7.2 Adaptive PEC 155

7.3 Numerical Validation 160

7.4 Experimental Testing 167

8 Conclusion **182**

Bibliography **186**

List of Tables

3.1 Moment-based distance by Distribution 56
3.2 Transition Probabilities for Uncorrelated Noise 69

4.1 The 16-parameter model derived from the ibm_washington data set has four types of quantum noise processes: (i) 'ABC': asymmetric binary channel model, (ii) 'CP': coherent phase error model, (iii) 'DP': depolarizing noise model and (iv) 'TR': thermal relaxation noise model. Note that the two-qubit model 'DP⊗DP' is a tensor product of depolarizing noise. 79
4.2 Hellinger distance values for the device parameters. 80
4.3 Gate lookup table 89

5.1 Register values for $\gamma_D(\tau)$ and $\gamma_{\max}$ 116
5.2 Device parameters 117

6.1 True means of the time-varying noise Pauli channel coefficients 144
6.2 Estimated Pauli coefficients for peiod 1 144
6.3 Estimated coefficients for period 2 145

List of Figures

1.1 Impact of the non-stationary noise, leading to irreproducible outcomes. The results show histograms (after state preparation and measurement (SPAM) noise mitigation) upon executing the Bernstein-Vazirani circuit on ibm_mumbai. 15
1.2 Evidence of spatial non-stationarity in the mean values for quibt decoherence times T_1 and T_2 for the 127-qubits of the ibm_washington device, generated on 14 Jan 2023 10:20 PM UTC. 15
1.3 state preparation and measurement (SPAM) fidelity distributions on ibm_toronto for qubits $0-8$ as measured on 8 April 8 2021, between 8:00-10:00pm (UTC-05:00). . . 16
1.4 state preparation and measurement (SPAM) fidelity distributions on ibm_toronto for qubits $9-17$ as measured on 8 April 8 2021, between 8:00-10:00pm (UTC-05:00). . . 16
1.5 state preparation and measurement (SPAM) fidelity distributions on ibm_toronto for qubits $18-26$ as measured on 8 April 8 2021, between 8:00-10:00pm (UTC-05:00). . 17

2.1 Interleaved sub-circuit post transpilation. 33
2.2 Spatial non-stationarity of decoherence times. Indivudal error-bars do not capture the variation across qubits. (a) In terms of the T_1 parameter, qubit 15 performed the best, exhibiting a median T_1 time of 184 microseconds, while qubit 4 performed the worst with a median T_1 time of 74 microseconds. Across all 27 qubits, the median T_1 time was 116 microseconds, with a standard deviation of 1.5 microseconds, and the range of T_1 times spanned from 32 to 297 microseconds. (b) For the T_2 parameter, qubit 3 demonstrated the highest performance, displaying a median T_2 time of 76 microseconds, while qubit 19 had the poorest performance with a median T_2 time of 15 microseconds. Across all 27 qubits, the median T_2 time averaged 30 microseconds, with a standard deviation of 1.1 microseconds, and the range of T_2 times varied from 3 to 191 microseconds. 34

2.3 This figure presents non-stationary temporal dynamics of decoherence for a qubit on IBM's belem device. The top figure shows T_1 relaxation time series for qubit 0, where two datasets were collected for 5 ms each on the same day, separated by a vertical line. The blue dataset varies between 116-126μs with a mean of 122 μs and a standard deviation of 2 μs, while the green dataset varies between 104-114μs with a mean of 108 μs and a standard deviation of 2 μs. The bottom figure displays Ramsey dephasing time (T_2) series for qubit 0, where two datasets were collected for 5 ms each on the same day, separated by a vertical line. The blue dataset varies between 67-73μs with a mean of 70 μs and a standard deviation of 1 μs, while the green dataset varies between 60-64μs with a mean of 62 μs and a standard deviation of 1 μs, collected around different times on the same day. The data shows significant non-stationarity in decoherence values over a 30-minute interval. 35
2.4 (a) Schematic of the 27-qubit device ibm_kolkata. (b) Estimated T_1 and T_2 time-series for qubit 0 as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023. 36
2.5 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 1 and (b) qubit 2. 37
2.6 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 3 and (b) qubit 4. 38
2.7 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 5 and (b) qubit 6. 39
2.8 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 7 and (b) qubit 8. 40
2.9 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 9 and (b) qubit 10. 41
2.10 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 11 and (b) qubit 12. 42
2.11 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 13 and (b) qubit 14. 43
2.12 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 15 and (b) qubit 16. 44
2.13 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 17 and (b) qubit 18. 45
2.14 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 19 and (b) qubit 20. 46

2.15 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 21 and (b) qubit 22. 47
2.16 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 23 and (b) qubit 24. 48
2.17 Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 25 and (b) qubit 26. 49

3.1 A classical Markov model for the correlated error process. 68
3.2 Simulation result for addressability F_A showing sensitivity to intra-register correlation. 68
3.3 Addressability: pairwise comparison when correlation parameter $u = 0.12$. Ideally, there should have been zero spatial variation (as idealized simulations do not differentiate between qubits at different physical locations). However, small fluctuations are seen about a mean value of 95.80 with a standard deviation of 0.02. This arises due to the readout error fluctuations in the ibm_yorktown noise model. 69
3.4 The Bernstein-Vazirani circuit for a 4-bit secret string $r = 1000$. H represents the Hadamard gate and Z represents the Z-gate. The meter symbols are measurement operations that project to the computational basis states. The measurement results are recorded in a classical register c. 70
3.5 Contribution to moment based distance (d) from increasing moment orders. The graph shows the results of comparing two normal distributions: $\mathcal{N}_1(\mu = \mu_0, \sigma = \sigma_0)$ and $\mathcal{N}_2(\mu = 2\mu_0, \sigma = 2\sigma_0)$ where $\mu_0 = 40$ and $\sigma_0 = 4$. 70
3.6 Simulated signal-to-noise ratios of the moment-based distance and total variation distance for two normal distributions of varying width. 71
3.7 When the two distributions are similar, then we expect each S_m to be zero. Empirically, that happens as we increase the sample size. The lower order moments take longer to go to zero. 71

4.1 Schematic layout of the ibm_yorktown device produced by IBM. Circles denote register elements and edges denote connectivity of 2-qubit operations. 78
4.2 Schematic of the ibm_toronto device, produced by IBM. Circles represent register elements, while edges denote the connectivity for performing two qubit operations. . 78
4.3 Experimentally observed probability density for SPAM (state preparation and measurement) fidelity of one of the register elements of the IBM transmon device named ibm_washington. 79

4.4 (a)-(e) Temporal stability of the SPAM fidelity F_{SPAM} of each register element in the ibm_yorktown device. The top panel shows the average F_{SPAM} of the register with associated variance, and the bottom panel shows a running calculation of the Hellinger distance using a one-month window. The dashed red line is the median value. . . . 80
4.5 Spatial stability of the SPAM fidelity for the ibm_yorktown device from May 2019 to December 2020, where the inset highlights the registers with the maximum distance. 81
4.6 (a)-(e) Temporal stability of the gate fidelity F_G for the CNOT gate for sequential register pairs in the ibm_yorktown device from March 2019 to December 2020. The top panel shows the average F_G of the register pair and the bottom panel shows a running calculation of the Hellinger distance with respect to May 2019. The dashed red line is the median value. 82
4.7 (a)-(d) Temporal stability of the CNOT duty cycle for sequential register pairs in the yorktown device. The top panel shows the harmonic mean of the register decoherence time T_2 for the elements, the upper-middle panel shows the gate duration T_G, the lower-middle panel plots the corresponding duty cycle τ, and the bottom panel presents the Hellinger distance for the duty cycle averaged over a one-month window. The dashed red line is the median value. 83
4.8 The spatial stability of the duty cycle τ for ibm_yorktown. The inset shows the experimental histograms for register pairs (0,1) and (2,3) which are separated by the largest Helligner distance of 0.789. 84
4.9 Spatial stability of the gate fidelity F_G for the CNOT gates of the ibm_yorktown device from March 2019 to December 2020. The inset shows the distribution of gate fidelities for pairs (1,2) and (3,2), which yield a Hellinger distance 0.467. 84
4.10 The figure presents results of reliability testing on a transmon platform over 16 months. Plots of the three Hellinger distance measures, with the unmodified measure H (blue line) being insensitive due to dimensionality issues. The blue and orange lines both capture the correlation structure of the joint distribution (with the orange line being normalized to enhance discrimination power), while the green line lacks correlation capture. The normalized measure $H_{\text{normalized}}$ ranges between 0.41 and 0.92, while the average measure H_{avg} varies between 0.431 and 0.51. The latter captures monthly variations in the marginal Hellinger distance for each of the 16 error parameters but fails to account for correlations. 85

4.11 Decomposition of the sources of quantum noise non-stationarity. The degree of non-stationarity varies amongst the sources. The plot shows the contributions made by each noise type to the composite Hellinger distance (a measure of the degree of non-stationarity). The Hellinger distance measures the statistical distance between the joint distribution of the noise observed in Apr-2023 to the joint distribution of the noise observed in Jan-2022. The various noise types contribute varying percentages but no single term dominates the sum. 86
4.12 The Pearson correlation coefficients between the 16 characterization parameters as observed in April-2023. Dark blue and dark red colors represent the Pearson coefficients 1 and -1, respectively. The axes labels $(x0, x1, \cdots x15)$ correspond to various noise sources as listed in Table 5.2. 87
4.13 Normalized mutual information of register pairs in the toronto device sampled 11:00-11:30 PM (UTC-5) on 11 December 2020. Data corresponds to register pairs prepared in the Bell state and the inset shows the range of the lowest and highest values for the normalized mutual information. 88
4.14 Experimentally observed duty cycle (τ) for the CNOT gate for washington. 88
4.15 The 95% temporal confidence interval (vertical orange lines) for CNOT duty cycle (τ) for the physical nearest-neighbor connections of ibm_washington. The data-set contains the values for all the 144 physical CNOT gates of ibm_washington between 1-Dec-2021 to 31-May-2022. The dashed black line represents the mean. 90
4.16 Temporal reliability of the SPAM fidelity F_{SPAM} of register element $q = 37$ 90
4.17 Temporal reliability of the gate fidelity F_G for the CNOT gate for register pairs 0 and 14. 91
4.18 Spatial reliability of the gate fidelity F_G for the CNOT gates of the ibm_washington device. The heat map shows the Hellinger distance between the distributions for the gate fidelities of nearest-neighbor connections. Only the lower triangular matrix is shown to avoid redundancy. The upper triangular matrix as well as any data gaps in the lower triangular matrix are colored white. The inset shows the distribution of gate fidelities for pairs $(11, 12)$ and $(19, 20)$, which yield a Hellinger distance 0.99. This metric captures the probability that the quantum processor register is spatially dissimilar. The estimation of the distribution for CNOT-i utilizes data for Sep-2022. 92

4.19 Spatial reliability of the duty cycle τ for the CNOT gates of the ibm_washington device. The heat map shows the Hellinger distance between the distributions for the duty cycles of nearest-neighbor connections. The inset shows the distribution of duty cycle for pairs (46, 47) and (96, 109), which yield a Hellinger distance 0.99. This metric captures the probability that the quantum processor register is spatially dissimilar. The estimation of the distribution for CNOT-i utilizes duty-cycle data for Sep-2022. 93

4.20 Spatial reliability of the SPAM fidelity F_{SPAM} for the ibm_washington device. The heat map shows the Hellinger distance between the distributions for each register pair. This metric captures the probability that the quantum processor register is spatially dissimilar. The inset shows the distributions of F_{SPAM} for registers 12 and 37, which represent distance=0.99 due to minimal overlap. The estimation of the distribution for qubit-i utilizes data for Sep-2022. 94

4.21 Spatial reliability of the gate fidelity F_G for the CNOT gates of the ibm_washington device. The heat map shows the Hellinger distance between the distributions for the gate fidelities of nearest-neighbor connections. Only the lower triangular matrix is shown to avoid redundancy. The upper triangular matrix as well as any data gaps in the lower triangular matrix are colored white. The inset shows the distribution of gate fidelities for pairs $(11, 12)$ and $(19, 20)$, which yield a Hellinger distance 0.99. This metric captures the probability that the quantum processor register is spatially dissimilar. The estimation of the distribution for CNOT-i utilizes data for Sep-2022. 95

4.22 Addressability of register pairs in the ibm_toronto device sampled 08:00-08:30 AM (UTC-5) on 11 December 2020. This data corresponds to the register prepared in the separable fiducial state. The inset shows the range i.e. the lowest and highest values for addressability. The average of (23,21) is the lowest value at 0.887 while all other values lie in the range [0.992, 1). The outlier is the only value that does not appear in the plot. 96

5.1 Circuit used for our experiment. In this figure, H represents the Hadamard gate. The meter symbol denotes measurement gate. 117

5.2 Plot illustrating significant spatial non-stationarity in the register-wise variation of the SPAM asymmetry for ibm_toronto. 118

5.3 The plot depicts distance variation across the 27 register elements of ibm_toronto device, illustrating the impact of spatially non-stationary noise on program outcomes and the dependence of output on register choice. 119

5.4 The plot illustrates the register-wise Hadamard gate angle error (in degrees) within ibm_toronto's quantum register of 27 elements, emphasizing spatial non-stationarity and cautioning against averaging qubit values for coherent noise mitigation. 120
5.5 This plot shows the register-wise variation of the composite accuracy metric γ_D for ibm_toronto device with 27 qubits, where a higher value significantly impacts program precision. The graph highlights the necessity of re-estimating the metric due to temporal non-stationary noise in unreliable devices, emphasizing the crucial role of analyzing noise parameter interactions for desired accuracy. 121
5.6 Characterizing circuit accuracy on toronto. Plot of $\gamma_{\max}$ (dashed line) and γ_D (blue dots) for ibm_toronto on 8 April 2021. The blue dots are experimentally-observed data plotted using the characterization data versus the actual observed distance (d) for each register element. Only a subset of qubits are shown. 122
5.7 Simulation demonstrating that when $H \leq H_{\max}$ (i.e. a reliable, slowly varying noise platform), then $s \leq s_{\max}$ (i.e. the temporal difference of the observable stays within the predicted upper bound). 123
5.8 A quantum circuit implementation of the Bernstein-Vazirani algorithm that employs 5 qubits, denoted q_0 to q_4. The first four qubits are used to compute the 4-bit secret string, while the fifth qubit serves as an ancilla and initially resides in the $|-\rangle$ superposition state. The symbol H denotes the Hadamard gate while the oracle unitary (U_r) implements the secret string (r). The depolarizing noise channel is denoted by $\mathcal{E}_{\mathrm{x}}(\cdot)$. A quantum measurement operation is represented by the meter box symbol at the circuit's end. 123
5.9 The ratio $s/s_{\max}$ for a simulated time-varying noisy circuit is plotted with respect to the increasing noise variance across 15 months. The results from noisy simulations of the Bernstein-Vazirani circuit with a secret string of 4, 8, or 12 bits validate the bound expected. 124
5.10 Contour plots to compare the probability densities of a two-dimensional subset of Hadamard gate errors for qubit 0 and 3 in Apr-2023, (a) with and (b) without correlation modeling using a copula function. 125
5.11 Simulations of the stability ratio s/s_{max} times 100 for a 4-qubit Bernstein-Vazirani circuit using the noise characterization from the ibm_washington platform. The box-and-whisker plot of the monthly statistics are based on noisy circuit simulations using the joint probability distribution derived from data from 1-Jan-2022 to 30-Apr-2023. Ratio values below unity confirm that the upper bound is never exceeded. 126

5.12 Schematic layout of the 127-qubit washington device produced by IBM. Circles denote register elements and edges denote connectivity of 2-qubit operations. The register elements 0, 1, 2, 3, and 4 are mapped to the physical qubits 4, 3, 2, 1, and 0, respectively, in the diagram above. The CNOT gates used in the circuit connect the physical qubits $(0, 1)$ and $(2, 1)$, in the diagram above, where the first number represents the control qubit and the second one represents the target qubit. 127

6.1 Time-varying density of SPAM fidelity. Data shown for register element #26. . . . 142

6.2 SPAM fidelity time-series for qubit #26 for Dec-May 2022. 142

6.3 We simulated SPAM and gate error channels for a quantum circuit with 4 qubits that creates a uniform superposition across all the computational basis states using Hadamard gates. The blue bars represent the probability distribution across the measurement outcomes for the ideal, noiseless circuit while the orange bars represent the same for a realization of an execution on an unstable device. 143

6.4 The plot shows improved accuracy achieved using Bayesian optimization in presence of time-varying noise. Mitigation using average parameters deteriorates accuracy. The number of samples used in the Bayesian estimation process is held constant at 10^4. . 143

6.5 The 4-qubit quantum circuit used for the simulation experiment. Each qubit is assumed to have a different, independent SPAM error process. Each Hadamard gate is similarly assumed to have a different, independent gate error process. 145

6.6 The figure displays the noisy basis circuits for the linear combination step (the first step) in PEC, where each figure represents the noisy operations $\mathcal{E}_{\mathrm{x}} \circ \mathcal{P}_{\sigma,\sigma'} \circ \mathcal{G}$ and $G = H \otimes H$ is the desired, ideal operation. Note that different basis set choices are possible, depending on the hardware. 146

6.7 We model the degradation of a non-stationary Pauli noise channel by assuming the coherence time steadily decreases over time. Using a non-stationary Dirichlet distribution, we model the joint distribution of coefficients for a two-qubit circuit, which fluctuate as coherence times deteriorate. The y-axis represents the degree of non-stationarity, and the x-axis shows four time-periods. The Hellinger distance between the Dirichlet distributions at time τ=0 and a later time is a measure of non-stationarity, increasing from 0 to 57%. This model is based on transmon platforms and is used as an experimental setup for our simulation experiments. 147

6.8 This graph compares adaptive and non-adaptive PEC for the $H \otimes H$ gate under time-varying noise. The y-axis shows the probability of observing a basis state. The black bars represent the ideal histogram for the test input, while red/orange bars are non-adaptive PEC results. Adaptive PEC improves accuracy. For example, the $|00\rangle$ probability increases from 57% to 72%. 148
6.9 This graph compares adaptive and non-adaptive PEC implementations over four time-periods with time-varying noise, using Hellinger distance to measure the difference from the ideal distribution. The black and orange bars represent the observed distributions with non-adaptive and adaptive PEC, respectively. Adaptive PEC significantly outperforms non-adaptive PEC in reducing Hellinger distance to 1.1%, and 3.2% compared to 7%, and 15% for non-adaptive PEC across time-periods. Adaptive PEC uses adaptive estimation of noise super-operators to improve accuracy compared to non-adaptive PEC. Time-varying noise underscores the need for adaptive PEC. . . . 149
6.10 The likelihood function used in the Bayesian inference procedure for the second period. 149
6.11 The cost function used in the Bayesian inference procedure for the second period. . . 150

7.1 Non-stationary distribution functions of the state preparation and measurement (SPAM) fidelity for qubit 2 on ibm_kolkata superconducting device collected on Jan 15, 2024. 170
7.2 A 5-qubit implementation of the Bernstein-Vazirani algorithm with secret bit-string r. 171
7.3 Circuit diagram for implementing one of the 16 noisy super-operators for CNOT noise mitigation using PEC. The initial noisy CNOT gate is followed up by $\mathbb{Y}$ and $\mathbb{Z}$ gates on the control and target qubits respectively. 171
7.4 One of the 512 noisy basis circuits for mitigation using PEC in the 5-qubit implementation of the Bernstein-Vazirani circuit. The initial noisy CNOT gate is followed up with $\mathbb{Y}$ and $\mathbb{Z}$ gates on the control and target qubits respectively. The readout lines for qubit $0, 1, 2, 3, 4$, and 5 are subjected to the Pauli gates $\mathbb{X}, \mathbb{I}, \mathbb{X}, \mathbb{X}, \mathbb{I}$ respetively, prior to measurement, in this specific noisy basis circuit for PEC. 172
7.5 Comparison of the adaptive PEC algorithm with a non-adaptive approach for a density matrix simulation that implements a noisy 5-qubit circuit for solving the Bernstein Vazirani problem. (a) Average accuracy across the 5-qubits and (b) Average stability across the 5-qubits. 172
7.6 Qubit layout of the 27-qubit superconducting device ibm_kolkata. Circles represent superconducting transmons and lines indicate possible gate operations between sites. 173

7.7 This figure depicts the non-stationary noise on the experimental device ibm_kolkata. In plot (a), the blue line represents the depolarizing parameter for the target qubit, while the black line denotes the depolarizing parameter for the control qubit in the CNOT gate. Plot (b) illustrates five lines, each indicating the SPAM fidelity for the register elements. The x-axis corresponds to intra-calibration timestamps for January 15. The shaded regions denote the time-varying standard deviations. 173

7.8 These plots depict the non-stationary nature of the quasi-probability distribution. To maintain clarity, we opted not to plot all 512 bins of the histogram in a single plot. In (a), we observe the time-varying weight $\mathbb{Q}_0$ for the raw Bernstein-Vazirani circuit which carries the most substantial weight, accounting for almost 80% of the distribution. In (b), we show the values of the quasi-probability bins for the subsequent 10 basis circuits, which are more than 10 times lower in magnitude compared to the first circuit. The non-stationary nature of the quasi-probability distribution becomes crucial given the lengthy data collection process required for PEC mitigation because the noise estimation becomes inaccurate in these time-frames. Our experiment took approximately 2 hours for each dataset comprising 512 circuits. 174

7.9 The figure comprises two plots, each depicting four graphs: (1) "No mitigation" presents raw Bernstein-Vazirani metrics without error mitigation, (2) "ROEM" shows metrics after readout error mitigation, with constant SPAM noise parameters, (3) "Non-adaptive PEC" displays metrics for PEC, and (4) "Adaptive PEC" exhibits metrics for PEC with adaptive noise mitigation. The x-axis denotes intra-calibration time-stamps (UTC) for Jan 15. Plot (a) illustrates the time-varying accuracy metric from Eqn. 7.16. It demonstrates the adaptive PEC's 42% accuracy improvement over non-adaptive PEC. Plot (b) shows the time-varying stability metric from Eqn. 7.19. It illustrates the adaptive PEC's 60% stability enhancement compared to non-adaptive PEC. These plots underscore the significant impact of non-stationary noise on PEC resilience. Due to PEC's lengthy completion time (a couple of hours), adaptive methods are able to handle non-stationary noise conditions better. 175

7.10 Non-stationary distribution function of SPAM fidelity for qubit 0 of ibm_kolkata as observed on Jan 15, 2024. 176

7.11 Non-stationary distribution function of SPAM fidelity for qubit 1 of ibm_kolkata as observed on Jan 15, 2024. 177

7.12 Time-varying density for probability of observing '00' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024. 178

7.13 Time-varying density for probability of observing '01' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024. 179
7.14 Time-varying density for probability of observing '10' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024. 180
7.15 Time-varying density for probability of observing '11' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024. 181

List of abbreviations

NSN	Non-stationary noise
PEC	probabilistic error cancellation
BC	Bhattacharyya coefficient
BV	Bernstein-Vazirani
SPAM	State preparation and measurement
Tr	Trace
WSS	Wide-sense stationary
MCMC	Markov chain Monte-Carlo

Notations

$\|\psi\rangle$	a pure quantum state
$\otimes$	tensor operator
Tr$(\cdot)$, Pr$(\cdot)$	Trace and probability operators respectively
ρ	density matrix
Ω, λ	Hamiltonian and its eigenvalue
T_1, T_2	Qubit relaxation and dephasing time respectively
N_s	Number of shots per circuit execution
L	Number of times a circuit is executed for collecting statistics
n	quantum register size
$\|v\rangle, \|v_{n-1} \cdots v_0\rangle, v_i, v$	$\|v\rangle$ is the short form for $\|v_{n-1} \cdots v_0\rangle$. The latter denotes an n-bit string with each $v_i \in \{0, 1\}$. Lastly, v denotes the decimal form for the bit-string e.g. if $\|v\rangle = \|0011\rangle$, then $v = 3$.
$\mathcal{U}$,$\tilde{\mathcal{U}}$	a unitary matrix and its noisy version (could be non-unitary)
$R_Y(\theta)$	operator for rotation by an angle (θ) about the Y-axis on the Bloch sphere
$\mathbb{I}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}$	2x2 identity and the Pauli X, Y, Z matrices

$\mathbb{H}$, CNOT, SX, RZ	Various quantum gates (Hadamard gate, Controlled-NOT, Square root of NOT and Unitary gate for rotation about Z-axis
T_D	delay gate duration
$\tilde{\mathbb{H}}$	noisy Hadamard gate
$\mathcal{I}(X, Y)$	Mutual Information between two random variables X and Y
$\mathcal{H}(X)$	Entropy of X
$\mathcal{O}$	A quantum observable
$\mathcal{E}_{\mathrm{x}}, \mathcal{E}_{\mathrm{x}_i}(\rho)$	Multi-qubit quantum noise channel parameterized by the vector x; $\mathcal{E}_{\mathrm{x}_i}(\rho)$ is a single-qubit quantum noise channel
M_k	Kraus operators characterizing the error channel $\mathcal{E}(\cdot)$
$F_{\mathrm{SPAM}}, \epsilon_{\mathrm{SPAM}}$	SPAM Fidelity and SPAM error respectively
F_{G}, e_G	Gate Fidelity and error per Clifford gate respectively
T_G	CNOT Gate Length
τ_G	Duty Cycle
$\mathrm{Beta}(\alpha, \beta)$	Beta function ($= \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$)
$\Gamma(z)$	Gamma function ($= \int_0^\infty t^{z-1}\exp^{-t}dt$)
$\mathbb{R}, \mathbb{C}$	Set of real and complex numbers respectively
$P_2(\mathbb{R})$	Set of real polynomials of degree not more than 2
x, x_i	$x = (\mathrm{x}_1, \cdots, \mathrm{x}_d)$ is a vector of parameters characterizing the noise of the quantum circuit
X, X_i	X is a vector of random variables (stochastic device parameters) corresponding to x above; X_i is a scalar random variable

$f_X(\mathrm{x}; t)$	multi-variate continuous probability distribution function for x; same as $f_X(\mathrm{x}; t)$
$F_X(\mathrm{x}; t)$	multi-variate continuous cumulative distribution function for x; same as $F_X(\mathrm{x}; t)$
$\eta(X, Y)$	Normalized Mutual Information between two random variables X and Y
F_{A}	Addressability
BC	Bhattacharyya coefficient
$H_X(t_1, t_2)$	Hellinger distance between the multi-variate joint distributions of X at times t_1 and t_2; if $t_1 = 0$ then we sometimes simply write it as $H_X(t)$
$H_{\mathrm{avg}}(t_1, t_2)$	average of the Hellinger distances over the d univariate marginal distributions $\left(= \frac{1}{d} \sum_{k=1}^{d} H_{X_k}(t_1, t_2)\right)$
$H_{\mathrm{normalized}}(t_1, t_2)$	Hellinger distance normalized with respect to the dimension $(= \sqrt{1 - BC^{1/d}})$
$\Theta(\cdot)$	copula function
$f^{\mathrm{BV}}(\cdot)$	Oracle function in the Bernstein-Vazirani problem
r	secret n-bit string in the Bernstein-Vazirani problem
U_f	Unitary for the Oracle function in the Bernstein-Vazirani problem
P_a, P_C, P_T	Pauli gates
$\mathcal{Q}$	quasi probability distribution
$\{\eta_k\}$	generally refers to a PEC linear combination coefficient
$\tilde{\mathcal{B}}_i$	noisy basis superoperators in PEC

Chapter 1

Introduction

1.1 Background

1.1.1 Promise of quantum computing

Error-resilient quantum computing holds great promise, offering significant advancements over conventional computing. Once realized, it is expected to super-polynomially reduce execution time, energy consumption, and memory storage needs compared to conventional state-of-the-art computers [1]. The potential impact of error-resilient quantum computing includes revolutionizing scientific applications such as simulating many-body quantum systems [2], solving large-scale optimization problems [3], efficiently sampling high-dimensional probability distributions [4], factorizing large integers, and enhancing the security of communication networks [5]. Consequently, this technology is expected to be disruptive to sectors such as telecommunications, cyber-security, pharmaceuticals, logistics, supply chain management, artificial intelligence, and materials science [6].

1.1.2 Quantum computing vs classical computing

Classical mechanics, rooted in the laws of Newtonian physics, has served as a successful framework for understanding the macroscopic world for centuries. However, when examining the behavior of particles at the atomic and subatomic scales, classical mechanics began to exhibit limitations and inconsistencies. For instance, classical mechanics predicted absurd outcomes, like suggesting that a blackbody emits an infinite amount of energy across all wavelengths.

As scientists delved deeper into the microscopic realm, counter-intuitive phenomena such as wave-particle duality, quantized energy levels, and non-locality emerged, challenging the classical paradigm.

These challenges necessitated the development of quantum mechanics, which offered a novel and revolutionary approach to describe the behavior of particles at the quantum level [7].

Quantum mechanics introduced probabilistic interpretations, superposition states, and entanglement, providing a more accurate and comprehensive understanding of the intricate workings of nature [8]. For example, consider a particle moving in one dimension under the influence of a conservative force, such as a harmonic oscillator. In classical mechanics, we can describe the particle's motion using Newton's second law and the equation of motion for a harmonic oscillator, which yields a sinusoidal trajectory and continuous energy levels. In contrast, in quantum mechanics, we describe the particle using the Schrödinger equation for a harmonic oscillator, which results in quantized energy levels and wave functions corresponding to discrete energy states. However, as the value of the Planck's constant (h), representing the fundamental scale of quantum mechanics, approaches zero ($h \to 0$), the quantum system converges towards the classical limit. In this limit, the quantized energy levels of the quantum harmonic oscillator become densely spaced and form a continuous energy spectrum, matching the classical behavior. The quantum wave function also converges to the classical trajectory, and the classical and quantum results become indistinguishable.

This convergence phenomenon is known as the correspondence principle [9], where classical mechanics emerges as the limiting case of quantum mechanics at large scales or when the quantum effects become negligible. Understanding the system's scale is crucial in selecting the appropriate framework for a given physical problem, with classical mechanics suitable for macroscopic objects with well-defined trajectories and negligible quantum effects, while quantum mechanics is employed for microscopic particles, providing a more accurate description of phenomena like wave-particle duality, quantization of energy levels, and quantum entanglement.

Quantum mechanics is built upon several fundamental postulates [10, 11] that were formulated to address experimental observations in the early 20th century. These postulates provide the framework for understanding the behavior of quantum systems. The first postulate states that every quantum system is associated with a complex Hilbert space [12]. The quantum state of a system is described by a density operator [13, 12], often denoted by ρ, which belongs to the set of density operators defined on the Hilbert space. The second postulate deals with measurements [13]. When a measurement is performed on a quantum system, it can have random outcomes with finite probabilities. The third postulate connects the quantum state with the measurement outcomes. When a measurement M is performed on a quantum system in the state ρ, then the observed outcome is a realization of the random variable $M\rho$ [11]. The fourth postulate addresses composite quantum systems. When we have two quantum systems, each associated with its own Hilbert space, say $\mathcal{H}_1$ and $\mathcal{H}_2$, the combined

Hilbert space of the composite system is given by the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. This tensor product construction allows us to represent the joint states of the individual systems [9].

These fundamental postulates provide the tools to analyze quantum systems. The concepts of quantum state, measurement, and composite systems are the key building blocks for understanding the intriguing phenomena that occur in the quantum world.

Classical computing is the conventional form of computing that relies on classical bits, represented by the binary numbers 0 and 1. Classical computers crucially rely on components like transistors, which function based on quantum mechanical principles. However, despite the quantum nature of the transistors, the interactions between these components within a classical computer follow a classical framework. This distinction can sometimes lead to confusion, as it may seem unsatisfactory to say that classical computers, which are built using components based on quantum principles, operate according to classical laws. But while specific underlying components leverage quantum phenomena, the interactions between these components, and the analysis of the data produced by them, are modeled satisfactorily using classical physics and classical information theory.

Analogously, the link between quantum computing [14] and quantum mechanics is fundamental, as the fundamental interaction between the components of quantum computers depend on quantum mechanical phenomenon [11] such as superposition and entanglement. Quantum computations use quantum bits or qubits [15] as the basic unit of information. Superposition enables qubits to exist in multiple states simultaneously, allowing quantum computers to perform parallel calculations. Entanglement [16] creates strong correlations between qubits, even when they are physically separated, potentially leading to increased computational power. Different qubit technologies, such as superconducting qubits, trapped ions, and topological qubits, each utilize distinct quantum phenomena. Researchers continuously draw upon quantum mechanical principles to optimize performance, addressing challenges like quantum decoherence [17, 18, 19, 20] resulting from interactions with the external environment.

Classical computing remains the practical and efficient choice for many computing needs. In fact, the computing power of classical computers has been doubling every two years, as per Moore's law. But still classical computing faces limitations with computationally challenging problems that grow exponentially with problem size such as unstructured large-scale optimization, factorization of large integers, and simulation of many-body systems. The quantum equivalent of Moore's law [11] states that adding just one perfect qubit to a quantum computer doubles its computational capability. Therefore, to match the progress of classical computers, a single error-resilient qubit needs to be integrated into quantum computers every two years.

However, quantum algorithms [10] are more efficient for tackling only a sub-set of classically computationally challenging problems (not universally). Example of problems that have an efficient quantum algorithm include prime factorization and discrete logarithm, both of which were developed by Shor. An efficient algorithm [21] operates within a time frame that corresponds to a polynomial function of the problem size, whereas an inefficient algorithm takes time that corresponds to a super-polynomial function of the problem size.

The study of algorithm efficiency [21] is a fundamental aspect of complexity theory [22], a branch of computer science. Problems are categorized based on their resource requirements, such as time and memory. For instance, problems solvable in polynomial time by classical computers fall into the class P, whereas those with solutions verifiable in polynomial time belong to class NP. While it's evident that P is a subset of NP, the question of whether there are problems in NP not in P remains unresolved.

In the realm of quantum computing, problems solvable by quantum algorithms within polynomial time (with bounded error probability) are classified as belonging to class BQP [11]. It has not been formally established whether BQP contains P. So, we are not certain that quantum computing contains classical computing as a special case, but evidence supports this assertion.

BQP of course contains QP which represents problems that a quantum computer can solve with a 100% probability of success in polynomial time. Examples that belong to the BQP class include Deutsch-Jozsa, Bernstein-Vazirani, and Simon's algorithm.

Why can we not simulate quantum computation using classical computers? The reason lies in the exponential space and time complexity involved in storing quantum gates as classical matrices and tracking entangled qubits after logical operations. For instance, even a system with just 500 atoms would need 2^{500} complex coefficients for perfect description. Attempting computations with such requirements would overwhelm classical computers. However, quantum computers excel in simulating such scenarios by storing, representing, and evolving states as native quantum states on qubit registers, bypassing the need for managing 2^{500} complex floating-point numbers with limited precision

Note that, even if we restricted ourselves to a small scale register, then also a perfect quantum computer can never be built using classical computers because quantum measurement cannot be perfectly simulated as there is no perfect random number generator.

In classical computing, logic gates are basic building blocks that manipulate classical bits (0s and 1s) to perform logical operations. Two examples of classical logic gates are the AND gate and the NOT gate. The AND gate takes two input bits, and its output is 1 (true) only when both input bits are 1;

otherwise, the output is 0 (false). The NOT gate takes a single input bit and produces the opposite value as output.

Input 1	Input 2	AND Gate Output
0	0	0
0	1	0
1	0	0
1	1	1

Input	NOT Gate Output
0	1
1	0

In quantum computing, instead of bits, we have qubits, which are represented as vectors in a complex two-dimensional Hilbert space. The basis vectors spanning this space are commonly expressed in three ways: the $\mathbb{Z}$-Basis (also called computational basis or standard basis), the $\mathbb{X}$-basis, and the $\mathbb{Y}$-basis. Each of these bases is made up of orthogonal vectors in two dimensions, ideal for representing a two-level quantum system in a two-dimensional vector space.

The need for these bases becomes clear when studying the underlying physics of quantum systems. Take, for example, the spin of an electron in a magnetic field. The electron's spin resembles a minuscule magnetic moment. Its orientation relative to the magnetic field influences the electron's energy.

For the $\mathbb{Z}$-Basis, consider a magnetic field applied vertically. An electron's spin might align with this field, represented by the lower energy state $|0\rangle$, or it could oppose the field, corresponding to the higher energy state $|1\rangle$. This basis provides an intuitive way to think about qubits, likening the lower and higher energy states to the classical binary values of 0 and 1 respectively. A qubit's state can be in a superposition of both $|0\rangle$ and $|1\rangle$, expressed as $c_1 |0\rangle + c_2 |1\rangle$, where c_1 and c_2 are complex amplitudes satisfying:

$$|c_1|^2 + |c_2|^2 = 1. \tag{1.1}$$

However, if we change our perspective and measure the spin horizontally, along the $\mathbb{X}$-axis, the electron's spin might point left or right. These orientations, when related back to the $\mathbb{Z}$-Basis, are actually superpositions of the $|0\rangle$ and $|1\rangle$ states. These superpositions,

$$|+\rangle = \frac{|0\rangle}{\sqrt{2}} + \frac{|1\rangle}{\sqrt{2}}, \tag{1.2}$$

and

$$|-\rangle = \frac{|0\rangle}{\sqrt{2}} - \frac{|1\rangle}{\sqrt{2}}, \tag{1.3}$$

define the $\mathbb{X}$-basis.

The $\mathbb{Y}$-basis offers yet another viewpoint. Measuring perpendicular to both the $\mathbb{X}$ and $\mathbb{Z}$ axes gives states that are complex superpositions of the $\mathbb{Z}$-Basis:

$$|i\rangle = \frac{|0\rangle}{\sqrt{2}} + i\frac{|1\rangle}{\sqrt{2}}, \tag{1.4}$$

and

$$|-i\rangle = \frac{|0\rangle}{\sqrt{2}} - i\frac{|1\rangle}{\sqrt{2}}, \tag{1.5}$$

Quantum operations can be visualized as rotations around these different basis axes, offering valuable insights for designing quantum algorithms. Furthermore, using different bases for measurements can be instrumental in pinpointing various types of errors in quantum systems.

Quantum gates, represented by unitary matrices, manipulate qubits, with essential examples including the Pauli gates ($\mathbb{I}$, $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{Z}$), the Hadamard gate ($\mathbb{H}$), and the two-qubit entangling CNOT gate ($\mathbb{U}_{\text{CNOT}}$). The Pauli-$\mathbb{X}$ gate acts as the quantum analog of the NOT gate, flipping the qubit's state between $|0\rangle$ and $|1\rangle$. The Pauli-$\mathbb{Y}$ gate transforms $|0\rangle$ to $i\,|1\rangle$ and $|1\rangle$ to $-i\,|0\rangle$. The Pauli-$\mathbb{Z}$ gate introduces a relative phase shift between the basis states of a qubit (it leaves $|0\rangle$ unchanged and flips the sign of $|1\rangle$). The Hadamard gate puts a qubit in an equal superposition of $|0\rangle$ and $|1\rangle$. The CNOT gate is a two-qubit gate that flips the target qubit if and only if the control qubit is in state $|1\rangle$.

Pauli-$\mathbb{X}$ Gate

Input	Output
$\|0\rangle$	$\|1\rangle$
$\|1\rangle$	$\|0\rangle$

Pauli-$\mathbb{Y}$ Gate

Input	Output
$\|0\rangle$	$i\,\|1\rangle$
$\|1\rangle$	$-i\,\|0\rangle$

Pauli-$\mathbb{Z}$ Gate

Input	Output
$\|0\rangle$	$\|0\rangle$
$\|1\rangle$	-$\|1\rangle$

Hadamard Gate

Input	Output
$\|0\rangle$	$(\|0\rangle + \|1\rangle)/\sqrt{2}$
$\|1\rangle$	$(\|0\rangle - \|1\rangle)/\sqrt{2}$

CNOT gate

Control qubit	Target qubit	Output (state of target qubit)
$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
$\|1\rangle$	$\|1\rangle$	$\|0\rangle$

In classical circuits, information flows through the movement of electrons from one transistor to another in a well-defined spatial layout. Classical logic gates, such as AND, OR, and NOT, manipulate classical bits (0 or 1) and perform logical operations. In contrast, quantum circuits

process information using qubits. Quantum circuits evolve the quantum state in situ, which modify the quantum amplitudes and phases of the state of the quantum register. Quantum gates, like Pauli-$\mathbb{X}$, CNOT, and Hadamard, implement quantum algorithms by performing quantum operations on the qubits.

Designing quantum algorithms [10] is significantly more challenging than classical algorithms [21] for several reasons. Firstly, quantum computing requires a departure from classical intuition, as quantum phenomena behave differently from classical physics. For instance, computer scientists experienced in conventional parallel programming understand the challenges associated with designing algorithms that can effectively harness GPU parallelism. They can thus empathize with the complexity of leveraging computing power through superposition, a form of parallel computing utilized in quantum systems. However, phenomena such as entanglement and quantum interference present unique opportunities that lack analogues in classical algorithm development. Secondly, to demonstrate the utility of a quantum algorithm, it must be more efficient than the best-known classical algorithm for a specific problem. The competitiveness of the latter introduces a moving target for quantum algorithm developers, where the best classical algorithms keep evolving, demanding continuous advancements in quantum algorithms to maintain claims of utility. Lastly, mapping a real-world (often classical) use-case into a quantum representation is a non-trivial task.

A quantum program is a series of instructions that can be executed by a quantum device in a specific sequence to perform a specific task. These instructions are typically written in a high-level programming language like Qiskit [23], designed to be readable and writable by humans.

A classical computer architecture consists of several key components that work together to perform various computational tasks. At its core, a classical computer contains a central processing unit (CPU). The CPU executes instructions, and coordinates data movement. It is supported by memory units, including random-access memory (RAM) and cache memory, where data and instructions are temporarily stored for faster access. The architecture also includes input and output (I/O) devices, such as keyboards, mice, monitors, and storage devices like hard drives or solid-state drives. These allow users to interact with the computer and store data for future use. The CPU communicates with other components via buses, which are pathways that transfer data and control signals between different parts of the computer. The system clock generates regular pulses that synchronize the activities of various components, ensuring smooth coordination of operations. Moreover, a classical computer architecture often involves a graphics processing unit (GPU) dedicated to handling graphics-intensive tasks, such as rendering images and videos.

The development and elaboration of a quantum computer architecture [24, 25, 26] are still in their early stages, mainly because we have not achieved fault-tolerant quantum computing yet. The full stack architecture will need focus beyond physical layer and must include error correction, feedback stabilization, hardware-aware compilation, logical level compilation, circuit optimization, application layer, and user interface. While progress is being made towards this vision, it remains a distant goal.

1.1.3 The problem of noise

The behavior of an ideal quantum computer can be modeled as follows. An n-qubit register spans a complex Hilbert Space denoted by $(\mathbb{C}^2)^{\otimes n}$. The initial state of the register can be represented as a tensor:

$$|\psi\rangle = |0\rangle^{\otimes n}. \tag{1.6}$$

A logical operation on the register state (also called a quantum gate) can be represented by a linear, unitary operation U:

$$|\psi\rangle \rightarrow U\,|\psi\rangle. \tag{1.7}$$

A measurement reads out a n-bit string v in the computational basis:

$$Pr(v) = |\,\langle v|\psi\rangle\,|^2, \tag{1.8}$$

where $v \in \{0,1\}^{\otimes n}$

Unlike modern classical computers with extremely low failure rates (e.g., 10^{-17} or less), superconducting quantum computers exhibit higher gate-level failure rates (nearly 0.01) [27, 28]. Thus, it is imperative to investigate the causes of noise and mitigate and correct them so that quantum computers can provide correct results.

Noise, in the context of this dissertation, refers to deviations from the ideal description of a quantum computer. Practical efforts to build quantum computers introduce noise, which affects technologies like superconducting qubits, trapped ions, and silicon quantum dots. Our focus is primarily on superconducting quantum computers [29, 30]. The underlying noisy processes that impact such a computer can be classified into three groups: noise affecting the the quantum register (such as leakage [31], undesired coupling, decay processes, non-uniformity, and cross-talk [32]), noise affecting the quantum operations (such as pulse distortion, attenuation, drift, and mis-calibration), and noise in the thermodynamic isolation system (due to issues with dilution refrigerators, vacuum chambers, shields, and vibration suppression mechanisms [33]).

Quantum computers today are referred to as existing in the NISQ era [34], which stands for noisy intermediate-scale quantum. The noise threshold for NISQ is defined by a single-qubit error rate being worse than 10^{-4}. The intermediate-scale label is often associated with having fewer than a hundred thousand qubits. Computing done with NISQ devices is called NISQ computing. The bare minimum requirements for NISQ computing includes: quantum registers for storing data, quantum gates to execute logical operations on the registers, and a measurement interface for extracting the computation outcomes. The field has witnessed rapid advancements, with NISQ devices now operating as systems with hundreds of interacting qubits. Remarkably, the field is already witnessing a transition towards a phase where NISQ devices are performing scientific computations at a scale that rivals classical supercomputers in terms of computational power [35]. These experiments apply error mitigation techniques to the outcomes of the noisy computations performed on the NISQ devices.

Note that error mitigation and error correction are distinct strategies to tackle challenges arising from noise: mitigation employs statistical techniques focused on minimizing noise effects rather than eliminating errors entirely, such as zero noise extrapolation and probabilistic error cancellation, while error correction seeks to actively detect and rectify errors during computations with an objective of fault tolerance, but requires the device noise to be below a threshold, which has not been achieved yet. One example of the unique challenges in the area of quantum error-correction is the no-cloning theorem states that it's impossible to create an exact duplicate of an unknown quantum state. This makes it challenging to incorporate redundancy into quantum computing systems to protect against information corruption.

1.2 Research focus

The fact that noise exhibits non-stationarity, underscores the core motivation of this dissertation. A quantum noise channel is often used to model how quantum information becomes distorted during its passage through a physical system. It describes how interactions with the environment can modify the quantum state of a system. Examples include depolarizing channel, Pauli noise channel, amplitude damping channel, phase damping channel, and the SPAM noise channel.

The term SPAM denotes state preparation and measurement. The statistics of the SPAM error channel are commonly quantified using the SPAM fidelity, a metric that evaluates a device's ability to prepare and measure a qubit. Specifically, SPAM fidelity quantifies the likelihood that the device readies the qubit (or a set of qubits) in the desired state and subsequently measures it in that same state. Fig. 1.3 - 1.5 depicts the time-varying noise densities of SPAM fidelity.

Concurrent experimental studies [17, 18, 19, 36, 37] support our concern that the assumption of fixed and invariant parameters for quantum processors is flawed. In fact, these studies show that the noise parameters can exhibit time variations of up to 50% of their mean value within an hour. Spatially varying noise in quantum devices have also been extensively studied, including the role of circuit geometry [38, 39] and cross-talk between neighbouring qubits [40, 41]. Unlike temporal variations in noise, the effects of which are magnified by the complexity of a quantum circuit, the spatial variations, such as seen in Fig. 1.2, are dependent on the geometry and scale of the circuit implementation.

To understand the impact of such non-stationary noise on program outcomes, consider Fig. 1.1 which depicts the time-varying histogram obtained from IBM's ibm_mumbai device for the Bernstein-Vazirani circuit. It is evident that the associated error bars on a particular day do not provide insights into results from a different date, highlighting the problem of reproducibility of results in quantum computing today.

The causes of non-stationary noise are not fully understood but are believed to stem from TLS defects in transmon registers which might be arising from deviations in crystalline order. The current consensus attributes these defects to the presence of certain oxides on the superconductors' surface [42, 36]. Thus, static quantum channel models do not accurately capture the dynamics in realistic quantum computations, particularly in superconducting qubits. Cosmic rays [43, 44] also contribute by ionizing the substrate upon impact, leading to the emission of high-energy phonons, which in turn triggers a burst of quasi-particles. These quasi-particles disrupt qubit coherence across the device. It has been shown that quantum computers can experience catastrophic errors in multi-qubit registers approximately every 10 seconds due to cosmic rays originating from outer space [44]. Studies that address non-stationary noise in superconducting quantum computers include investigations on output reproducibility [45], noise modeling [46], tracking the non-stationary profile of quantum noise [47], and quantum error mitigation using continuous control [48]. Non-stationary quantum channel models have been proposed [46, 49, 50, 51] that use stochastic processes. Our dissertation focuses on understanding the effect of non-stationary noise on program outcomes, as well as devising strategies to address it. Specifically, we model the noise channel as a random variable and implement adaptive methods to manage it.

Before we can begin, we need to fix the precise language for performance assessment. This task is not trivial due to the complexity inherent in quantum technology, which both distinguishes it from classical computing and hinders its rigorous checking [52, 53]. Challenges include the inherent randomness in quantum measurements, error accumulation without clear source attribution, the curse of dimensionality, and the inability to step-through program execution in quantum circuits.

In fact, the performance evaluation [54] of noisy quantum computations is a vast topic that is crucial for several additional reasons, apart from our motivation of studying the impact of non-stationary noise. Firstly, as quantum computing is still in its early stages [55, 56, 57], understanding the sources of errors and noise is vital. Through rigorous evaluation, researchers can model [58], identify, model and quantify sources of noise, such as decoherence, gate errors, and readout errors. Secondly, this understanding is essential for developing error mitigation techniques [59]. Thirdly, reproducibility of results is critical, and rigorous performance evaluation ensures experiments can be replicated by other researchers, contributing to the validation and verification of quantum algorithms. An additional challenge is the diverse range of terms encountered in quantum computing today which can blur distinctions between them, making it challenging to appreciate their nuanced differences. Examples include verification [60, 61, 62] (ensuring correct transpilation), validation [63, 64] (validating correctness of output or the quantum nature of a device), benchmarking [65, 66, 67] (assigning a performance measure to a processor), accreditation [68], and certification [54]. Thus out first task is to precisely define computational accuracy, result reproducibility, device reliability, and observable stability in the presence of non-stationary noise.

Our next objective is to experimentally assess hardware reliability, with a particular emphasis on analyzing spatial and temporal variations in noise statistics. IBM [55] has introduced a range of processors in recent years, each with an expanding register size. These include the Canary processors with 2-16 qubits, Falcon processors with 27 qubits, Egret processors with 33 qubits, Hummingbird processors with 65 qubits, Eagle processors with 127 qubits, and Osprey processors with 433 qubits. Quantifying the spatial and temporal reliability of these quantum computers is crucial to understand system-wide performance changes over time. This evaluation should encompass both component-level metrics, such as individual gates and qubits, and composite-level metrics, such as circuits, to assess the degree of non-stationarity in noise and its implications on program outcomes. Holistically measuring reliability at the circuit level is essential, as examining thousands or millions of qubits and gates may not provide conclusive insights at the application level [69].

Our third objective in this dissertation is to establish stability bounds for error-mitigated outcomes affected. We aim to determine the upper and lower bounds for our performance evaluation metrics. Our inquiries include determining the minimum sample size necessary to ensure histogram reproducibility with a confidence level of $1 - \delta$, bounding outcome stability based on the variation in time-varying noise densities, and establishing reliability bounds to achieve stable outcomes.

Numerous studies on noise modeling in quantum computing systems have highlighted the challenges associated with noise estimation [70, 71]. A natural question arises regarding how can we effectively counteract the detrimental impacts of non-stationary noise using adaptive algorithms? In the final

chapter of our dissertation, we consolidate the various elements of our investigation in the context of adaptive probabilistic error cancellation (PEC) [72]. Our focus is on enhancing the stability of PEC outcomes, using a Bayesian [73, 74, 75, 76] updating of the quasi-probability distributions, in the presence of non-stationary noise.

This research focuses on a limited scope. Firstly, the experiments exclusively uses the superconducting platforms provided by IBM. Other platforms such as trapped ion, neutral atom, photonic, or quantum dot are not considered. Secondly, our performance evaluation framework mainly concerns with the output measured in computational basis and how it is impacted by device noise. Thirdly, not all the superconducting devices provided by IBM have been characterized; only a sub-set of the devices, mainly ibm_kolkata, ibm_mumbai, ibm_washington, ibm_toronto, and ibm_yorktown, are used in this study. Lastly, we do not focus on the problem of optimal selection of a statistical model for a given noisy device. Instead, a generic error channel formalism is employed whenever possible. For verifying the theory, quantum noise channel models like the Pauli noise [77] channel is chosen, with a specific focus on parameter estimation. The research does not explore the question of identifying the best noise model for a given device.

For our research, we have made use of the daily characterization data stored on IBM's servers as-is. For data at time-scales of minutes and below, we collected the data ourselves and offer all the associated collection and preparation software. For the latter case, the data collection frequency was limited by network time lags and constraints in the qiskit software, such as the maximum number of circuits and shots allowed. These limitations have been gradually improving over time

1.3 Organization and notation

The document is structured as follows. Chapter 2 provides background on noise in quantum computing, quantum channel modeling, and experimental characterization of quantum decoherence, emphasizing the non-stationary statistics of noise. Chapter 3 establishes a systematic framework for assessing noisy quantum computer performance. Chapter 4 focuses on the testing of reliability. The evaluation encompasses both component-level metrics (such as individual gates and qubits), and composite-level metrics (such as circuits). Chapter 5 seeks to determine the bounds on the assessment framework developed in Chapter 3, using available noise characterization data. Specifically, it discusses how to bound outcome stability in terms of the distance between time-varying noise densities [78]. Chapter 6 explores methods to improve accuracy in the presence of non-stationary noise. Chapter 7 brings together the various concepts discussed till date in the context of adaptive probabilistic error cancellation. Chapter 8 provides concluding remarks.

Notations in this dissertation vary in meaning depending on font, although it should be clear from the context (see Table 2). This was needed because the work draws upon concepts from physics, information theory, computer science, and statistics, each of which has established conventions. $|\psi\rangle$ represents a pure quantum state, and ρ represents density matrices. The symbol $\otimes$ signifies a tensor operator, and $\mathrm{Tr}(\cdot)$ is an abbreviation for the trace operator. The system Hamiltonian is denoted as Ω, with its eigenvalues represented by λ (however ω signifies angular frequency). Quantum observables are typically denoted by $\hat{O}$, and the uppercase letter U typically stands for a unitary matrix. A single-qubit rotation by an angle θ on the Bloch sphere is often denoted as $R(\theta)$. However, the uppercase Greek letter $\Theta(\cdot)$ denotes the copula [79] function from statistics. $\mathcal{E}_{\mathrm{x}}(\cdot)$ represents a quantum noise channel, while Λ symbolizes the SPAM noise channel, a classical channel operating on probabilities. The canonical Pauli matrices are denoted as $\mathbb{X}, \mathbb{Y}, \mathbb{Z}, \mathbb{I}$. Note that we do not use the small Greek letter σ for Pauli matrices, reserving it for standard deviation instead. The Pearson correlation matrix is denoted by the capital Greek letter Σ. Note that the small z signifies a standard normal variable. The identity matrix is represented as $\mathbb{I}$, while $\mathcal{I}(X, Y)$ in calibrated font signifies the mutual information between random variables X and Y. The letter x typically signifies noise parameter(s), whereas an uppercase X corresponds to a specific realization of x. If x is not deterministic, then $f(\mathrm{x})$ denotes the probability distribution of x. This distribution can exhibit temporal fluctuations, denoted as $f(\mathrm{x}; t)$, with its cumulative distribution function indicated by $F(\mathrm{x}; t)$. Additionally, the curly capital $\mathcal{F}$ stands for Fisher Information, while f_s with the subscript s represents the data sampling frequency. The symbol Π_r with a subscript is reserved for the projector operator onto the eigenstate $|\lambda_r\rangle$ (however, when presented without a subscript, Π signifies the normal product operator). The measurements are conducted in the computational basis (or $\mathbb{Z}$ basis), resulting in qubits yielding classical bits. We employ the notation $b_i(t)$ to represent the observed classical bit value on qubit i at time t. The state of an n-qubit quantum register is denoted by $|v\rangle = |v_{n-1} \cdots v_0\rangle$, with n generally denoting the quantum register size. Upon measurement, this state yields an n-bit string, with each $v_i \in 0, 1$. We utilize the symbol W to denote the dataset consisting of collected bit-strings from repeated circuit executions. The total number of samples collected is typically denoted as L, where l denotes the l-th circuit execution (however, note that the curly $\mathcal{L}$ represents the likelihood function). The Hellinger distance between probability distributions is denoted as H, while the curly capital $\mathcal{H}$ is exclusively reserved for entropy. The symbol $\mathbb{H}$ represents the Hadamard gate. A quantum circuit is represented by the capital C, and while, the small c usually signifies the control qubit in a CNOT gate. Additionally, c serves as a constant in certain information theoretic results. The symbols α and β may assume different meanings depending on the context, referring either to quantum amplitudes for $|0\rangle$ and $|1\rangle$ or the parameters of the beta distribution, both of which find common usage in this dissertation. The small greek γ is used as

a proxy parameter encapsulating various device noise parameters, while the capital Γ is reserved for the gamma function. The symbol η typically represents a PEC linear combination coefficient. Capital D denotes a noisy quantum device, while the small letter d typically dimensionality. In the context of the Bernstein-Vazirani problem, r is used to denote the secret n-bit string, with the latter problem being extensively employed as an illustrative quantum circuit. The small letter s is used to denote the stability metric. Absolute time is typically represented by the small letter t, while time intervals are denoted as τ. For instance, τ_c signifies the circuit execution time, and τ_N denotes network delay. Time duration is represented as δt. However, the capital letter T is primarily reserved for parameters related to decoherence characterization, such as T_1 and T_2 (representing qubit relaxation and dephasing time, respectively). In some instances, T (without a subscript) is used to denote the target qubit in a CNOT gate, which is generally clear from the context. Finally, the non-standard abbreviations that have been used in this dissertation are: NSN (non-stationary noise), PEC (probabilistic error cancellation), BC (Bhattacharyya coefficient), BV (Bernstein-Vazirani), SPAM (State preparation and measurement) and WSS (Wide-sense stationary).

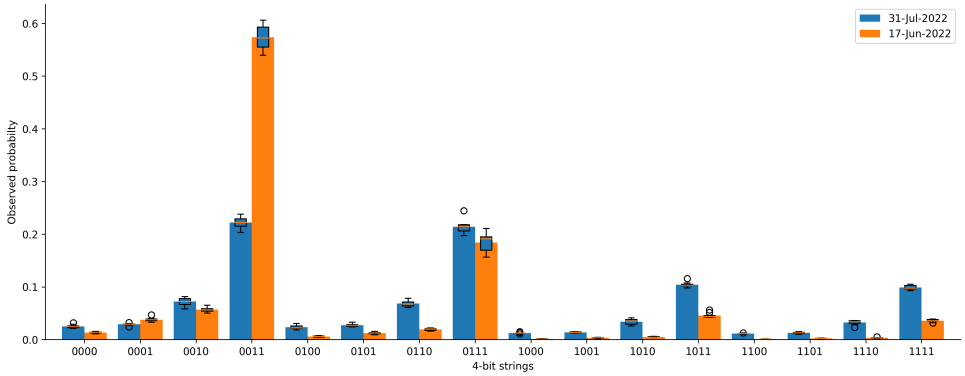


Figure 1.1: Impact of the non-stationary noise, leading to irreproducible outcomes. The results show histograms (after state preparation and measurement (SPAM) noise mitigation) upon executing the Bernstein-Vazirani circuit on ibm_mumbai.

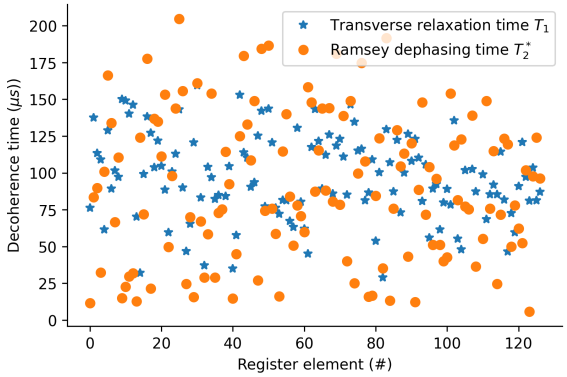


Figure 1.2: Evidence of spatial non-stationarity in the mean values for quibt decoherence times T_1 and T_2 for the 127-qubits of the ibm_washington device, generated on 14 Jan 2023 10:20 PM UTC.

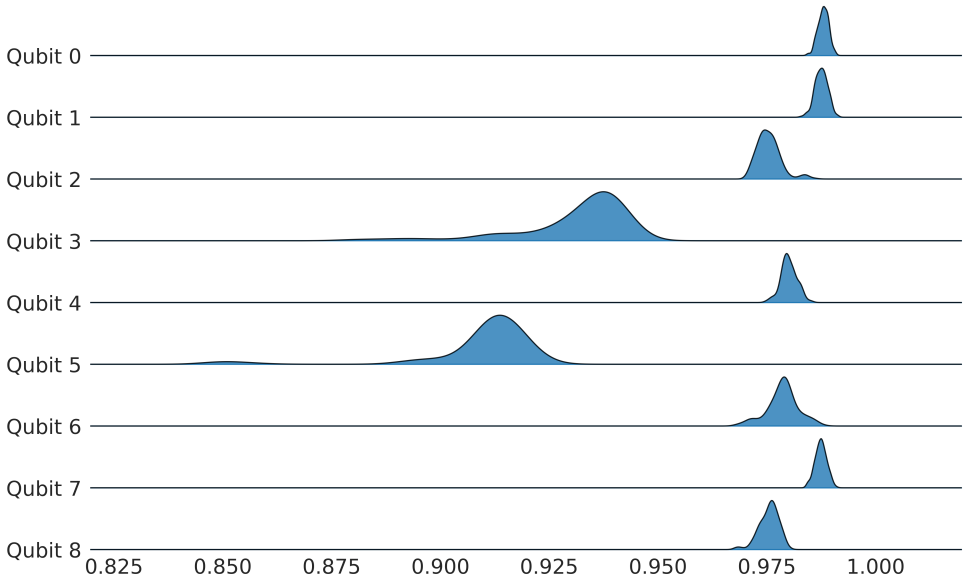


Figure 1.3: state preparation and measurement (SPAM) fidelity distributions on ibm_toronto for qubits $0-8$ as measured on 8 April 8 2021, between 8:00-10:00pm (UTC-05:00).

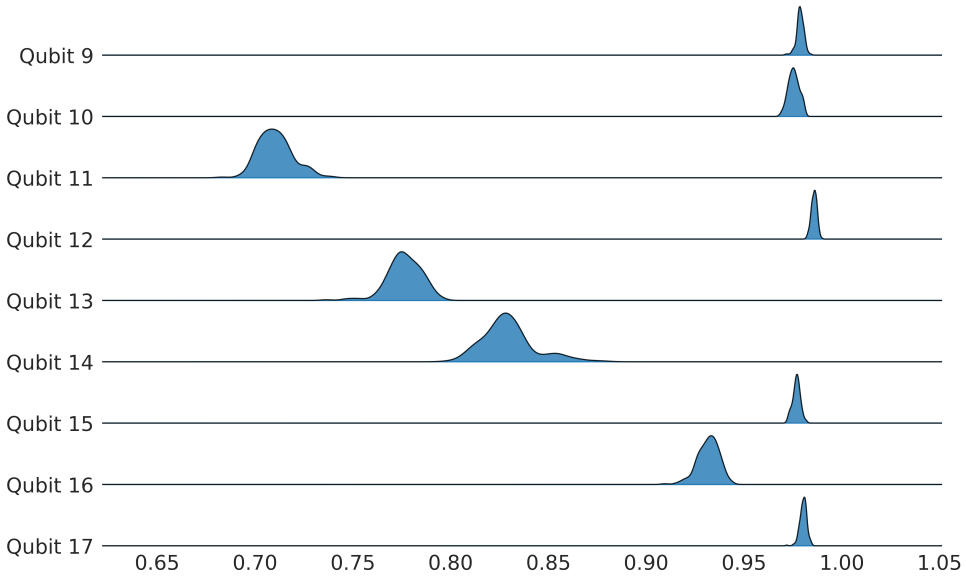


Figure 1.4: state preparation and measurement (SPAM) fidelity distributions on ibm_toronto for qubits $9-17$ as measured on 8 April 8 2021, between 8:00-10:00pm (UTC-05:00).

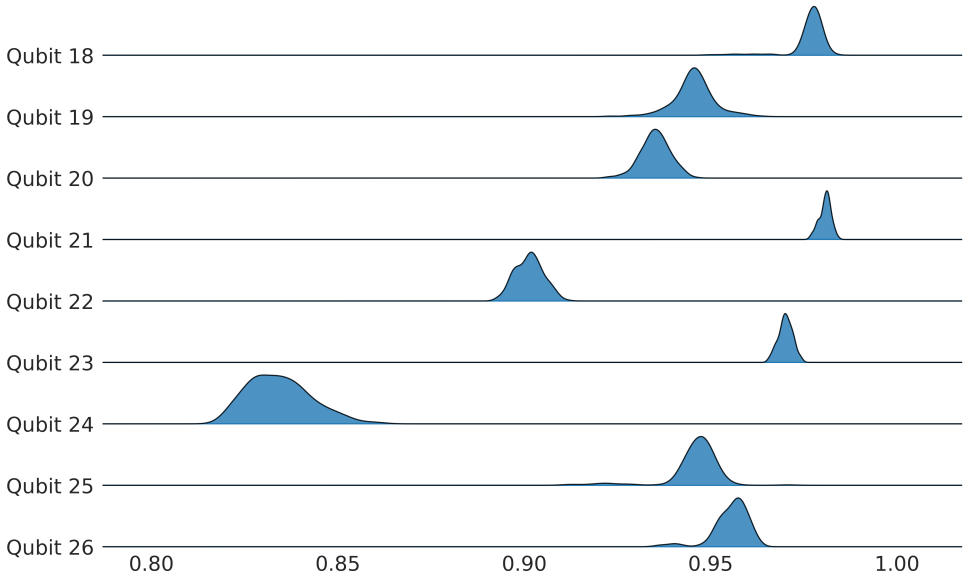


Figure 1.5: state preparation and measurement (SPAM) fidelity distributions on ibm_toronto for qubits $18 - 26$ as measured on 8 April 8 2021, between 8:00-10:00pm (UTC-05:00).

Chapter 2

Noise in quantum computing

Practical efforts to realize a quantum computer (e.g. transmons, trapped ions, silicon quantum dots [18, 80, 81, 1]) introduce various physical processes, referred to as noise, which deviate from the ideal description of a quantum computer. Unlike modern classical computers, which boast device components with extremely low failure rates (e.g., 10^{-17} or less), the current state-of-the-art quantum computers exhibit higher gate-level failure rates (e.g., 10^{-2}). In this dissertation, we mainly use experimental data from transmon[30, 33] based realizations of a quantum computer. Transmon qubits are a variant of superconducting charge qubits designed to reduce sensitivity to charge noise[29].

2.1 Physical sources of noise

The various noise processes [17, 18, 19] can be classified into three groups:

2.1.1 Quantum register

One of the pathways for noise is the implementation of the quantum register, which encounter phenomena such as: (i) Leakage i.e. unintended energy states outside the computational subspace, (ii) Undesired coupling to the external environment (such as spurious charge, magnetic fields, stray photons, lattice vibrations (phonons), nuclear spins) leading to loss of coherence, (iii) Spontaneous decay processes that transition a qubit from an excited state to a lower energy state, (iv) Non-uniformity in qubit's coupling strength to the control field, results in errors in quantum gates, (v) Inter-qubit cross-talk arising from shared control lines or capacitive coupling between neighboring qubits. In trapped ion systems, cross-talk could arises from motional coupling between ions, affecting the states of neighboring ions.

2.1.2 Control system

The quantum register undergoes four fundamental control operations: initialization or reset, measurement, single-qubit rotation gates, and 2-qubit entangling gates. These operations require the application of precisely calibrated control pulses on the qubits. Imperfections in the control system used for logic implementation can arise from several sources. Firstly, pulse distortion occurs when the desired shape and duration of pulses encoding quantum information are altered due to the finite time resolution and frequency response limitations, as well as pulse timing errors. Secondly, control pulses may experience attenuation caused by electromagnetic interference and material imperfections in the quantum system's vicinity. Thirdly, qubits can drift either physically (in the case of trapped ions) or in parameter space. Fourthly, the noise may be an effect mis-calibration.

2.1.3 Thermodynamic isolation system

Transmon qubits, a specific kind of superconducting qubit, require cooling to approximately 10 milli-kelvin in order to mitigate the presence of thermal noise. The thermodynamic isolation system [82] helps achieve this using a system of dilution refrigerators, vacuum chambers, electromagnetic shields, and vibration suppression mechanisms. The dilution refrigerators employ a multi-stage cooling process that gradually reaches colder temperatures, using substances like liquid helium to progressively lower the temperature. The vacuum chambers effectively eliminate gas molecules and particles that could potentially couple with the qubits. The electromagnetic shields are responsible for blocking external radiation and fields from disturbing the quantum state. Inadequate electromagnetic shielding could allow disruptive external radiation to interfere with the qubits. The vibration suppression systems minimize mechanical vibrations and movements that could potentially jeopardize the quantum states. Noise from imperfect thermodynamic control systems can be non-Markovian in nature, which are difficult to rectify using quantum error correction tools.

2.2 Cause of non-stationarity

Non-stationary noise refers to noise in a quantum system that exhibits time-varying statistical properties. The temporal fluctuations of the mean and variance of the energy relaxation times (T_1), dephasing times (T_2), and qubit frequencies, are well-studied topics[18, 19, 36, 37] that suggest suggest that noise in NISQ[34] devices can fluctuate unpredictably. For example, T_1 times have been found to fluctuate by approximately 50 percent within an hour [18]. Similarly, many advances have also been made for spatially varying noise in quantum devices [38], its effect on the choice of circuit

geometry [39], as well as the interplay with cross-talk between qubits for single and two-qubit gates [40, 41].

The causes and mechanisms behind quantum noise non-stationarity are poorly understood. In transmon registers, potential sources of fluctuations include TLS (two-level system) defects, quasi-particles, parasitic microwave modes, phonons, nuclear spins, paramagnetic impurities, spurious resonances, critical current noise, background charges, gate voltage fluctuations, and the electromagnetic environment [57]. Among these, TLS defects have been identified as the primary cause of decoherence [17, 18, 19]. These defects arise from deviations from crystalline order in the naturally occurring oxide layers of transmons, resulting in trapped charges, dangling bonds, tunneling atoms, or collective motion of molecules.

The findings not only highlight the necessity for frequent re-calibration in qubit setups but also question the reproducibility of device characterizations, and their use in error mitigation. Consequently, modeling time-varying quantum noise has become an active area of interest [46], such as through the inclusion of T_1 and T_2 fluctuations in quantum channel models to investigate the concept of time-varying quantum channels (TVQC).

2.3 Decoherence characterization

Decoherence refers to loss of unitarity in state evolution. The traditional definition of decoherence, which describes the decay of off-diagonal terms in the density matrix, is now referred to as dephasing and considered one kind of decoherence [57]. Decoherence studies typically focus on three metrics: transverse relaxation time (T_1), longitudinal relaxation time (T_2), and dephasing time (T_ϕ).

T_1, also known as the transverse relaxation time or relaxation time, measures the attenuation of amplitude in a quantum system. It represents the probability that an excited state $|1\rangle$ will decay to the ground state $|0\rangle$ after time t, and is modeled by the function:

$$\Pr(|1\rangle \to |0\rangle) = 1 - \exp(-t/T_1) \; . \tag{2.1}$$

The decay-time probability density $f_T(t)$ can be described by the exponential function:

$$f_T(t) = T_1 \exp^{-t/T_1} \; , \tag{2.2}$$

whose mean is the density parameter $\mathbb{E}(T) = T_1$.

T_2 is a measure of how long it takes for a qubit in the superposition state to decay. Specifically, it measures the decay of the off-diagonal elements of the density matrix and is modeled by an exponential decay function. Therefore, it captures the loss of synchronization between the basis states of an arbitrary quantum ensemble. There are two types of T_2 time often quoted in literature [20]:

- Ramsey dephasing time T_2^*: measures the time-scale at which a quantum register experiences de-phasing effects when left to evolve freely
- Hahn-echo dephasing time T_2^{echo}: uses intermediate π pulses for re-focusing to increase relaxation time.

When simulating noisy circuits, the appropriate T_2 value to use depends on whether the physical implementation of the circuit uses Hahn-echo for noise suppression or not. We will specifically focus on the Hahn-echo with one echo T_2 time.

Finally, the pure dephasing time (T_ϕ) is an upper bound on the decoherence time for a qubit, since thermal fluctuations in the environment inevitably cause a loss of phase coherence. In practice, the dominant relaxation time is usually T_2 (or sometimes T_1), rather than T_ϕ [20].

The three decoherence benchmarks are related by:

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_\phi} \ . \tag{2.3}$$

2.3.1 Experimental characterization

We analyzed decoherence (i.e. T_1 and T_2 times) in the transmon processor ibm_kolkata. We had 24-hour access on Tuesday, September 12, to September 13 (from 12 noon to 12 noon) through OLCF. We chose this time-frame as it is typical for user program queues for execution on the IBM platform. We measured fluctuations in T_1 and T_2 times for all 27 qubits on the device during this 24-hour period. We validated our software through numerical simulations (detailed in the end of this section).

The complete quantum circuit used to gather the decoherence parameters T_1 and T_2 is too large to display in its entirety. However, in Fig. 2.1, we provide a concise representation of a section of the circuit, specifically for qubit $|0\rangle$. It's important to clarify that this circuit structure is replicated for all 27 qubits in ibm_kolkata, and the entire circuit is executed multiple times to obtain statistical data.

In the sub-circuit presented in Fig. 2.1, we illustrate only one mid-circuit reset for the sake of clarity. In reality, the full circuit employs three conditional resets to ensure a high probability of mid-circuit

reset success. This choice of three resets aligns with qiskit guidelines to optimize the likelihood of successful resets.

It's worth noting that mid-circuit measurement allows for the simultaneous collection of the decoherence parameters T_1 and T_2 with a time interval of just a few hundred microseconds. This simultaneous data collection facilitates the empirical calculation of temporal correlations between these parameters.

The basic T_1 measurement circuit begins by initializing a qubit to the $|0\rangle$ state and then applying an X gate to transition it to the $|1\rangle$ state. Subsequently, a phase gate is introduced, during which the qubit is affected by noise. Following this, a measurement is performed in the Z basis. In the absence of noise, the measurement would yield the $|1\rangle$ state with complete certainty. However, in the presence of noise, the probability of obtaining the $|1\rangle$ state is less than 100%, and this probability depends solely on T_1. Therefore, by analyzing the observed probability of measuring $|1\rangle$, we can deduce an estimate for the T_1 time.

The basic T_2 circuit starts by setting a qubit to the $|0\rangle$ state. It then uses the Hadamard gate to create an equal superposition of $|0\rangle$ and $|1\rangle$. After a brief phase-shift delay, another H gate is applied, followed by a measurement in the Z basis. In the absence of noise, the circuit guarantees a 100 percent chance of retrieving the $|0\rangle$ state. However, if there's dephasing noise, the probability decreases, and this can be used to estimate the T_2 time.

It's important to note that these calculations also consider SPAM noise, which will be discussed in detail later.

To calculate T_1 and T_2, data is fitted to an exponential decay plot using four different evolution times: 10 μs, 50 μs, 100 μs, and 160 μs. For each qubit, the process involves the following sequence:

1. basic T_1 circuit for a 10 μs, followed by measurement and reset
2. basic T_2 circuit for a 10 μs, followed by measurement and reset
3. basic T_1 circuit for a 50 μs, followed by measurement and reset
4. basic T_2 circuit for a 50 μs, followed by measurement and reset
5. basic T_1 circuit for a 100 μs, followed by measurement and reset
6. basic T_2 circuit for a 100 μs, followed by measurement and reset
7. basic T_1 circuit for a 160 μs, followed by measurement and reset
8. basic T_2 circuit for a 160 μs, followed by measurement

The statistical estimation of T_1 and T_2 are impacted by SPAM noise. We model the SPAM noise using a binary asymmetric model with the parameters:

$$\begin{aligned} u_{00} + u_{01} &= 1 \\ u_{10} + u_{11} &= 1 \end{aligned} \tag{2.4}$$

where, u_{00} denotes the probability of getting 0 given an input of 0, u_{01} denotes the probability of getting 1 given an input of 0, u_{10} denotes the probability of getting 0 given an input of 1, and u_{11} denotes the probability of getting 1 given an input of 1. We define p as the survival probability for the T_1 circuit, indicating the probability of an excited state enduring beyond time t in T_1 measurement. Similarly, q denotes the survival probability for the T_2 circuit, reflecting the probability of observing a ground state after time t in T_2 measurement.

$$\begin{aligned} \tilde{p}_1 &= \frac{\#1's \text{ observed post-measurement of } T_1 \text{ circuit}}{\#\text{Circuit repetitions}} \\ \tilde{q}_0 &= \frac{\#0's \text{ observed post-measurement of } T_2 \text{ circuit}}{\#\text{Circuit repetitions}} \end{aligned} \tag{2.5}$$

The survival probabilities in absence of SPAM error are given by:

$$\begin{aligned} p_1 &= e^{-\tau/T_1} \\ q_0 &= \frac{1}{2}(1 + e^{-\tau/T_2}) \end{aligned} \tag{2.6}$$

While the formula for p_1 (for T_1) is straightforward, the derivation for q_0 (for Hahn-echo T_2 with one echo) is a little more involved. The steps are as follows. First we initialize the qubit in the ground state:

$$\rho_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.7}$$

Then, we subject it to a Hadamard gate:

$$\rho_1 = H\rho_0 H^\dagger = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{2.8}$$

Then, we evolve the density matrix for time $\tau/2$ using the Hamiltonian $\Omega = \lambda \left|1\right\rangle \left\langle 1\right|$. The unitary operator for this phase gate D is:

$$D = \exp(-i\Omega\frac{\tau}{2}) = \exp(-i\lambda\frac{\tau}{2}) \left|1\right\rangle \left\langle 1\right| + \left|0\right\rangle \left\langle 0\right| \equiv \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\lambda\frac{\tau}{2}} \end{pmatrix} \tag{2.9}$$

Thus, the state becomes:

$$\rho_2 = D\rho_1 D^\dagger = \frac{1}{2}\begin{pmatrix} 1 & e^{-i\lambda\frac{\tau}{2}} \\ e^{i\lambda\frac{\tau}{2}} & 1 \end{pmatrix} \tag{2.10}$$

However, during this time-evolution, it is acted upon by an amplitude and phase damping (APD) channel. The state after taking APD noise into account is:

$$\rho_3 = \mathcal{E}_{\text{APD}}(\rho_2) = \begin{pmatrix} 1 - \frac{1}{2}e^{-\frac{\tau}{2}/T_1} & \frac{1}{2}e^{-i\lambda t}e^{-\frac{\tau}{2}/T_2} \\ \frac{1}{2}e^{i\lambda t}e^{-\frac{\tau}{2}/T_2} & \frac{1}{2}e^{-\frac{\tau}{2}/T_1} \end{pmatrix} \tag{2.11}$$

After this comes a deliberate bit-flip through a $\mathbb{X}$ gate (assumed noiseless):

$$\rho_4 = X\rho_3 X = \begin{pmatrix} \frac{1}{2}e^{-\frac{\frac{\tau}{2}}{T_1}} & \frac{1}{2}e^{i\lambda\frac{\tau}{2}}e^{-\frac{\tau}{2}/T_2} \\ \frac{1}{2}e^{-i\lambda\frac{\tau}{2}}e^{-\frac{\tau}{2}/T_2} & 1 - \frac{1}{2}e^{-\frac{\tau}{2}/T_1} \end{pmatrix} \tag{2.12}$$

This is followed up with another phase gate D subject to APD noise:

$$\rho_5 = D\rho_4 D^\dagger = \begin{pmatrix} \frac{1}{2}e^{-\frac{\tau}{2}/T_1} & \frac{1}{2}e^{-\frac{\frac{\tau}{2}}{T_2}} \\ \frac{1}{2}e^{-\frac{\tau}{2}/T_2} & 1 - \frac{1}{2}e^{-\frac{\frac{\tau}{2}}{T_1}} \end{pmatrix} \tag{2.13}$$

$$\rho_6 = \mathcal{E}_{\text{APD}}(\rho_5) = \begin{pmatrix} 1 - e^{-\frac{\tau}{2}/T_1}\left(1 - \frac{1}{2}e^{-\frac{\tau}{2}/T_1}\right) & \frac{1}{2}e^{-\frac{2\frac{\tau}{2}}{T_2}} \\ \frac{1}{2}e^{-2\frac{\tau}{2}/T_2} & e^{-\frac{\tau}{2}/T_1}\left(1 - \frac{1}{2}e^{-\frac{\tau}{2}/T_1}\right) \end{pmatrix} \tag{2.14}$$

Then another Hadamard is applied:

$$\rho_7 = \mathbb{H}\rho\mathbb{H}^\dagger \tag{2.15}$$

Finally, we measure the probability of the qubit being in the ground state. The probability of observing $|0\rangle$ in the final measurement is:

$$q_0 = \Pr(0) = \frac{1}{2}\left(1 + e^{-\tau/T_2}\right) \tag{2.16}$$

The survival probabilities in presence of SPAM error are given by:

$$\begin{aligned} \tilde{p}_1 &= p_1 u_{11} + p_0 u_{01} = e^{-\tau/T_1}\left[1 - u_{10} - u_{01}\right] + u_{01} \\ \tilde{q}_0 &= q_0 u_{00} + q_1 u_{10} = e^{-\tau/T_2}\frac{1 - u_{10} - u_{01}}{2} + \frac{1 + u_{10} - u_{01}}{2} \end{aligned} \tag{2.17}$$

2.3.2 T_1 estimation

For the T_1 circuit, the observed data y_l measured in the Z-basis follows a Bernoulli distribution:

$$y_l = \begin{cases} 1, & \text{with probability } \tilde{p}_1 \\ 0, & \text{with probability } 1 - \tilde{p}_1 \end{cases} \tag{2.18}$$

where $\tilde{p}_1 = e^{-\tau/T_1}\left[1 - u_{10} - u_{01}\right] + u_{01}$. The Likelihood function is given by:

$$\mathcal{L} = \prod_{l=1}^{L} \Pr\left(Y_l = y_l\right) = \prod_{l=1}^{L} \tilde{p}_1^{y_l}(1 - \tilde{p}_1)^{1-y_l} \tag{2.19}$$

Setting:

$$\frac{\partial \log \mathcal{L}}{\partial \tilde{p}_1} = 0, \tag{2.20}$$

we get,

$$\tilde{p}_1^* = \sum \frac{y_l}{L} \tag{2.21}$$

Thus we have obtained the estimator (denoted by the * sign) for $\tilde{p}_1$. Note that

$$\mathbb{E}(\tilde{p}_1^*) = \sum \frac{\mathbb{E}(y_l)}{L} = \sum \frac{\tilde{p}_1}{L} = \tilde{p}_1 \tag{2.22}$$

and hence it is an unbiased estimator.

The variance of this estimate is given by:

$$\mathrm{Var}(\tilde{p}_1^*) = \sigma^2(\tilde{p}_1^*) = \sum \frac{\mathrm{Var}(y_l)}{L^2} = \frac{\tilde{p}_1(1 - \tilde{p}_1)}{L} \approx \frac{\tilde{p}_1^*(1 - \tilde{p}_1^*)}{L} \tag{2.23}$$

Using the above, we can get K different equations, one for each phase gate with evolution time τ_k and whose post-measurement results are denoted by $\{y_{l,k}\}$:

$$\begin{aligned} e^{-\tau_1/T_1}\left[1 - u_{10} - u_{01}\right] + u_{01} &= \frac{\sum y_{l,1}}{L} \\ e^{-\tau_2/T_1}\left[1 - u_{10} - u_{01}\right] + u_{01} &= \frac{\sum y_{l,2}}{L} \\ &\cdots \\ e^{-\tau_K/T_1}\left[1 - u_{10} - u_{01}\right] + u_{01} &= \frac{\sum y_{l,K}}{L} \end{aligned} \tag{2.24}$$

Each of these K equations has three unknowns: T_1, u_{10} and u_{01}. From these K equations, we find the best fit value for T_1 using the scipy.optimize.minimize module in python. K has to be at least 3 so that the problem is not underspecified (overspecified is okay).

2.3.3 T_2 estimation

In an analogous manner, for the T_2 circuit, the observed data y_l measured in the Z-basis follows a Bernoulli distribution:

$$y_l = \begin{cases} 1, & \text{with probability } 1 - \tilde{q}_0 \\ 0, & \text{with probability } \tilde{q}_0 \end{cases} \tag{2.25}$$

where $\tilde{q}_0 = e^{-\tau/T_2} \frac{[1-u_{10}-u_{01}]}{2} + \frac{[1+u_{10}-u_{01}]}{2}$.

The Likelihood function is given by:

$$\mathcal{L} = \prod_{l=1}^{L} \Pr\left(Y_l = y_l\right) = \prod_{l=1}^{L} \tilde{q}_0^{1-y_l} (1 - \tilde{q}_0)^{y_l} \tag{2.26}$$

Setting:

$$\frac{\partial \log \mathcal{L}}{\partial \tilde{q}_0} = 0, \tag{2.27}$$

we get,

$$\tilde{q}_0^* = 1 - \sum \frac{y_l}{L} \tag{2.28}$$

Thus we have obtained the estimator (denoted by the * sign) for $\tilde{q}_0$. Note that

$$\mathbb{E}(\tilde{q}_0^*) = \sum \frac{\mathbb{E}(y_l)}{L} = \sum \frac{\tilde{q}_0}{L} = \tilde{q}_0 \tag{2.29}$$

and hence it is an unbiased estimator.

The variance of this estimate is given by:

$$\mathrm{Var}(\tilde{q}_0^*) = \sigma^2(\tilde{q}_0^*) = \sum \frac{\mathrm{Var}(y_l)}{L^2} = \frac{\tilde{q}_0(1-\tilde{q}_0)}{L} \approx \frac{\tilde{q}_0^*(1-\tilde{q}_0^*)}{L} \tag{2.30}$$

Using the above, we can get K different equations, one for each phase gate with evolution time τ_k and whose post-measurement results are denoted by $\{y_{l,k}\}$:

$$\begin{aligned} e^{-\tau_1/T_2}\frac{1-u_{10}-u_{01}}{2}+\frac{1+u_{10}-u_{01}}{2} =& \frac{\sum y_{l,1}}{L} \\ e^{-\tau_2/T_2}\frac{1-u_{10}-u_{01}}{2}+\frac{1+u_{10}-u_{01}}{2} =& \frac{\sum y_{l,2}}{L} \\ & \cdots \\ e^{-\tau_K/T_2}\frac{1-u_{10}-u_{01}}{2}+\frac{1+u_{10}-u_{01}}{2} =& \frac{\sum y_{l,K}}{L} \end{aligned} \tag{2.31}$$

Each of these K equations has three unknowns: T_2, u_{10} and u_{01}. From these K equations, we find the best fit value for T_2 using the scipy.optimize.minimize module in python. K has to be at least 3 so that the problem is not underspecified (overspecified is okay).

2.3.4 Error bars on decoherence estimates

For the T_1 circuit:

$$\begin{aligned} p_1 =& e^{-\tau/T_1} \\ \Rightarrow \log p_1 =& -\frac{\tau}{T_1} \\ \Rightarrow \frac{\delta p_1}{p_1} =& \frac{\tau}{T_1^2}\delta T_1 \\ \Rightarrow \sigma_{T_1}^2 =& \frac{T_1^4\left(e^{\tau/T_1}-1\right)}{L\tau^2} \end{aligned} \tag{2.32}$$

Now suppose that we use K different evolution times in the circuit: $\tau_1, \tau_2, \cdots \tau_K$. Let the desired standard deviation of the T_1 estimate be σ_{desired}. Since the data underlying the estimation obtained at different times are independent:

$$\sigma_{\text{desired}}^2 = \frac{1}{K}\sum \sigma_i^2 \tag{2.33}$$

where σ_i^2 is the variance obtained when T_1 is estimated using a delay gate with delay time τ_i. This gives:

$$L_{\min} = \frac{T_1^4}{K\sigma_{\text{desired}}^2}\sum \frac{e^{\tau_i/T_1}-1}{\tau_i^2} \tag{2.34}$$

For the T_2 circuit:

$$\begin{aligned} q_0 =& \frac{1}{2}(1+e^{-\tau/T_2}) \\ \Rightarrow \log(2q_0-1) =& -\frac{\tau}{T_2} \\ \Rightarrow \frac{2\delta q_0}{2q_0-1} =& \frac{\tau}{T_2^2}\delta T_2 \\ \Rightarrow \sigma_{T_2}^2 =& \frac{T_2^4\left(e^{2\tau/T_2}-1\right)}{L\tau^2} \end{aligned} \tag{2.35}$$

Now suppose that we use K different evolution times in the circuit: $\tau_1, \tau_2, \cdots \tau_K$. Let the desired standard deviation of the T_2 estimate be σ_{desired}. Since the data underlying the estimation obtained at different times are independent:

$$\sigma^2_{\text{desired}} = \frac{1}{K} \sum \sigma_i^2 \tag{2.36}$$

where σ_i^2 is the variance obtained when T_2 is estimated using phase-shift time τ_i. This gives:

$$L_{\min} = \frac{T_2^4}{K \sigma^2_{\text{desired}}} \sum \frac{e^{2\tau_i/T_2} - 1}{\tau_i^2} \tag{2.37}$$

Since we have a circuit that measures T_1 and T_2 in one go, we have to take the max of the $L_{\min}$ for each of the two cases (i.e. max of $L_{\min}$ for T_1 and $L_{\min}$ for T_2).

Using a noisy simulation with known error estimates (detailed in the program validation section next), we arrived at $L_{\min}$ at each time to be $20,000$ samples. There were four delay gates (aka four separate decay experiments), so a total of $80,000$ samples went into the computation of each T_1 and T_2 value for each time-stamp. The standard-deviation we aimed for is $1\mu s$.

2.3.5 Program validation

Program validation is crucial to ensure that, within the assumptions of our theoretical noise model, our simulations of quantum circuits, utilizing known noise parameters, yield precise results. Specifically, we aimed for our statistical analysis to accurately recover the expected values. Additionally, program validation helps us determine the required number of circuit repetitions for achieving a specified level of outcome precision. While minor deviations in the final standard deviation are expected due to various factors, they should generally align with our target precision.

Since we individually measure the T_1 and T_2 times for each of the 27 qubits of ibm_kolkata, validating the program for a single qubit suffices. We found that a sample size of 10,000 is sufficient to attain the desired precision of 1 microsecond, but for safety, we opted for 20,000 samples in the final run on real device. Given the use of four delay gates, each T_1 or T_2 data point estimation relies on a total of $80,000$ samples. For consistency with the ibm_kolkata processor's specifications, we set the readout error at 0.028 T_1 time at 134 microseconds and T_2 time at 93 microseconds in our noise simulation.

Our program validation yielded an estimated T1 time of 134.52 microseconds, which falls within the 1-microsecond precision target. The estimated T2 time was 94.05 microseconds, slightly exceeding the 1-microsecond precision by a difference of 1.03 microseconds which we deemed acceptable.

2.3.6 Summary of results

In terms of the T_1 parameter, qubit 15 performed the best, exhibiting a median T_1 time of 184 microseconds, while qubit 4 performed the worst with a median T_1 time of 74 microseconds. Across all 27 qubits, the median T_1 time was 116 microseconds, with a standard deviation of 1.5 microseconds, and the range of T_1 times spanned from 32 to 297 microseconds (see Fig. 2.2 (a)). As for the T_2 parameter, qubit 3 demonstrated the highest performance, displaying a median T_2 time of 76 microseconds, while qubit 19 had the poorest performance with a median T_2 time of 15 microseconds. Across all 27 qubits, the median T_2 time averaged 30 microseconds, with a standard deviation of 1.1 microseconds, and the range of T_2 times varied from 3 to 191 microseconds (see Fig. 2.2 (b)). The time-series for all the 27 qubits over the 24 hour period is shown in Figs. 2.4 (b)- 2.17.

2.4 Modeling quantum noise channels

2.4.1 Amplitude and Phase damping channel

The amplitude damping channel $\mathcal{E}^{\mathrm{AD}}(\cdot)$ and de-phasing channel $\mathcal{E}^{\mathrm{PD}}(\cdot)$ are two fundamental sources of quantum de-coherence and information loss in transmons [18, 36, 11]. A realistic model for this noise channel, denoted as APD, involves a combination of amplitude damping and de-phasing. Amplitude damping can be described by the Kraus operators E_0^{AD} and E_1^{AD}, while phase damping can be described by E_0^{PD} and E_1^{PD}, as follows [11]:

$$\mathcal{E}^{\mathrm{AD}}(\rho) = \sum_{k=0}^{1} E_k^{\mathrm{AD}} \rho E_k^{\mathrm{AD}\dagger} \ , \tag{2.38}$$

$$E_0^{\mathrm{AD}} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \ , \tag{2.39}$$

$$E_1^{\mathrm{AD}} = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} \ , \tag{2.40}$$

$$\mathcal{E}^{\mathrm{PD}}(\rho) = \sum_{k=0}^{1} E_k^{\mathrm{PD}} \rho E_k^{\mathrm{PD}\dagger} \ , \tag{2.41}$$

$$E_0^{\mathrm{PD}} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix} \ , \tag{2.42}$$

$$E_1^{\mathrm{PD}} = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{pmatrix} \ . \tag{2.43}$$

Here, $\gamma = 1 - \exp(-t/T_1)$ and $\lambda = 1 - \exp(-t/T_2)$, where t is the time scale of the decoherence process. The relation between T_1, T_ϕ, and T_2 was previously discussed in Eq. (2.3). The Kraus decomposition of the combined amplitude and phase damping channel $\mathcal{E}^{\mathrm{APD}}(\cdot)$, valid for a single qubit, can be expressed as E_0^{APD}, E_1^{APD}, and E_2^{APD}.

$$\mathcal{E}^{\mathrm{APD}}(\rho) \coloneqq \mathcal{E}^{\mathrm{PD}} \circ \mathcal{E}^{\mathrm{AD}} = \sum_{k=0}^{2} E_k^{\mathrm{APD}} \rho E_k^{\mathrm{APD}\dagger} \ , \tag{2.44}$$

where,

$$E_0^{\mathrm{APD}} = E_0^{\mathrm{PD}} E_0^{\mathrm{AD}} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{[1-\gamma][1-\lambda]} \end{pmatrix} \ , \tag{2.45}$$

$$E_1^{\mathrm{APD}} = E_0^{\mathrm{PD}} E_1^{\mathrm{AD}} = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} \ , \tag{2.46}$$

$$E_2^{\mathrm{APD}} = E_1^{\mathrm{PD}} E_0^{\mathrm{AD}} = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{[1-\gamma]\lambda} \end{pmatrix} \ . \tag{2.47}$$

Using the fact that:

$$E_0^{\mathrm{APD}} = \frac{1 + \sqrt{1-\lambda-\gamma+\lambda\gamma}}{2} I + \frac{1 - \sqrt{1-\lambda-\gamma+\lambda\gamma}}{2} Z \ , \tag{2.48}$$

$$E_1^{\mathrm{APD}} = \frac{\sqrt{\gamma}}{2} X + \frac{\sqrt{\gamma}}{2} iY \ , \tag{2.49}$$

$$E_2^{\mathrm{APD}} = \frac{\sqrt{\lambda-\lambda\gamma}}{2} I - \frac{\sqrt{\lambda-\lambda\gamma}}{2} Z \ . \tag{2.50}$$

the APD channel can be expressed as:

$$\begin{aligned} \mathcal{E}^{\mathrm{APD}}(\rho) = & \frac{2-\gamma+2\sqrt{1-\lambda-\gamma+\lambda\gamma}}{4} \rho + \frac{\gamma}{4} X\rho X - \frac{\gamma}{4} Y\rho Y \\ & + \frac{2-\gamma-2\sqrt{1-\lambda-\gamma+\lambda\gamma}}{4} Z\rho Z \\ & + \frac{\gamma}{4} I\rho Z + \frac{\gamma}{4} Z\rho I - \frac{\gamma}{4i} X\rho Y - \frac{\gamma}{4i} Y\rho X \ , \end{aligned} \tag{2.51}$$

where λ and γ are the APD parameters, and I, X, Y, and Z are the Pauli matrices.

2.4.2 Depolarizing channel

Next, the depolarizing channel is a common type of quantum noise channel. It works by randomly applying one of the Pauli operators $(\mathbb{X}, \mathbb{Y}, \mathbb{Z})$ to the quantum state with a certain (but equal)

probability, causing a loss of information about the state. For a qubit:

$$\mathcal{E}_D(\rho) = (1 - \mathrm{x})\rho + \mathrm{x}\frac{\mathbb{I}}{d}, \tag{2.52}$$

where ρ is the input quantum state, x is the probability of noise occurring, $\mathbb{I}$ is the identity operator, and d is the dimension of the Hilbert space ($d = 2$ for a qubit).

2.4.3 Pauli noise channel

The Pauli noise channel is a generalization of the depolarizing channel. It encompasses the effect of the bit-flip ($\mathbb{X}$), phase-flip ($\mathbb{Z}$), and bit-phase-flip ($\mathbb{Y}$) errors with unequal probabilities.

The impact of Pauli noise on quantum information encoded in an n-qubit register is shown below:

$$\mathcal{E}_{\mathrm{x}}(\rho) = \sum_{i=0}^{N_p-1} \mathrm{x}_i P_i(n) \rho P_i(n)^\dagger \tag{2.53}$$

where N_p denotes the total number of Pauli coefficients and $P_i(n)$ represents n-qubit Pauli operators. The coefficients contribute to a simplex:

$$\sum_{i=0}^{N_p-1} \mathrm{x}_i = 1, \quad \mathrm{x}_i \geq 0 \tag{2.54}$$

Pauli noise channel is widely used in quantum error correction because it is a simple and natural model for random quantum noise [83, 84, 49]. It is a well-understood and easily implementable noise model that can simulate a variety of realistic physical processes that lead to quantum errors, such as dephasing, amplitude damping, and phase-flip errors. Additionally, the Pauli noise channel is mathematically tractable and can be efficiently simulated, making it a useful tool for developing and testing quantum error correction protocols. It can be used to estimate the average fidelity of a quantum gate subject to the original APD channel and identify codes that work for the APD channel [84]. The Pauli noise channel, although not a completely general noise model, still manages to model many practical situations. It is widely used because of two reasons: (a) it is efficiently simulatable on a classical computer (per the Gottesman-Knill theorem) and (b) when used as a proxy for physically accurate noise models (such as the amplitude and phase damping noise) which are not efficiently simulatable on a classical computer, it still manages to preserve important properties like entanglement fidelity [85].

Remarkably, Pauli twirling can map[86, 87, 88] a more complex quantum noise channel (e.g. APD) to a simple Pauli channel while preserving certain features such as the average channel fidelity and

the entanglement fidelity [85]. Consider a single-qubit amplitude and phase damping channel (APD) [58]. Upon Pauli twirling [89]:

$$\mathcal{E}_{\text{twirl}}(\rho) = \frac{1}{4} \sum_{A \in \{I,X,Y,Z\}} A^\dagger \mathcal{E}_{\text{APD}} \left(A \rho A^\dagger \right) A \tag{2.55}$$

$$= \sum_{k=0}^{3} c_k \sigma_k \rho \sigma_k \tag{2.56}$$

an APD channel becomes a Pauli noise channel. Here, $\{\sigma_k\}_{k=0}^3 = \{I, X, Y, Z\}$ are the Pauli matrices. Thus, the coefficients of the Pauli noise channel are functions of the coefficients of the original APD channel, which in turn are functions of the decoherence times T_1 and T_2 [46]:

$$c_1 = c_2 = \frac{1}{4} \left[1 - \exp\left(-t/T_1 \right) \right] \tag{2.57}$$

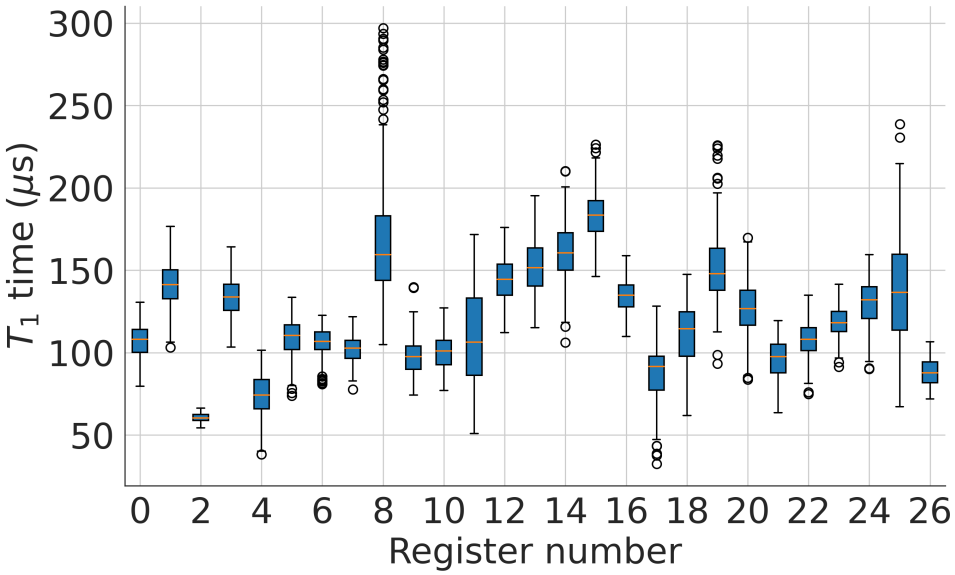
$$c_3 = \frac{1}{4} \left[1 - \exp\left(-t/T_2 \right) \right] \tag{2.58}$$

$$c_0 = 1 - (c_1 + c_2 + c_3) \tag{2.59}$$

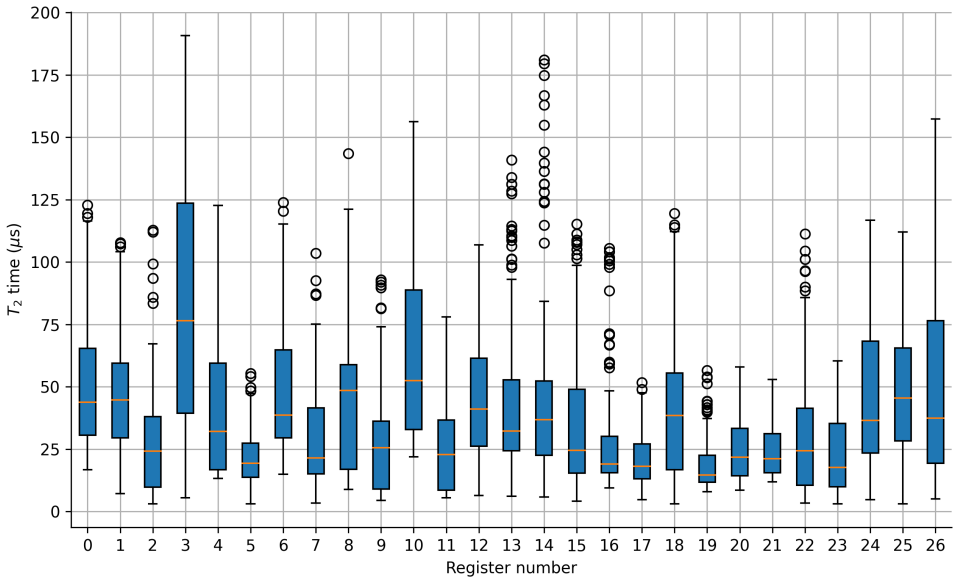
This directly links the estimation of Pauli channels to the decoherence data collected in Sec. 2.3.



Figure 2.1: Interleaved sub-circuit post transpilation.



(a)



(b)

Figure 2.2: Spatial non-stationarity of decoherence times. Indivudal error-bars do not capture the variation across qubits. (a) In terms of the T_1 parameter, qubit 15 performed the best, exhibiting a median T_1 time of 184 microseconds, while qubit 4 performed the worst with a median T_1 time of 74 microseconds. Across all 27 qubits, the median T_1 time was 116 microseconds, with a standard deviation of 1.5 microseconds, and the range of T_1 times spanned from 32 to 297 microseconds. (b) For the T_2 parameter, qubit 3 demonstrated the highest performance, displaying a median T_2 time of 76 microseconds, while qubit 19 had the poorest performance with a median T_2 time of 15 microseconds. Across all 27 qubits, the median T_2 time averaged 30 microseconds, with a standard deviation of 1.1 microseconds, and the range of T_2 times varied from 3 to 191 microseconds.

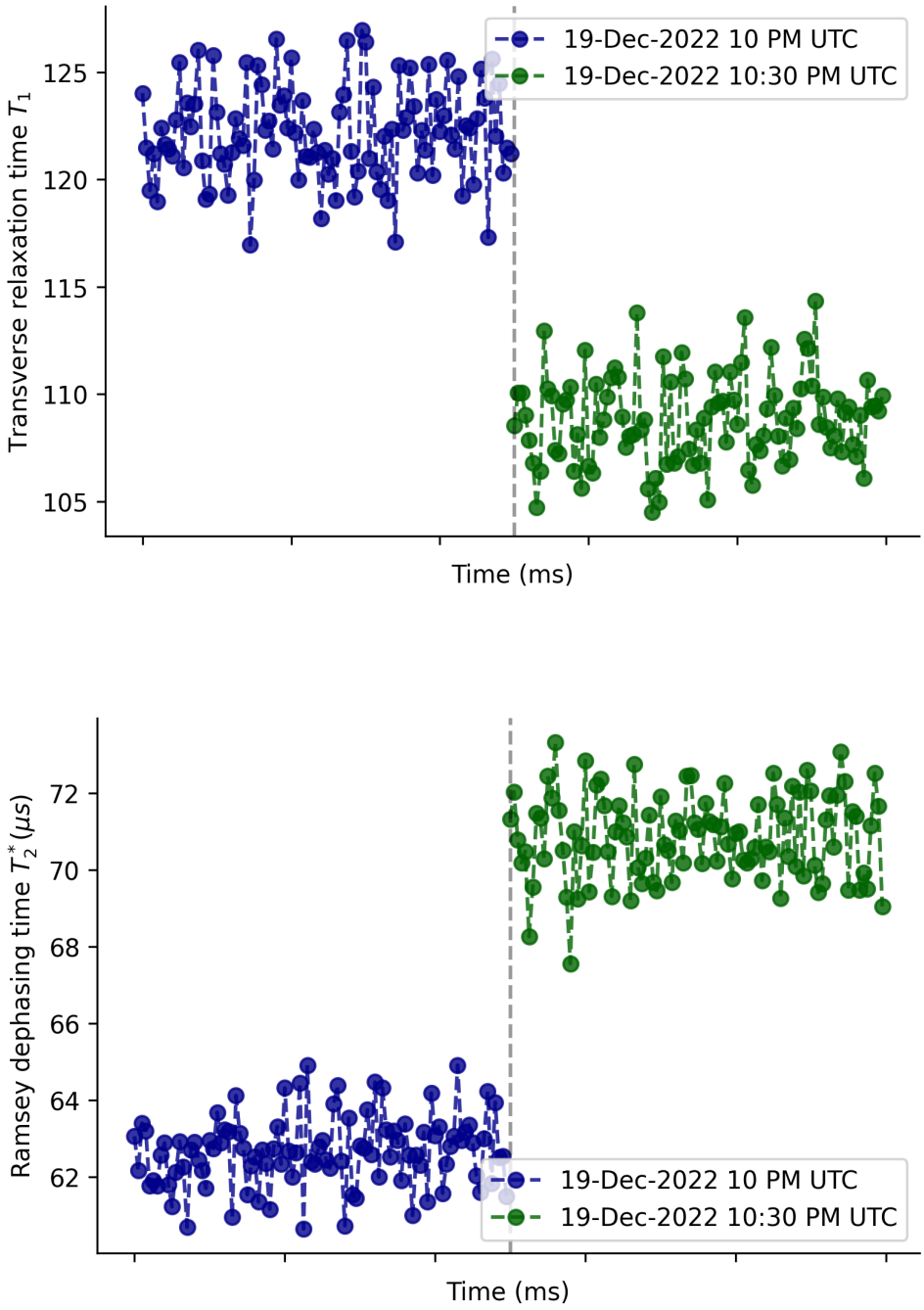
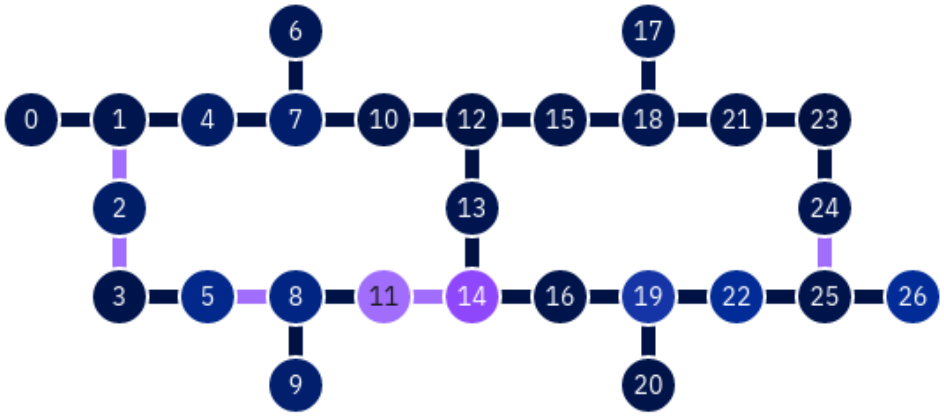
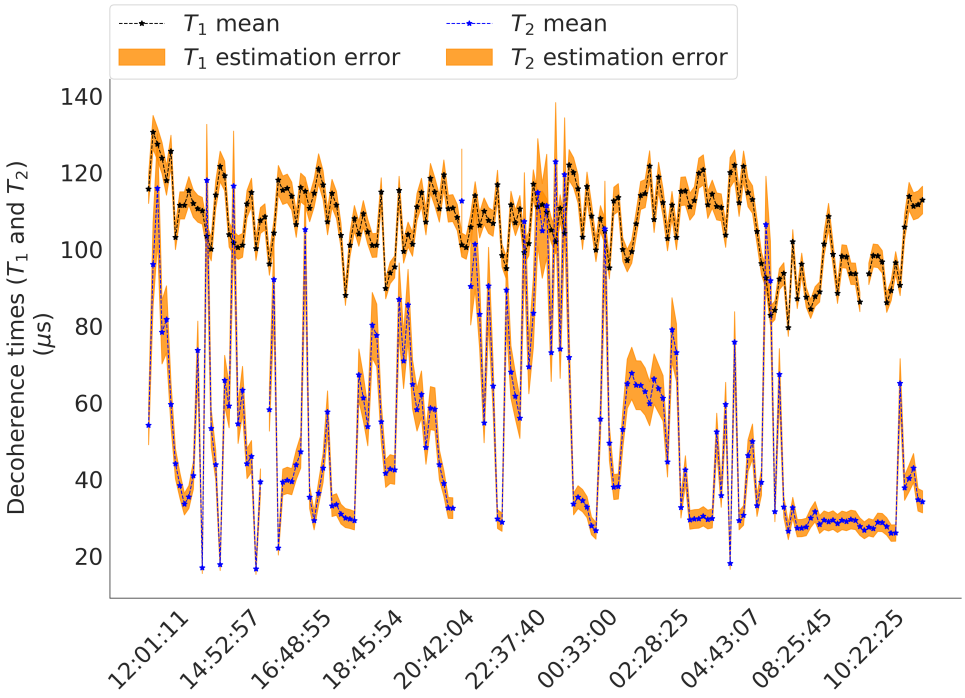


Figure 2.3: This figure presents non-stationary temporal dynamics of decoherence for a qubit on IBM's belem device. The top figure shows T_1 relaxation time series for qubit 0, where two datasets were collected for 5 ms each on the same day, separated by a vertical line. The blue dataset varies between 116-126μs with a mean of 122 μs and a standard deviation of 2 μs, while the green dataset varies between 104-114μs with a mean of 108 μs and a standard deviation of 2 μs. The bottom figure displays Ramsey dephasing time (T_2) series for qubit 0, where two datasets were collected for 5 ms each on the same day, separated by a vertical line. The blue dataset varies between 67-73μs with a mean of 70 μs and a standard deviation of 1 μs, while the green dataset varies between 60-64μs with a mean of 62 μs and a standard deviation of 1 μs, collected around different times on the same day. The data shows significant non-stationarity in decoherence values over a 30-minute interval.

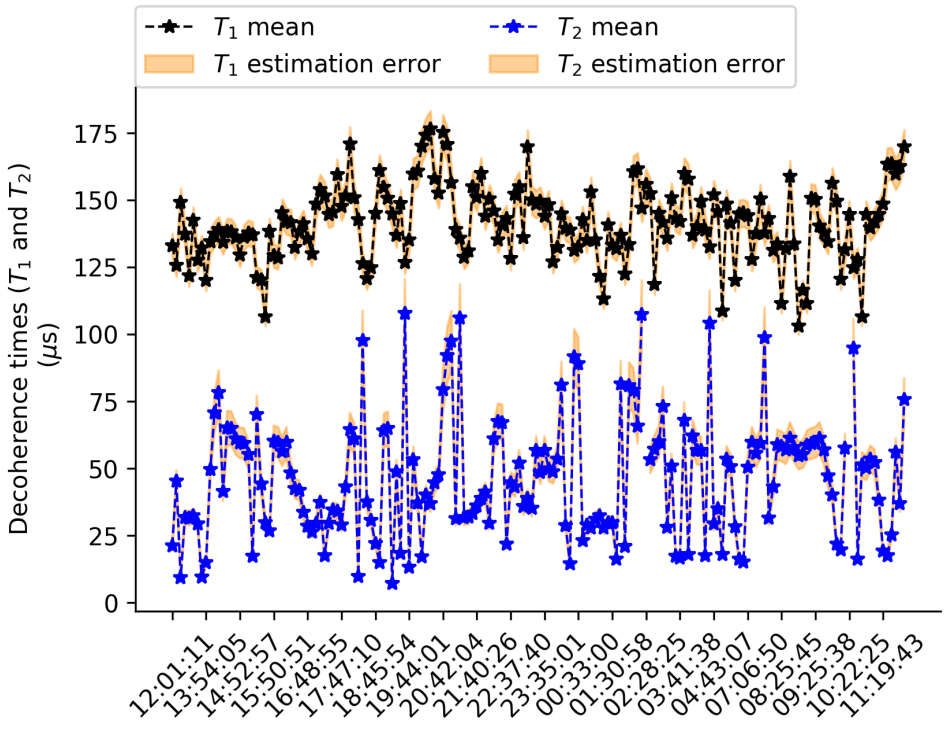


(a)

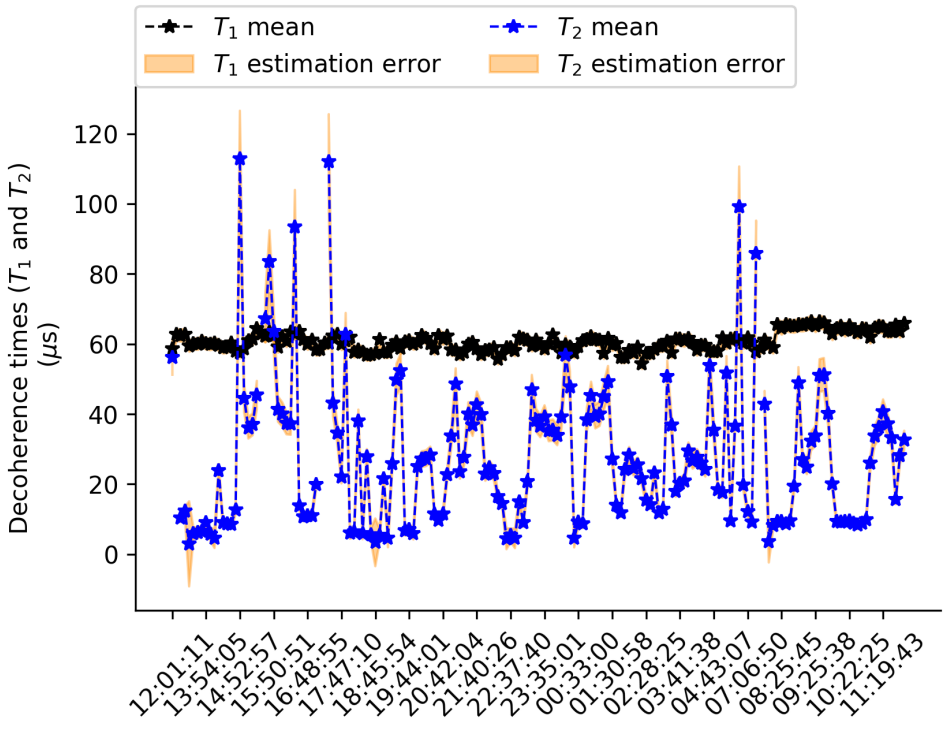


(b)

Figure 2.4: (a) Schematic of the 27-qubit device ibm_kolkata. (b) Estimated T_1 and T_2 time-series for qubit 0 as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023.

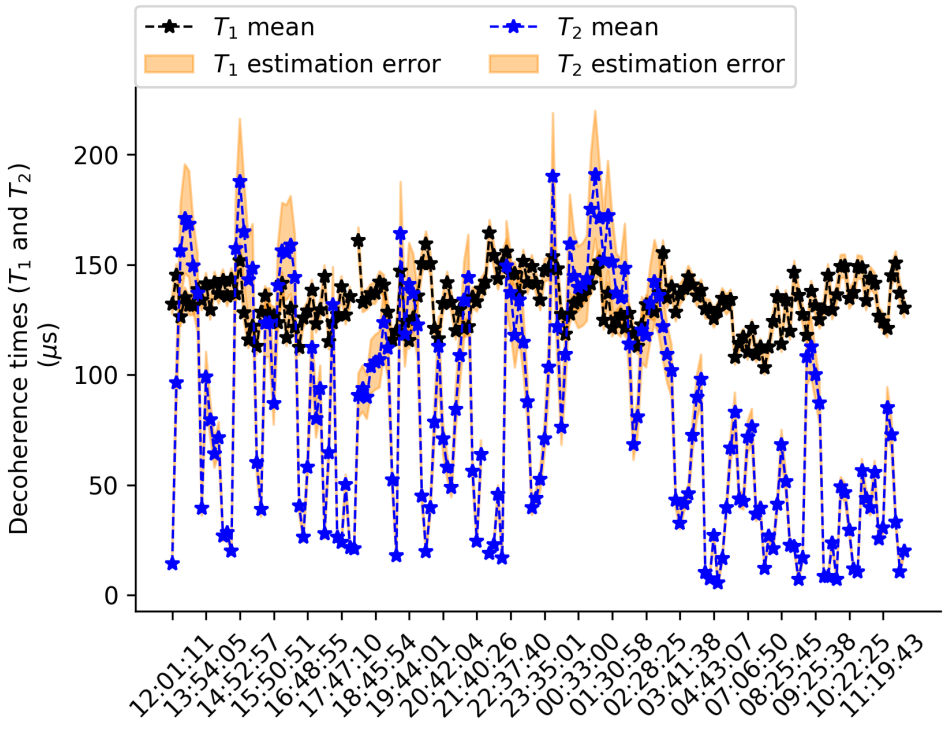


(a)

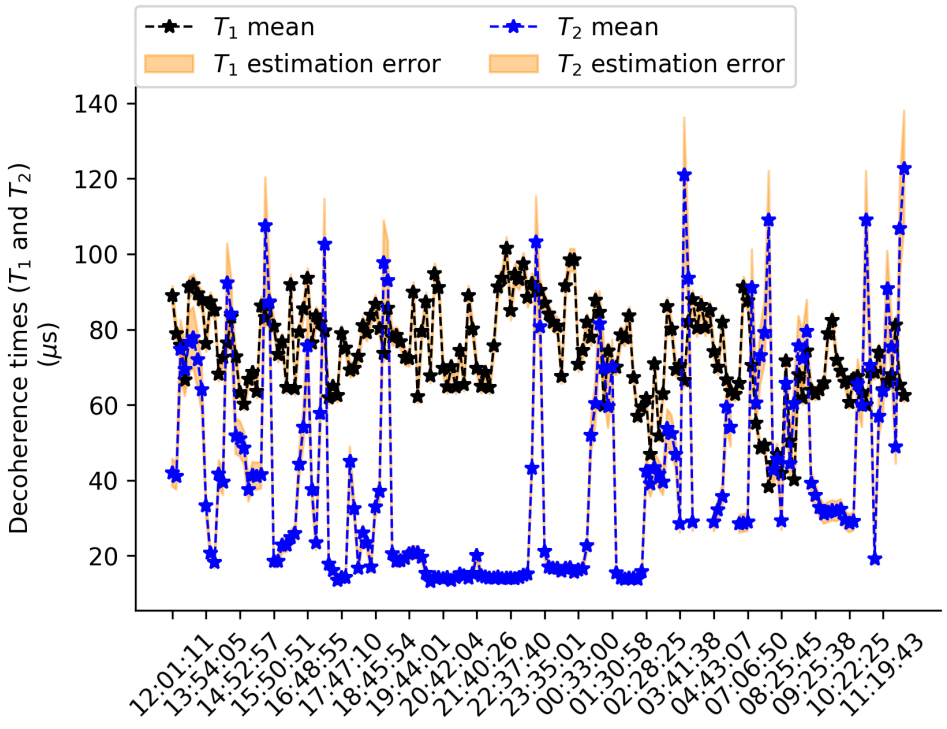


(b)

Figure 2.5: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 1 and (b) qubit 2.

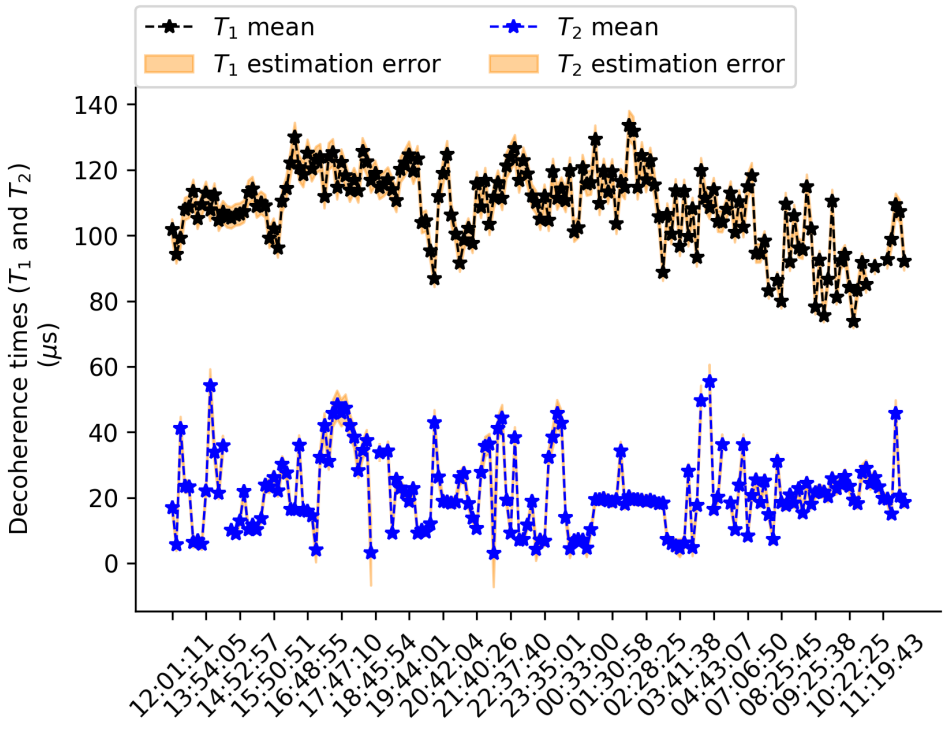


(a)

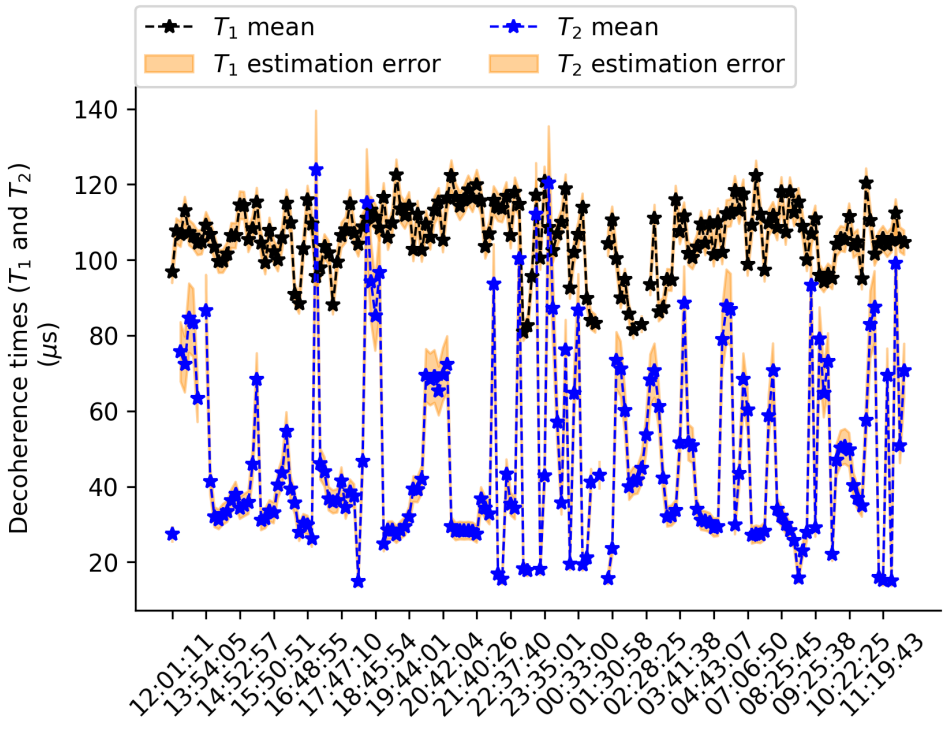


(b)

Figure 2.6: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 3 and (b) qubit 4.

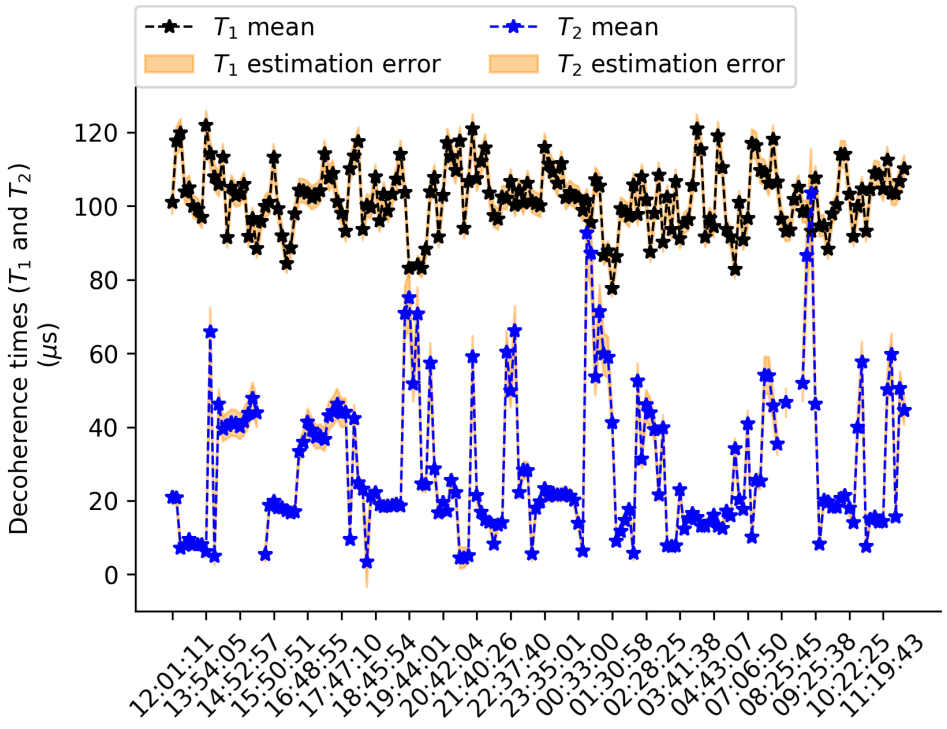


(a)

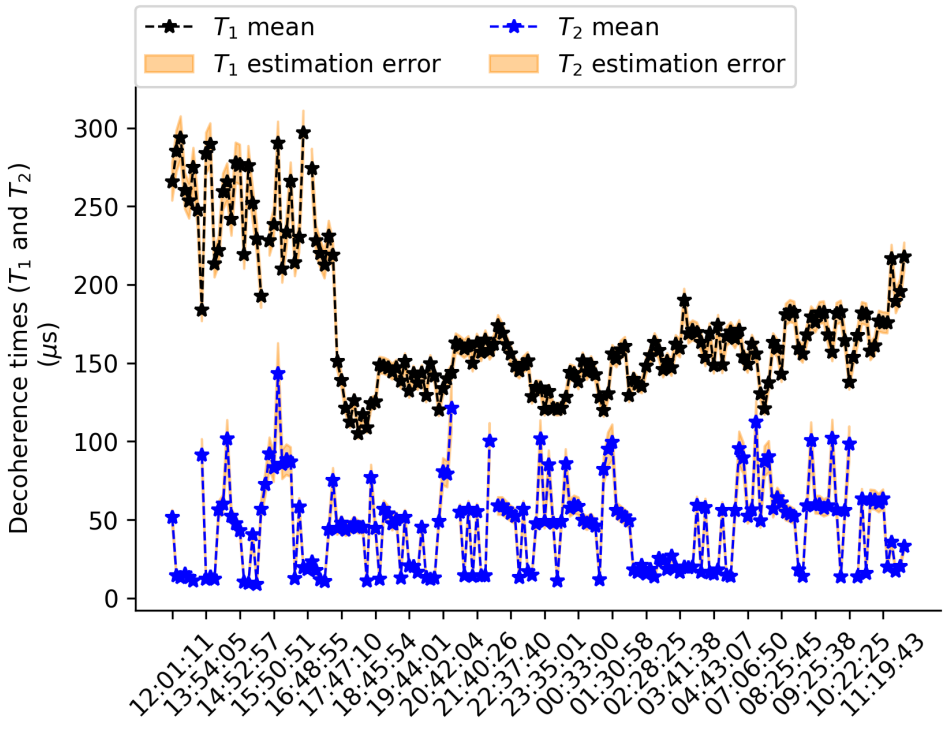


(b)

Figure 2.7: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 5 and (b) qubit 6.

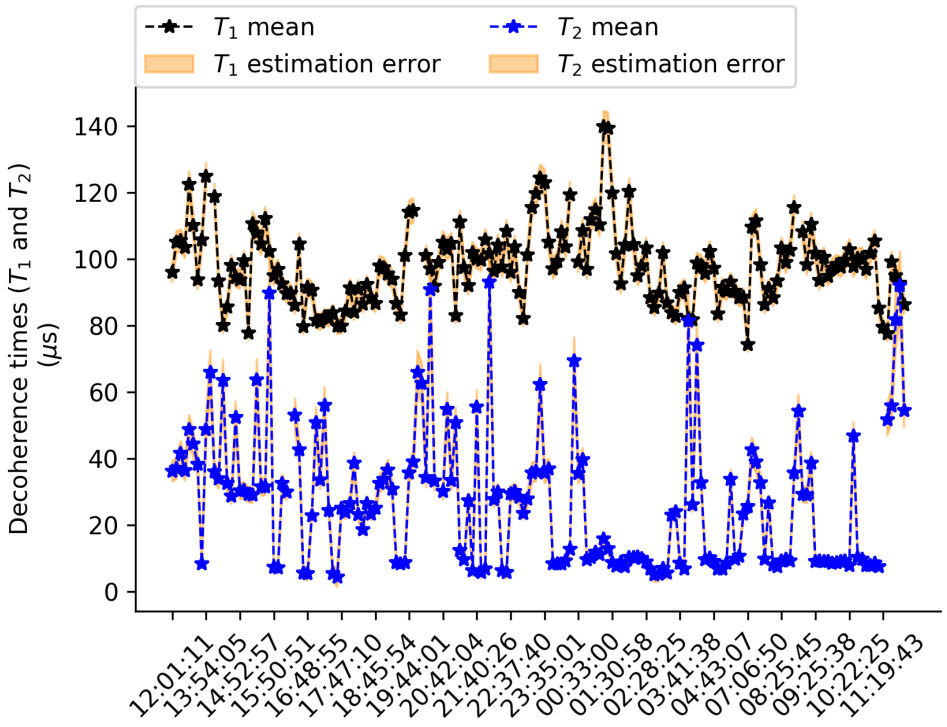


(a)

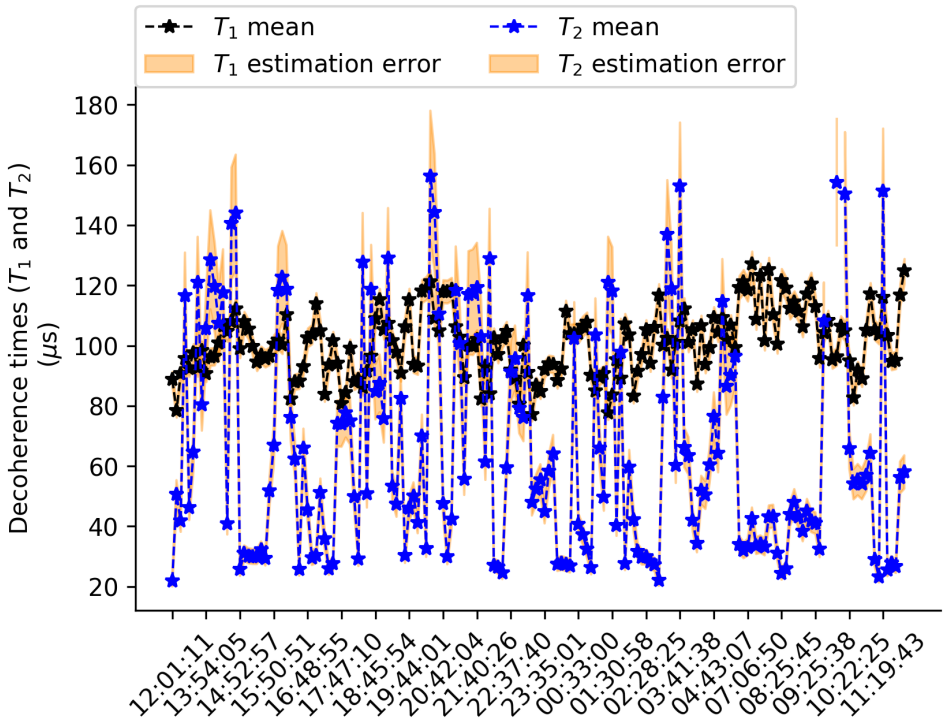


(b)

Figure 2.8: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 7 and (b) qubit 8.

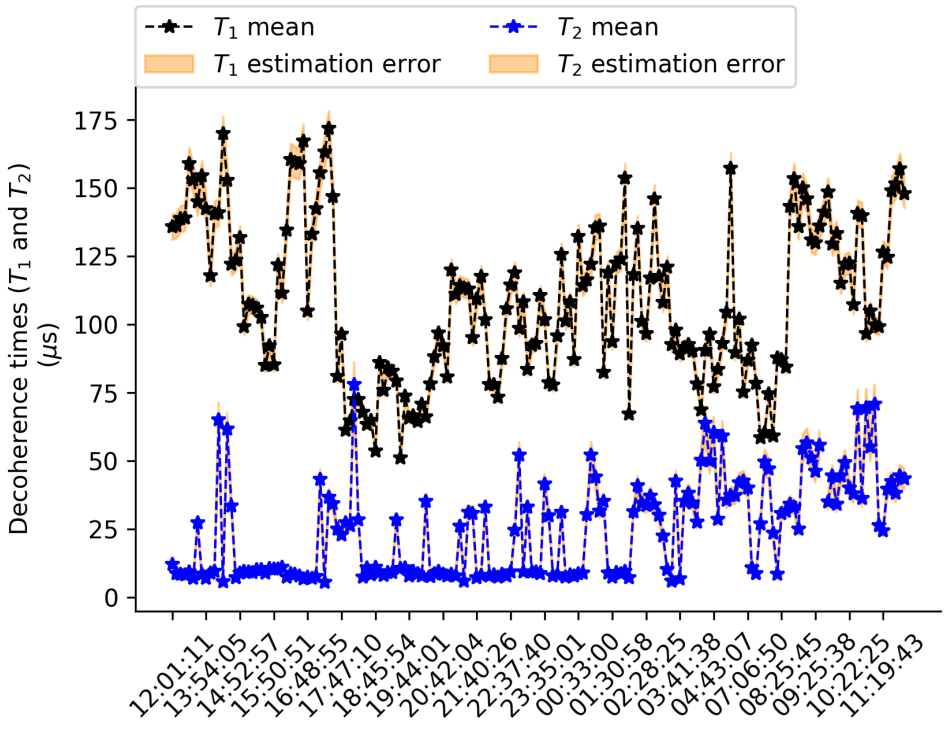


(a)

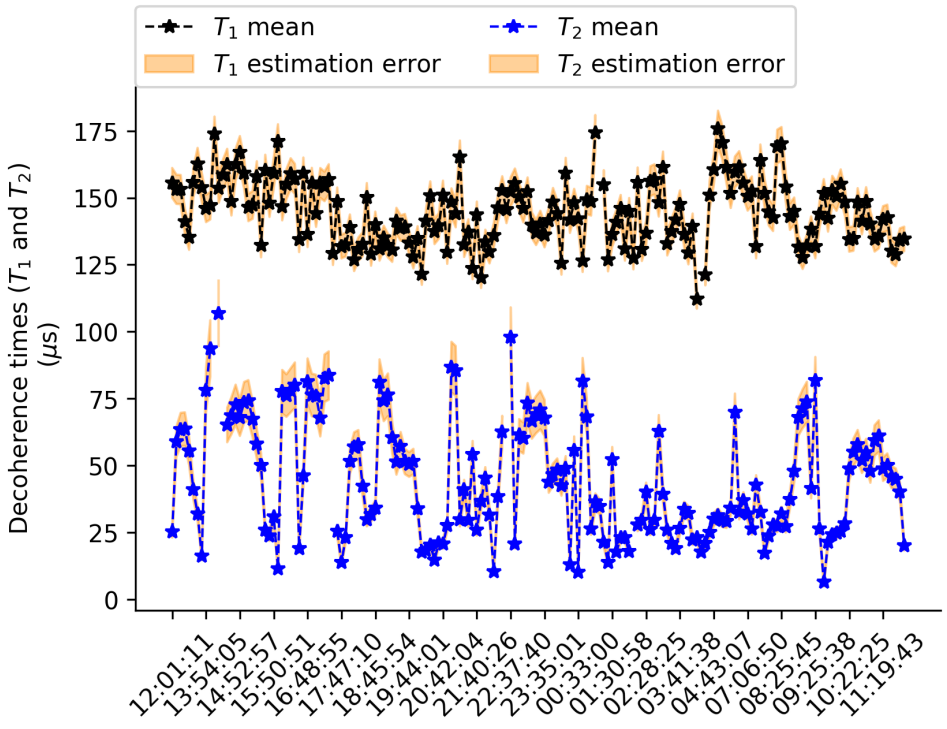


(b)

Figure 2.9: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 9 and (b) qubit 10.

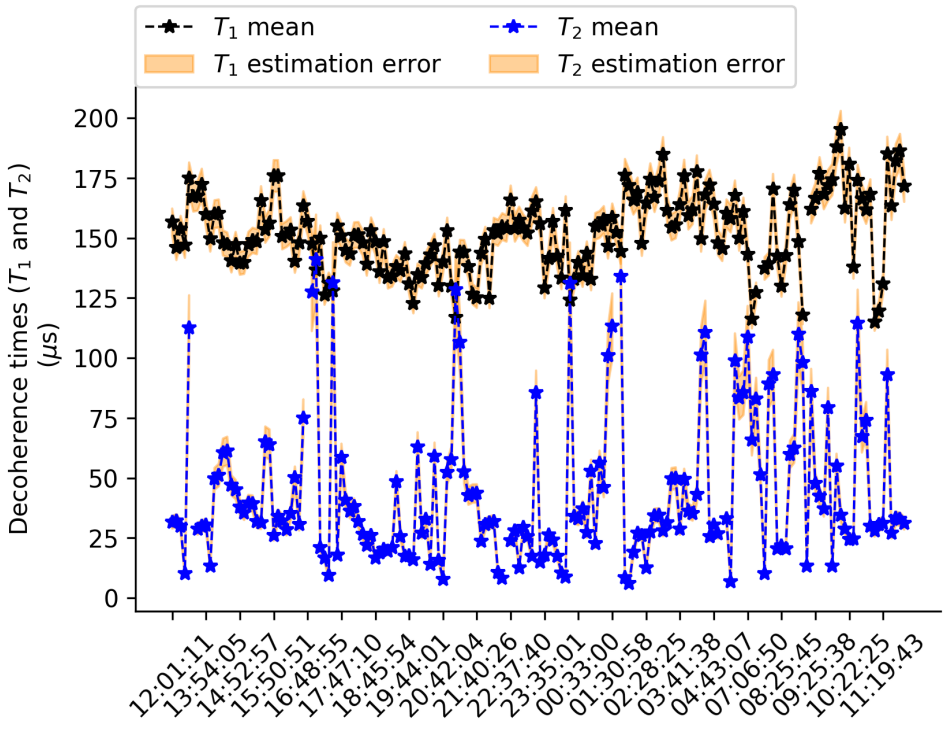


(a)

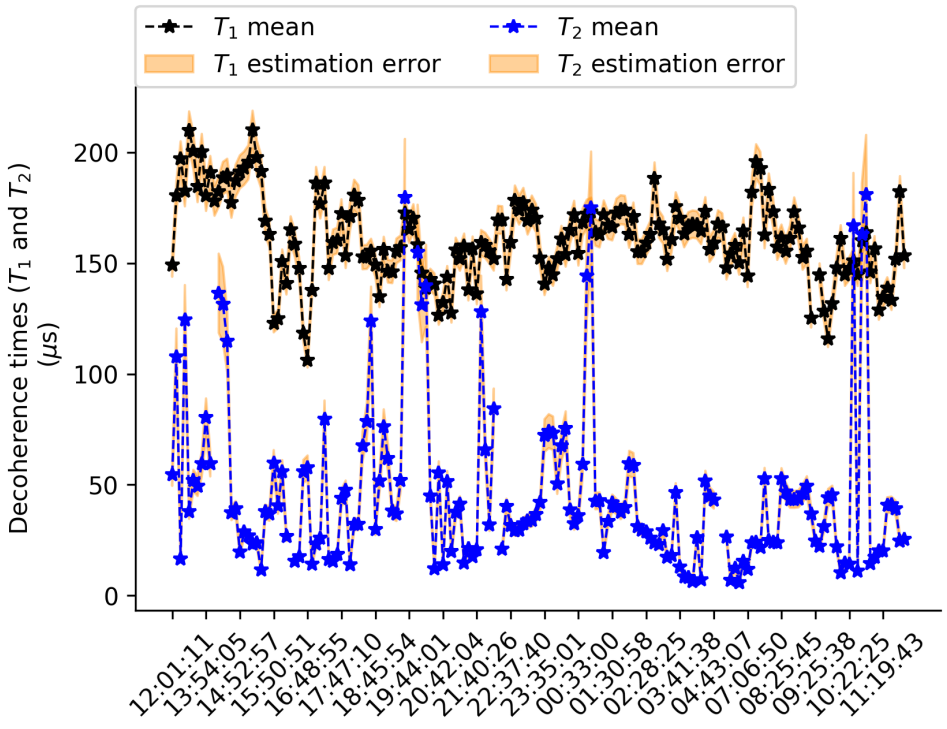


(b)

Figure 2.10: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 11 and (b) qubit 12.

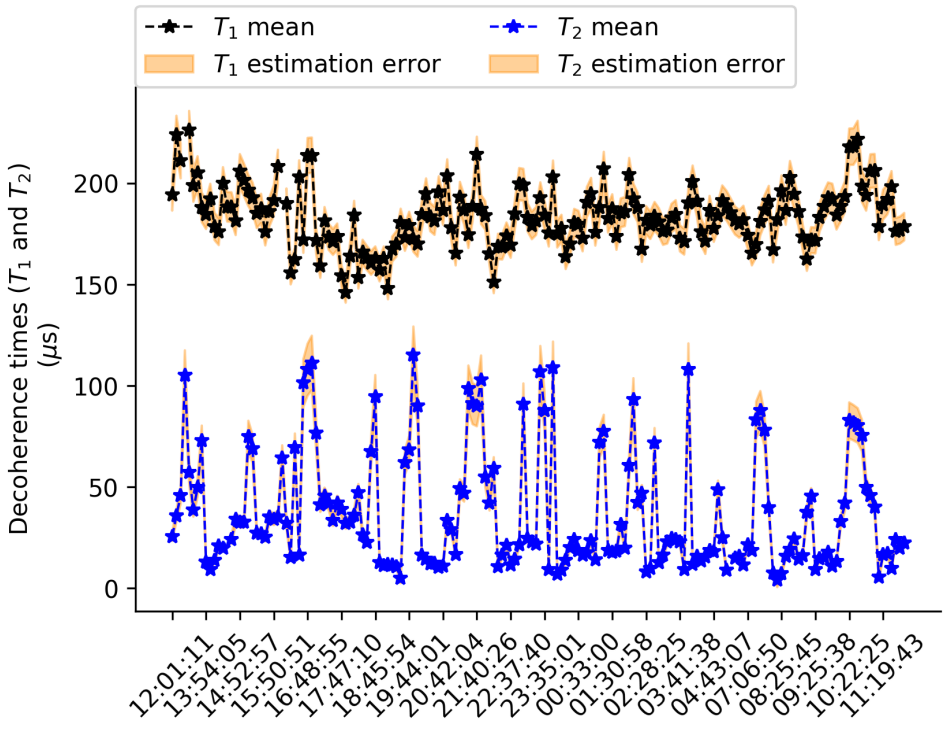


(a)

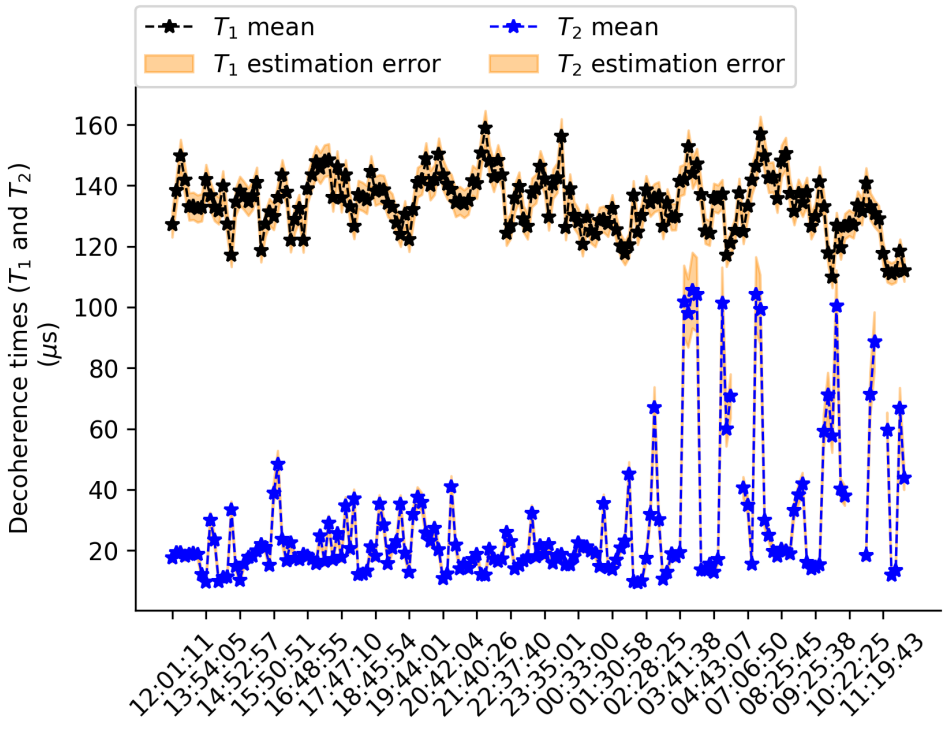


(b)

Figure 2.11: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 13 and (b) qubit 14.

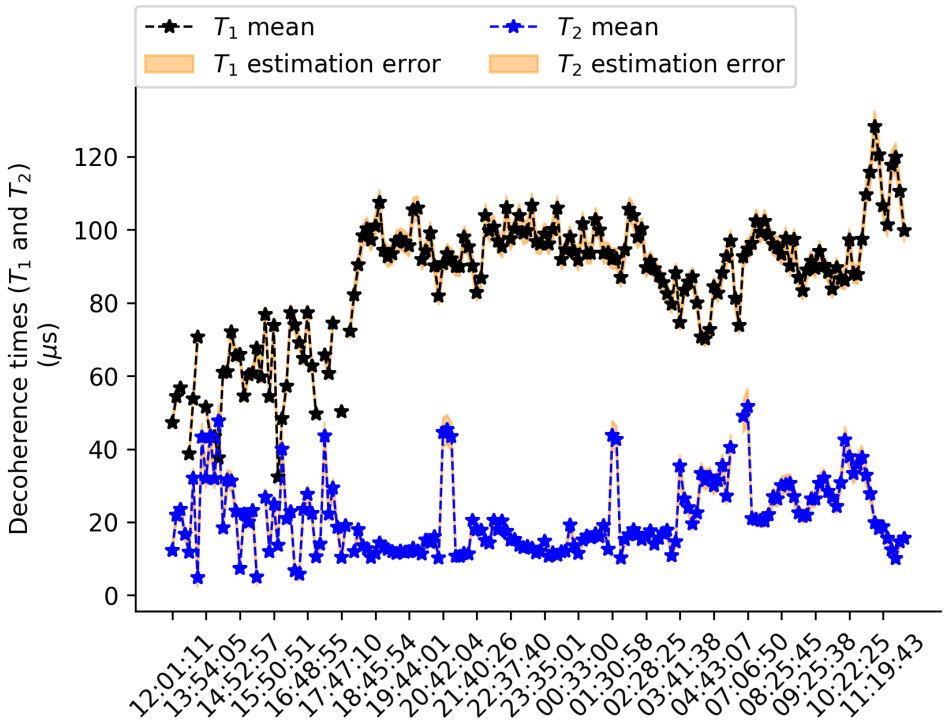


(a)

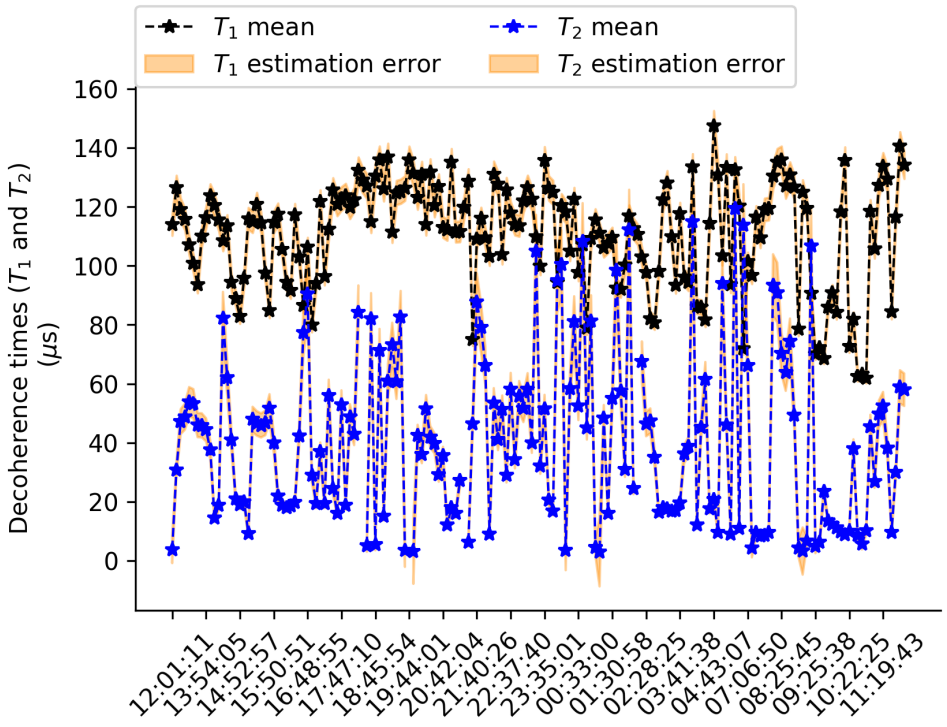


(b)

Figure 2.12: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 15 and (b) qubit 16.

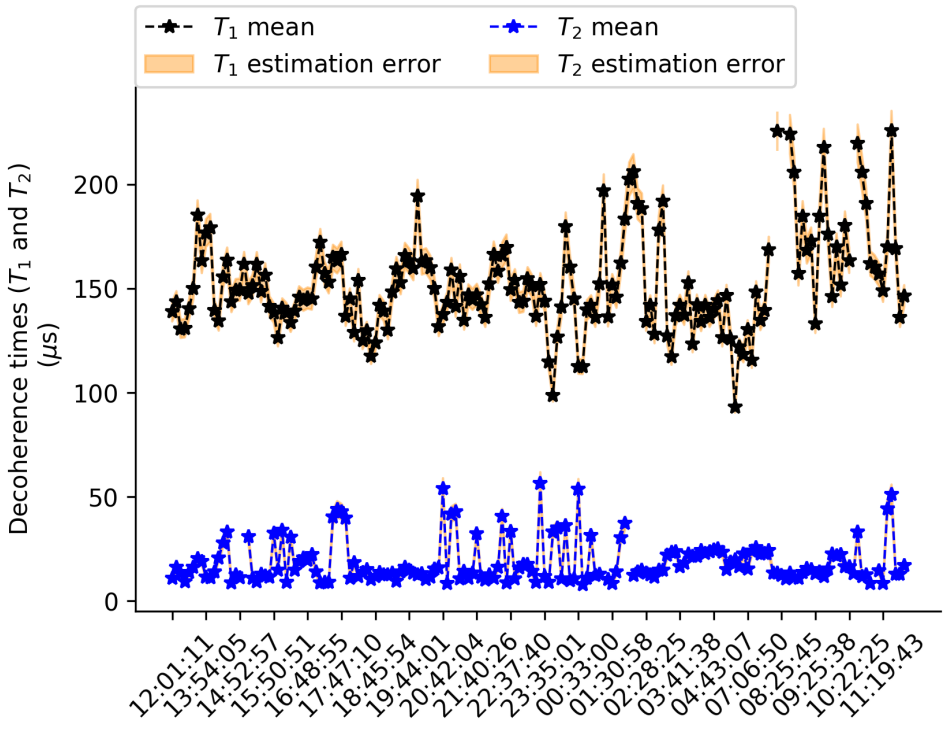


(a)

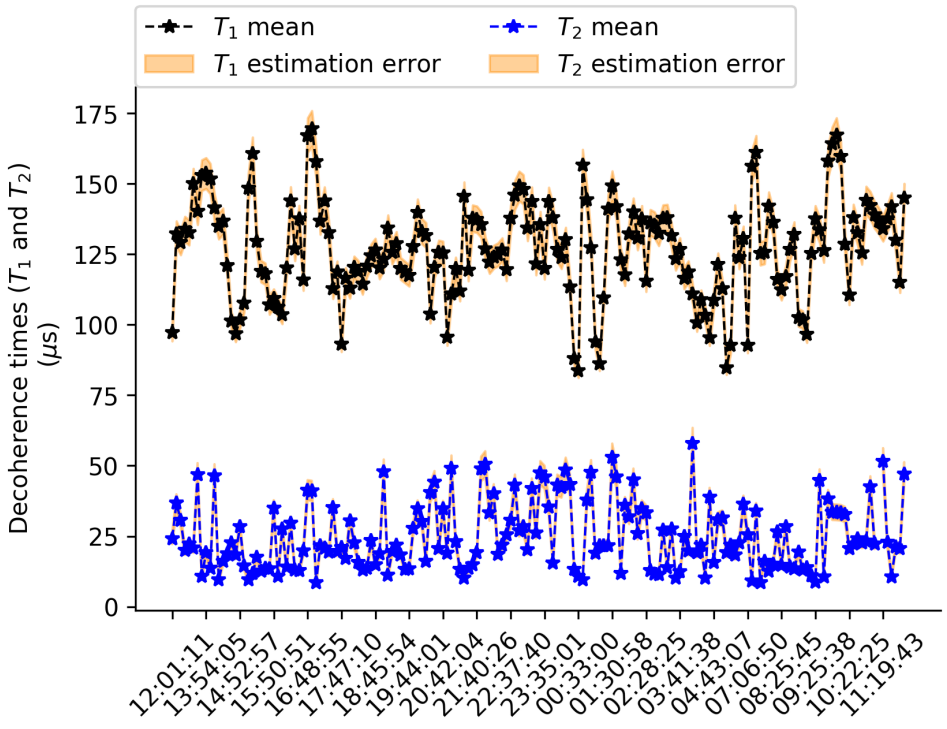


(b)

Figure 2.13: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 17 and (b) qubit 18.

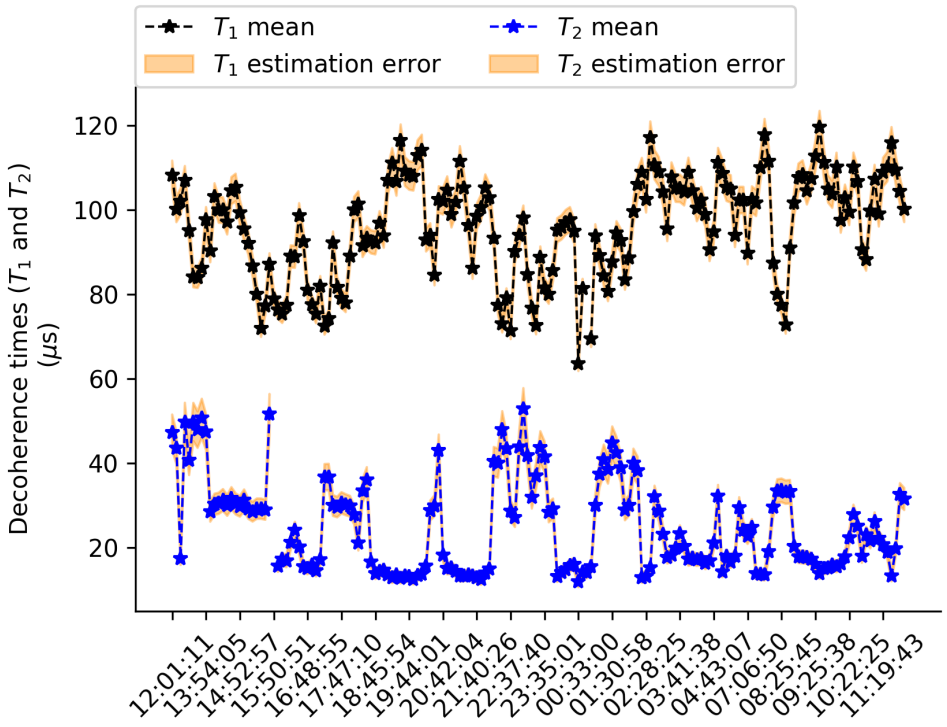


(a)

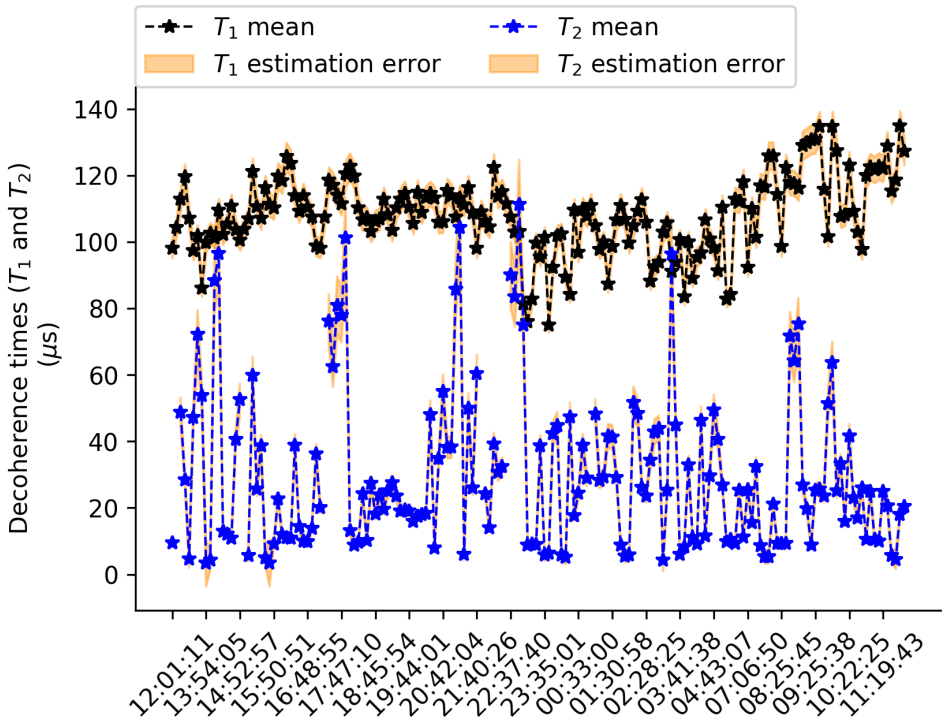


(b)

Figure 2.14: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 19 and (b) qubit 20.

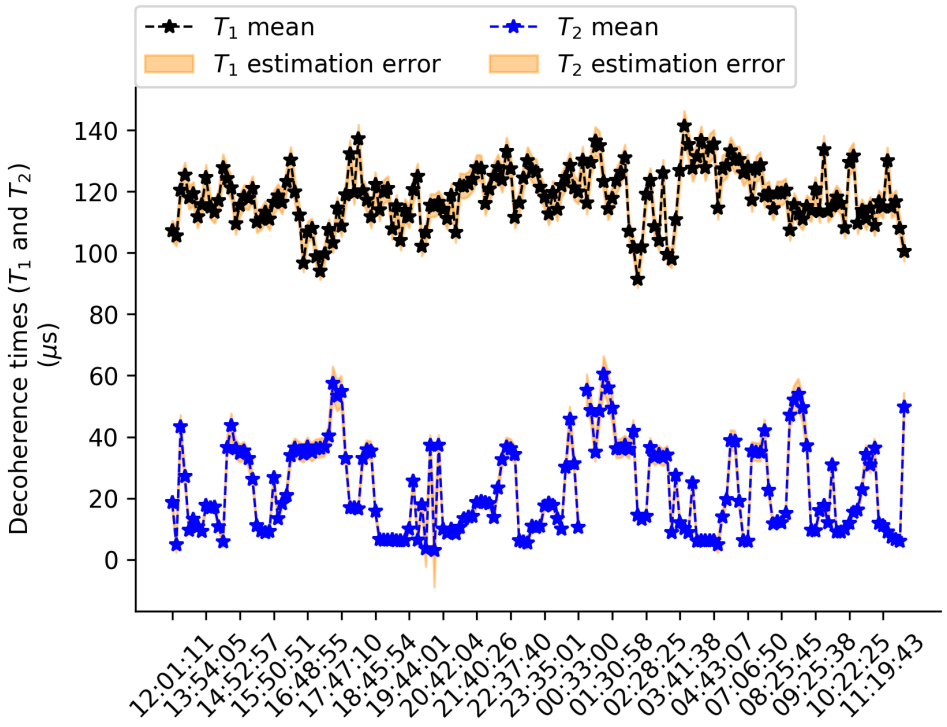


(a)

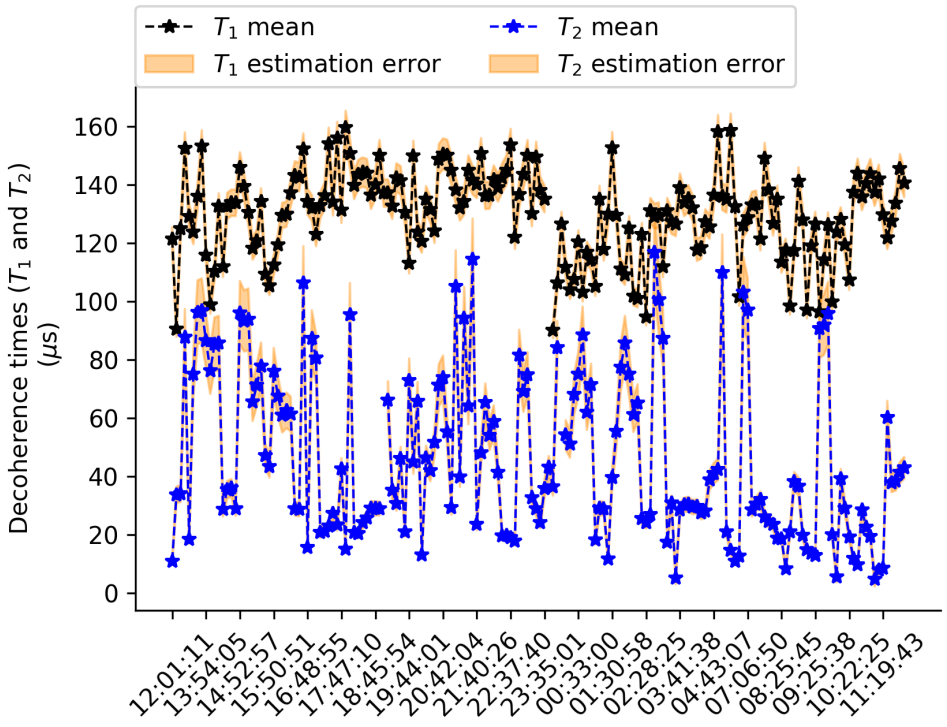


(b)

Figure 2.15: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 21 and (b) qubit 22.

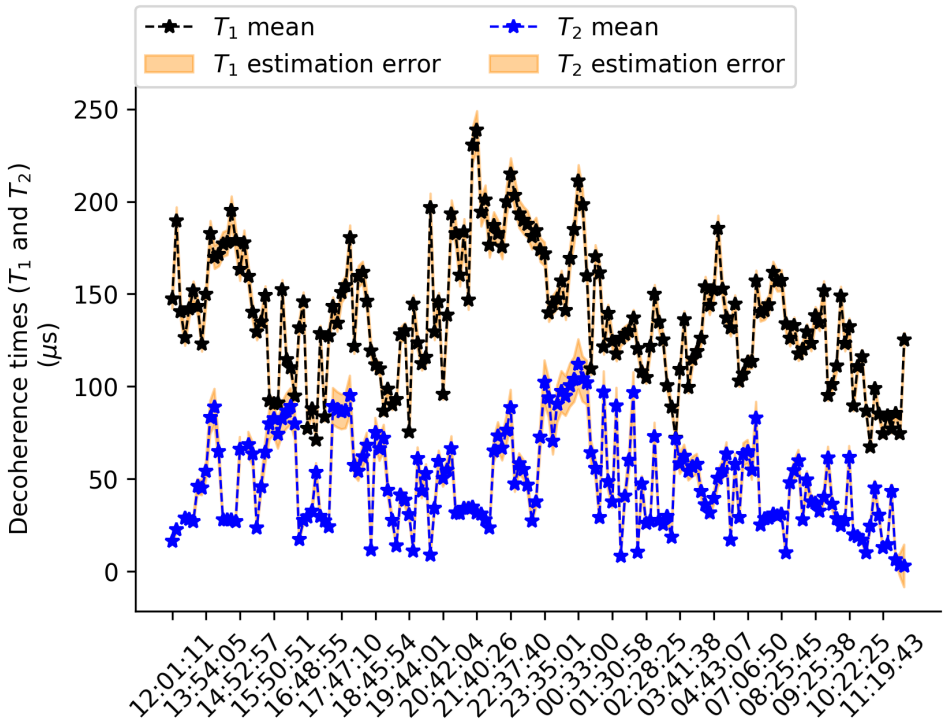


(a)

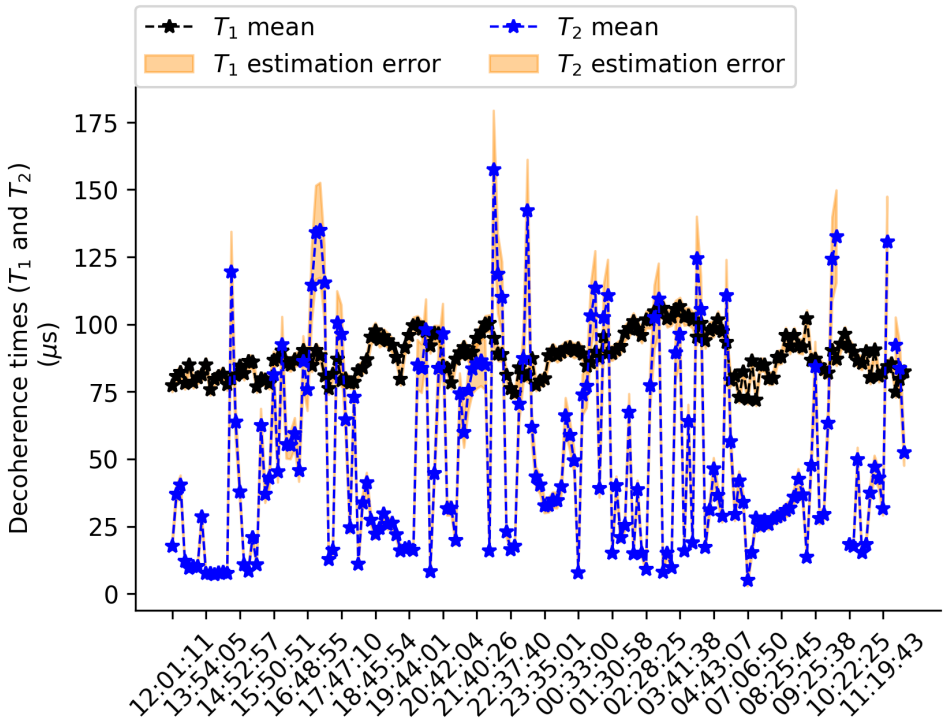


(b)

Figure 2.16: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 23 and (b) qubit 24.



(a)



(b)

Figure 2.17: Estimated T_1 and T_2 time-series as collected between 12:00 P.M. ET on Sep 12, 2023 and 12:00 P.M. ET on Sep 13, 2023 for (a) qubit 25 and (b) qubit 26.

Chapter 3

Performance evaluation framework

Performance evaluation of noisy quantum computing in crucial for several reasons. Firstly, quantum computing is still in its early stages of development [55, 56, 57], and understanding the sources of errors and noise[58] is vital to improve the performance of quantum computers. By conducting rigorous performance evaluation, researchers can identify and quantify the various sources of noise, such as decoherence, gate errors, and readout errors. Secondly, this understanding is essential for developing error mitigation techniques [59], which are necessary for scaling up small-scale quantum computations. Thirdly, reproducibility of results from quantum computing is critical. Rigorous performance evaluation ensure that experiments can be replicated by other researchers, contributing to validation and verification of quantum algorithms.

However, the task is not simple as the same complexity that gives quantum technology an advantage over classical computing also hinders its rigorous checking [54]. Reasons include the inherent randomness in quantum results due to the Born rule, error accumulation without clear source attribution, the curse of dimensionality [52], and the inability to step-through program execution in quantum circuits [53].

Also, the diverse range of terms encountered in quantum computing today can blur the distinctions between them, making it challenging to appreciate their nuanced differences. Examples include verification[60, 61, 62] (which pertains to ensuring correct transpilation), validation[63, 64] (which can have two connotations: (a) validating correctness of output by comparison to theory or concurrent classical simulation (akin to accuracy), or validating the quantum nature of a device), benchmarking[65, 66, 67] (which involves assigning a performance measure, often a simple scalar number, to a quantum

processor/ subsystem/ subroutine, with reproducibility as a key defining characteristic), accreditation [90, 68] and certification [54].

3.1 Distance measures

In this chapter, our focus is on the development of a systematic performance evaluation framework for the outcomes from noisy quantum computers. To accomplish this, it is necessary to compare probability distributions. Various options exist for quantifying the distance between high-dimensional densities, including Hellinger distance[91], total variation distance[92], Kolmogorov-Smirnov statistic[93], Mahalanobis distance[94], Wasserstein metric[95], Levy-Prokhorov metric[96], and non-metric divergence measures like Kullback-Leibler divergence[97], Jensen-Shannon divergence[98], Renyi's divergence[97] and Tsallis divergence[99] for quantum states. Metric measures are advantageous because they obey the triangle inequality, enabling rigorous comparisons, while non-metric measures (often called divergence) are useful for obtaining performance bounds in specific problem settings. However, all distance measures suffer from the curse of dimensionality[52], which results in exponentially increasing resource requirements to accurately represent information as the number of qubits in the quantum system grows [100, 101].

3.1.1 Hellinger distance

In our work, we primarily employ the Hellinger distance. The Hellinger distance between two probability distributions $f_X(\mathrm{x})$ and $f_Y(\mathrm{x})$ for the random variables X and Y is defined by:

$$H(f_X, f_Y) = \sqrt{1 - BC(f_X, f_Y)}, \tag{3.1}$$

where the Bhattacharyya coefficient BC is

$$BC(f_X, f_Y) = \int_{\mathrm{x}} \sqrt{f_X(\mathrm{x}) f_Y(\mathrm{x})} d\mathrm{x}. \tag{3.2}$$

The Hellinger distance provides a practical and meaningful approach to measuring the similarity of distributions. Firstly, it operates directly on observed data, eliminating the need to compute intermediate abstractions like entropy. Secondly, the Hellinger distance is easy to interpret and apply in practical scenarios. Thirdly, it can be easily extended to quantum states through Tsallis divergence [99]. Lastly, the Hellinger distance is proportional to the Fisher information[102], which quantifies the partial knowledge a density carries about some unknown.

In particular, suppose x is a realization of the random noise parameter X drawn from the time-varying distribution $f_X(\mathrm{x};t)$, where t denotes time. Let,

$$H_X(t_1,t_2) = H(f_X(t_1), f_X(t_2)), \tag{3.3}$$

measure the distance between the densities of X at time t_1 and t_2.

For example, the Hellinger distance between two beta distributions (used for SPAM and CNOT fidelity characterizations) is given by:

$$\begin{aligned} f(\mathrm{x};\alpha_1,\beta_1) &= \frac{\mathrm{x}^{\alpha_1-1}(1-\mathrm{x})^{\beta_1-1}}{\mathrm{Beta}(\alpha_1,\beta_1)}, \quad g(\mathrm{x};\alpha_2,\beta_2) = \frac{\mathrm{x}^{\alpha_2-1}(1-\mathrm{x})^{\beta_2-1}}{\mathrm{Beta}(\alpha_2,\beta_2)} \\ d_H &= \sqrt{1-\int_0^1 \sqrt{f(\mathrm{x};\alpha_1,\beta_1)g(\mathrm{x};\alpha_2,\beta_2)}d\mathrm{x}} = \sqrt{1-\frac{\mathrm{Beta}(\alpha_1+\alpha_2-1,\beta_1+\beta_2-1)}{\mathrm{Beta}(\alpha_1,\beta_1)\mathrm{Beta}(\alpha_2,\beta_2)}} \end{aligned} \tag{3.4}$$

while that between two gamma distributions (used for duty cycle characterization) is given by:

$$\begin{aligned} f(\mathrm{x};m_1,\alpha_1) &= \frac{1}{\Gamma(m_1)}\mathrm{x}^{m_1-1}\alpha_1^{m_1}e^{-\alpha_1\mathrm{x}}, \quad g(\mathrm{x};m_2,\alpha_2) = \frac{1}{\Gamma(m_2)}\mathrm{x}^{m_2-1}\alpha_2^{m_2}e^{-\alpha_2\mathrm{x}} \\ d_H &= \sqrt{1-\int_0^1 \sqrt{f(\mathrm{x};n,\alpha)g(\mathrm{x};m,\beta)}d\mathrm{x}} = \sqrt{1-\frac{\alpha_1^{m_2}\alpha_2^{m_1}}{(\alpha_1+\alpha_2)^{m_1+m_2-1}}\frac{\Gamma(m_1+m_2-1)}{\Gamma(m_1)\Gamma(m_2)}} \end{aligned} \tag{3.5}$$

For non-standard distributions $f(x)$ and $g(x)$ where an analytical closed form solution is not available, we numerically integrate the empirical distributions using the standard trapezoidal method:

$$d_H^2 - 1 = \int_a^b \sqrt{f(x)g(x)}dx = \lim_{n\to\infty}\sum_{i=0}^{n-1}\left[\frac{1}{2}\left[f(a-ih/2)g(a-ih/2) + f(a+ih/2)g(a+ih/2)\right]\right]^{1/2} h \tag{3.6}$$

where $h = \frac{b-a}{n}$.

Despite its ease of interpretation, the Hellinger distance scales exponentially in the number of noise parameters. This has the effect that even small changes in a distribution yield large changes in the distance value. This is called the curse of dimensionality. To see this, consider d independent and identically distributed noise parameters $\{\mathrm{x}_1,\cdots,\mathrm{x}_d\}$, whose marginals are given by $f_{X_i}(\mathrm{x};t)$. Let h be the Hellinger distance between the marginals at time t_1 and t_2. Thus,

$$H_{X_i}(t_1,t_2) = h \quad \forall i \tag{3.7}$$

Since the parameters are independent:

$$\begin{aligned} \log\left(1 - H_X^2\right) &= -d|\log(1-h^2)| \\ \Rightarrow H_X &= \sqrt{1 - \exp\left[-d|\log(1-h^2)|\right]}. \end{aligned} \tag{3.8}$$

Thus the distance approaches 1 quickly as the number of dimensions increases.

A more sensitive measure can be defined using H_{avg}, defined as the average over the distances for the d univariate (X_k) marginal distributions:

$$H_{\text{avg}}(t_1, t_2) = \frac{1}{d}\sum_{k=1}^{d} H_{X_k}(t_1, t_2). \tag{3.9}$$

When the joint distributions are time-invariant, then the marginals must also be time-invariant, resulting in a small average value for H_{avg}. This test is more sensitive as it mitigates the curse of dimensionality and offers higher dispersion for improved calibration.

Another sensitive approach is to normalize the distance relative to the dimensionality d of the distribution:

$$H_{\text{normalized}}(t_1, t_2) = \sqrt{1 - BC^{1/d}}. \tag{3.10}$$

We refer to this statistic as the normalized Hellinger distance (note that although we call it distance, this statistic is not technically a metric as it does not satisfy the triangle inequality).

3.1.2 Moment-Based Distance

We also developed a new disance measure as part of our research which we call Moment-Based Distance (MBD). The key advantage of MBD is its ability to incorporate the geometric shape of the underlying noise distribution while still being a metric. Thus, it takes into account higher order effects like kurtosis and skewness. Specifically, we define the moment-based metric (d) between two histograms (f and g) based on the equality of their moments.

$$d(f, g) = \sum_{m=0}^{\infty} S_m(f, g) \tag{3.11}$$

where

$$S_m(f, g) = \frac{1}{(m)!}\int_a^b \left|\left(\frac{x}{\gamma}\right)^m (f(x) - g(x))\right| dx \tag{3.12}$$

and,

$$\gamma = \max(|a|, |b|) \tag{3.13}$$

for a bounded real variable x. Here, a and b are the minimum and maximum values of x and can be derived from theoretical considerations (e.g., when the random variable is a probability then $\gamma = 1$) or from empirical histogram data.

The moment-based-distance $d(f, g)$ satisfies the following properties:

1. $d(f, g) \geq 0$ follows from the definition of d.

2. $d(f, g) = d(g, f)$ follows from the definition of d.

3. $d(f, g) = 0$ iff $f(x) = g(x)$.

   Proof: If $f(x) = g(x)$, then $d = 0$ because $S_m = 0$ for every m. Conversely, if $d = 0$, then $S_m = 0$ for all m.If $S_m = 0$, then for all x, the integrand must satisfy

   $$\left| \left(\frac{x}{\gamma} \right)^m (f - g) \right| = 0$$

   As $(x)^m \neq 0$ for all x, it must be that $|f(x) - g(x)| = 0$ for all x and, hence, $f(x) = g(x)$. ■

4. $d(f, g) \leq d(f, h) + d(h, g)$

   Proof: For every m,

   $$\begin{aligned} S_m(f,g) &= \int_a^b \left| \left(\frac{x}{\gamma} \right)^m \frac{1}{m!} (f(x) - g(x)) \right| dx \\ &= \int_a^b \left| \left(\frac{x}{\gamma} \right)^m \frac{1}{m!} (f(x) - h(x) \right. \\ &\qquad \left. + h(x) - g(x)) \right| dx \\ &\leq \int_a^b \left| \left(\frac{x}{\gamma} \right)^m \frac{f(x) - h(x)}{m!} \right| dx \\ &\quad + \int_a^b \left| \left(\frac{x}{\gamma} \right)^m \frac{h(x) - g(x)}{m!} \right| dx \\ &\leq S_m(f,h) + S_m(h,g) \end{aligned}$$

   and whence the sum satisfies the inequality as well. ■

5. The series $d = S_0 + S_1 + S_2 + \cdots$ converges.

   Proof: The distance d converges if, after some fixed term, the ratio of each term to the preceding term is less than some quantity r, which is itself numerically less than unity. If $S_{m+1} < S_m$ for

all $m \geq 1$, then

$$\begin{aligned} d =& S_0 + S_1 + S_1 \cdot \frac{S_2}{S_1} + S_1 \cdot \frac{S_2}{S_1} \cdot \frac{S_3}{S_2} + \cdots \\ <& S_0 + S_1(1 + r + r^2 + r^3 + \cdots) \\ =& S_0 + \frac{S_1}{1-r} \text{ since } r < 1 \end{aligned}$$

To prove that $S_{m+1} < rS_m$ for $m \geq 1$, we proceed as follows:

$$\begin{aligned} S_{m+1} &= \int_a^b \left| \left(\frac{x}{\gamma}\right)^{m+1} \frac{1}{(m+1)!}(f(x) - g(x)) \right| dx \\ &= \int_a^b \left| \frac{x}{\gamma(m+1)} \right| \left| \left(\frac{x}{\gamma}\right)^{m} \frac{f(x) - g(x)}{m!} \right| dx \\ &\leq \left| \frac{x}{\gamma} \right|_{max} \frac{1}{m+1} S_m \\ &\leq \frac{1}{m+1} S_m \text{ since } \left| \frac{x}{\gamma} \right|_{max} = 1 \\ &\leq \frac{1}{2} S_m \text{ since } m \geq 1 \end{aligned}$$

where $|x|_{max}$ is the maximum of x. ■

An important consequence of the latter convergence property is that the moment-based distance satisfies the practical requirement that lower-order moments contribute more than higher-order moments to the distance (for $m > 1$). This proves essential to our subsequent use of the moment-based distance below, as we rely on the approximate distance defined to order n as

$$d_n = \sum_{m=0}^{n} S_m \tag{3.14}$$

We next present a series of simulation studies to develop intuition for how the moment-based distance behaves in the presence of both stable and unstable distributions. In particular, we will show that moment-based distance is small but non-zero for distributions that are similar but not identical, while such deviations grow with dissimilarity. For our studies, we computed the distance of 10 different distributions with respect to a reference distribution. Table 3.1 summaries the list of tests as well as their moment-based distance from the reference normal distribution $\mathcal{N}(\mu, \sigma)$. For testing purpose, the parameters are $\mu = 0.4, \Delta = 0.2$ and $\sigma = 0.04$. We note that, as expected, the distribution 'closest' to $\mathcal{N}(\mu, \sigma)$ is $\mathcal{N}(1.01\mu, \sigma)$ and the 'farthest' are $\mathcal{N}(2\mu, \sigma)$ and $\mathcal{N}(\mu + 2\Delta, 2\sigma)$.

Table 3.1: Moment-based distance by Distribution

Distribution	d_4	d_{20}	Error(%)
$N(\mu, \sigma)$	0.00000	0.00000	NA
$N(\mu + \Delta, \sigma)$	2.70868	2.70876	-0.00289
$N(\mu, 2\sigma)$	0.83252	0.83253	-0.00104
$N(\mu, 4\sigma)$	1.47301	1.47304	-0.00180
$N(2\mu, \sigma)$	2.93489	2.93520	-0.01033
$N(\mu, 1.5\sigma)$	0.49215	0.49216	-0.00091
$N(1.01\mu, \sigma)$	0.11739	0.11740	-0.00079
$SkewedNormal(\mu, 2\sigma)$	0.80887	0.80888	-0.00140
$Gumbel(\mu, 2\sigma)$	0.95131	0.95134	-0.00246

We next study how the order of the series d_n increases the accuracy of the distance measured. In our simulation studies of well-defined distributions, we find that d_n converges for $n = 4$ when the distributions are sufficiently dissimilar. As shown in Fig. 3.5, the relative contributions of each S_m to d_n decreases with increasing m as expected from the convergence property. Thus, $m = 0$ accounts for about 60% of the total distance while $m = 1$ accounts for 90% and $m = 2$ reaches 98%. For $m = 4$, d_m is nearly 100% of the d_∞. Consequently, we will consider $m = 4$ sufficient to accurately characterize the moment-based distance for the remainder of our analysis. This is certainly an approximation in the sense that two histograms which start to differ only after the fourth order moment will be erroneously classified as same. Is d_4 still a valid *distance metric*? Yes. A glance at the proofs will reveal that properties (1) to (4) are still satisfied when we truncate the d series at a finite m (say m=4). Moreover, it converges too (i.e. Property (5) is satisfied too) because a finite number of terms (in this case 5 terms) is by definition convergent when the individual terms are finite. The latter is true because each S_m is bounded between finite a and b as per Equation (2).

As a point of comparison, we contrast the moment-based distance to total variation distance (TVD), a state-of-the-art metric which has proven useful in earlier experimental investigations [103, 104]. We note that the magnitudes of the moment-based distance and total variation distance are not directly comparable as they follow very different methodologies but one can compare the corresponding signal-to-noise ratio (SNR) of the two metrics as the inverse of the coefficient of variation. For our numerical studies, we generated two time series, each of length 8192, by sampling two different probability distributions. The first was a normal distribution with mean 10 and standard deviation 1, and the second a normal distribution with mean 10 and standard deviation 4. We calculate the moment-based distance and total variation distance between these two time series, and then we repeated this numerical experiment 400 times to generate a distribution of the TVD and MBD

distances. Using the average μ and standard deviation σ of these distributions, we calculated the respective SNR as

$$\text{SNR} = \frac{\mu}{\sigma} \tag{3.15}$$

As shown in Fig. 3.6, our results indicate that the moment-based distance has more statistical power as indicated by a higher SNR. As an aside, a practical concern is the dependence of precision of the moment-based distance on sampling size. Although each S_m should vanish when two distributions are similar, finite sampling lead to approximations and ultimately a non-zero distance. As shown in Fig. 3.7, increasing sampling may be used to reduce the relative error in each moment to a desired relative precision. Since MBD lacks direct comparability with measures such as Fisher information (which have deep physical interpretations), we primarily use the Hellinger distance for the rest of the document.

3.2 Evaluation framework

Let ρ be a density matrix representing the state of an n-qubit quantum register. Suppose ρ undergoes a unitary transformation U, which can be decomposed into K unitaries:

$$U = U_K \cdots U_1. \tag{3.16}$$

The noiseless output state of the quantum register is:

$$\rho_{\text{out}}^{\text{ideal}} = U\rho U^\dagger. \tag{3.17}$$

The projection operators $\{\Pi_i = |i\rangle\langle i|\}$ project the output state into one of the 2^n computational basis states $\{|0\rangle, \cdots, |2^n - 1\rangle\}$. The probability distribution for the results generated by a noiseless quantum computer is denoted by:

$$\mathbb{P}^{\text{ideal}} = \{p_i^{\text{ideal}}\} \quad i \in \{0, 1, \cdots, 2^n - 1\} \tag{3.18}$$

where $p_i^{ideal} = \text{Tr}[\Pi_i \rho_{\text{out}}^{\text{ideal}}]$. In general, it is not efficient to construct the set $\mathbb{P}^{\text{ideal}}$ using classical computing as the resource needs scale exponentially with n. However, such demanding calculations are feasible if either $n < 50$ or if the circuit has exhibits high-degree of structure (such as the quantum search).

In the presence of noise, the evolution of the quantum register no longer adheres to a unitary evolution[11]. This leads to mixed states in the output. Let $\mathcal{E}$ denote the super-operator that

characterizes a noisy quantum channel. It can be defined by a set of Kraus operators $\{M_k\}$. In particular,

$$\rho_{\text{out}}^{\text{noisy}} = \mathcal{E}_K \left(\cdots \mathcal{E}_2 \left(U_2 \mathcal{E}_1 \left(U_1 \rho_{\text{in}} U_1^\dagger \right) U_2^\dagger \right) \cdots \right) \tag{3.19}$$

where the action of each $\mathcal{E}_k$ is given by:

$$\mathcal{E}_k(\rho) = \sum_k M_k \rho M_k^\dagger. \tag{3.20}$$

We will sometimes use the notation $\mathcal{E}_{\mathrm{x}}$ to emphasize the dependence of the error channel on a vector of noise parameters $(\mathrm{x}_1, \cdots, \mathrm{x}_d)$.

The corresponding probability distribution for a noisy computer is:

$$\mathbb{P}^{\text{noisy}} = \{p_i^{\text{noisy}}\} \text{ for } i \in \{0, 1, \cdots, 2^n - 1\} \tag{3.21}$$

where $p_i^{\text{noisy}} = \text{Tr}[M_i^\dagger M_i \rho_{\text{out}}^{\text{ideal}}]$, and M_i is the measurement operator for a noisy readout channel [105]:

$$\begin{aligned} M_0 =& \sqrt{f_0} \left|0\right\rangle \left\langle 0\right| + \sqrt{1 - f_1} \left|1\right\rangle \left\langle 1\right| \\ M_1 =& \sqrt{1 - f_0} \left|0\right\rangle \left\langle 0\right| + \sqrt{f_1} \left|1\right\rangle \left\langle 1\right| \end{aligned} \tag{3.22}$$

The Hellinger distance between $\mathbb{P}^{ideal}$ and $\mathbb{P}^{noisy}$ is:

$$H(\mathbb{P}^{\text{ideal}}, \mathbb{P}^{\text{noisy}}) = \sqrt{1 - BC(\mathbb{P}^{\text{ideal}}, \mathbb{P}^{\text{noisy}})} \tag{3.23}$$

with the Bhattacharyya coefficient $BC(\mathbb{P}^{\text{ideal}}, \mathbb{P}^{\text{noisy}}) \in [0, 1]$ defined as:

$$BC(\mathbb{P}^{\text{ideal}}, \mathbb{P}^{\text{noisy}}) = \sum_{i=0}^{2^n - 1} \sqrt{p_i^{\text{ideal}} p_i^{\text{noisy}}}. \tag{3.24}$$

Next, we turn our attention to the notation for the mean of a quantum observable as an outcome of a noisy quantum computer. Let $\hat{O}$ symbolize the operator associated with an observable computed from the results of the quantum circuit. The operator can be broken down into its spectral decomposition:

$$O = \sum_m \lambda_m \left|\lambda_m\right\rangle \left\langle \lambda_m\right|, \tag{3.25}$$

where λ_m represents the real eigenvalues of O and $\left|\lambda_m\right\rangle$ denotes the corresponding eigen-states. The expectation of the observable O, relative to the noisy quantum state described by the density matrix

$\rho_{\text{out}}^{\text{noisy}}$, is given by:

$$\langle O_{\text{x}} \rangle = \text{Tr}\left(\hat{O}\rho_{\text{out}}^{\text{noisy}}\right) = \text{Tr}\left(\hat{O}\mathcal{E}_{\text{x}}(\rho_{\text{out}}^{\text{ideal}})\right) = \sum_m \lambda_m \text{Tr}\left(\Pi_m \mathcal{E}_{\text{x}}(\rho_{\text{out}}^{\text{ideal}})\right), \tag{3.26}$$

where $\Pi_m = |\lambda_m\rangle \langle\lambda_m|$ stands as the projective operator.

For example, consider the case of a register with $n = 1$ qubits in the presence of depolarizing noise. The latter channel operator is characterized by a noise parameter x for which the Kraus operators $M_k \in \{\sqrt{1-\text{x}}\mathbb{I}, \sqrt{\text{x}}\mathbb{X}, \sqrt{\text{x}}\mathbb{Y}, \sqrt{\text{x}}\mathbb{Z}\}$ yield

$$\rho_{\text{out}}^{\text{noisy}} = \mathcal{E}_{\text{x}}(\rho) = (1-\text{x})\rho + \frac{\text{x}}{3}\mathbb{X}\rho\mathbb{X} + \frac{\text{x}}{3}\mathbb{Y}\rho\mathbb{Y} + \frac{\text{x}}{3}\mathbb{Z}\rho\mathbb{Z} \tag{3.27}$$

Assuming $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, the state-dependent noisy observable Z is:

$$\langle Z_{\text{x}} \rangle = (2|\beta|^2 - 1)\left(1 - \frac{4}{3}\text{x}\right). \tag{3.28}$$

With the notations out of the way, now we can focus on assessing the quality of the digital histograms in the presence of time-varying quantum noise[46]. We reduce the complexity in assessment by developing an intuitive performance evaluation framework. Specifically, we differentiate between computational accuracy, result reproducibility, program stability, and device reliability. These notions are related yet still distinct.

3.2.1 Computational accuracy

We begin by defining computational accuracy. We say that a quantum computation is $\epsilon-$accurate if the Hellinger distance between $\mathbb{P}^{\text{noisy}}$ and $\mathbb{P}^{\text{ideal}}$ is upper bounded by ϵ:

$$H(\mathbb{P}^{\text{ideal}}, \mathbb{P}^{\text{noisy}}) \leq \epsilon \tag{3.29}$$

The above definition requires a-priori knowledge of the noiseless reference distribution $\mathbb{P}^{\text{ideal}}$. This may be an impractical requirement when testing the accuracy for large problem sizes. In such cases, instead of looking at histogram accuracy, we might choose to look at the accuracy of the mean of an

observable. The accuracy condition is then described as:

$$|\langle O_{noisy}\rangle - \langle O_{noiseless}\rangle| \leq \epsilon \tag{3.30}$$

$$\Rightarrow |\sum_m \lambda_m \left(p^{\text{noisy}}(m) - p^{\text{noiseless}}(m)\right)| \leq \epsilon \tag{3.31}$$

$$\Rightarrow |\sum_m \lambda_m \text{Tr}\left[\Pi_m \left(\mathcal{E}_{\text{x}}(\rho_{\text{out}}) - \rho_{\text{out}}\right)\right]| \leq \epsilon \tag{3.32}$$

where $\mathcal{E}_{\text{x}}(\cdot)$ denotes the effective noise channel, λ_m is an eigenvalue of the observable $\hat{O}$, and Π_m is the projection operator corresponding to the m-th eigenstate.

Consider the single-qubit example in the presence of depolarizing noise. The accuracy metric (in terms of the Hellinger distance) for this case is state-dependent and is given as:

$$H = \left(1 - |\alpha|^2\sqrt{1 - \frac{2\text{x}}{3}\left(1 - \left|\frac{\beta}{\alpha}\right|^2\right)} - |\beta|^2\sqrt{1 - \frac{2\text{x}}{3}\left(1 - \left|\frac{\alpha}{\beta}\right|^2\right)}\right)^{1/2} \tag{3.33}$$

The accuracy metric in terms of the Z observable is:

$$|\langle Z\rangle_{\text{noisy}} - \langle Z\rangle_{\text{noiseless}}| = \left|\frac{4\text{x}(1-2|\beta|^2)}{3}\right| \tag{3.34}$$

Requiring ϵ-accuracy places an upper bound on the depolarizing channel parameter as:

$$\text{x} \leq \frac{3\epsilon}{4}. \tag{3.35}$$

3.2.2 Distribution reproducibility

Next, consider the problem of reproducibility in quantum computing. We will call our empirical histogram δ-reproducible if:

$$\Pr(H \leq \epsilon) \geq 1 - \delta, \tag{3.36}$$

where $1 - \delta$ is the statistical confidence level. This analysis requires an ensemble of histograms to be created through multiple executions on a noisy quantum computer.

With respect to the mean of a quantum observable, we may similarly pose the reproducibility condition as:

$$\Pr(\left|\langle O^{ideal}\rangle - \langle O^{noisy}\rangle\right| \leq \epsilon) \geq 1 - \delta \tag{3.37}$$

where $\langle O_{\mathrm{x}}^{noisy} \rangle$ is a random variable due to the presence of both shot noise as well as the non-stationarity of x. This reproducibility condition may be used to derive a stronger bound on the device noise. For example, consider again the single-qubit example in the presence of depolarizing noise. Suppose the depolarizing parameter x follows an exponential distribution:

$$f_X(\mathrm{x}) = \nu \exp^{-\nu \mathrm{x}}. \tag{3.38}$$

Then the δ-reproducibility condition requires that the mean of the depolarizing parameter x should be bounded as:

$$\mathbb{E}(\mathrm{x}) = \frac{1}{\nu} \leq \frac{3\epsilon}{4|\log \delta|}. \tag{3.39}$$

3.2.3 Hardware reliability

While device characterization metrics can be technology specific, there is a subset of five abstractions that represent the fundamenatal criteria for achieving a functional quantum computer [69]. These are: (1) Register size, n, a measure of the information capacity, (2) SPAM fidelity, F_{SPAM}, a measure of the noise in preparing a fiducial state and subsequently measuring it, (3) gate fidelity, F_{G}, a measure of the noise in implementing a quantum operation, (4) duty cycle, τ_G, a measure of the number of operations feasible before a quantum state decoheres, and (5) addressability, F_{A}, a measure of unwanted inter-qubit cross-talk. We use this subset to characterize the reliability of a NISQ[34] computer.

SPAM fidelity is defined as:

$$F_{\mathrm{SPAM}} = 1 - \epsilon_{\mathrm{SPAM}}. \tag{3.40}$$

where F_{SPAM} stands for the probability of preparing and measuring the n-qubit register in a fiducial state while ϵ_{SPAM} is the probability of observing any other erroneous outcome.

Gate fidelity (F_G) is defined by the error per Clifford gate ϵ_G, often measured using randomized benchmarking:

$$F_G = 1 - \epsilon_G. \tag{3.41}$$

We focus our attention specifically on characterizing the two-qubit CNOT gate, which often plays a decisive role in the performance limits of NISQ computing (second only to SPAM noise).

We define the duty cycle τ_G as the ratio of the duration T_G of a given gate to the register de-coherence time:

$$\tau_G = T_G / T_D. \tag{3.42}$$

The composite metric τ_G measures the number of quantum operations that can be executed before the register de-coheres, providing a quantification of the circuit depth achievable.

Addressability characterizes how well each register element can be measured individually. We quantify this in terms of the intra-register correlations [106, 107] that arise during quantum operations and measurement, due to either unwanted entanglement or classical cross-talk. Specifically, we define addressability F_A as:

$$F_A = 1 - \frac{2\mathcal{I}(X,Y)}{\mathcal{H}(X) + \mathcal{H}(Y)}. \tag{3.43}$$

where $\mathcal{I}(X,Y)$ is the mutual information between X and Y and $\mathcal{H}(\cdot)$ denotes the entropy.

We use a numerical simulation to illustrate addressability F_A. Let us characterize a two-qubit device where q_0 is the first qubit and q_1 is the second qubit. Assume that the state prior to measurement ($|\alpha\rangle$) is in one of four computational basis states: $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$. Suppose this prior state is impacted by an uncorrelated binary noise process to become $|\beta\rangle$. The transition probability from $|\alpha\rangle$ to $|\beta\rangle$ is given by Table 3.2. The input to this model is the intermediate state $|\beta\rangle$, which is subjected to a correlated noise process as shown in Fig. 3.1. Let $P(X)$ denote the probability of state X. Using the noise models represented by Table 3.2:

$$\begin{aligned} P(|s\rangle = |00\rangle) =& \frac{1+2u}{4} \\ P(|s\rangle = |01\rangle) =& \frac{1-2u}{4} \\ P(|s\rangle = |10\rangle) =& \frac{1-2u}{4} \\ P(|s\rangle = |11\rangle) =& \frac{1+2u}{4} \end{aligned} \tag{3.44}$$

The probability of observing individual measurement outcomes Q_0 and Q_1 are:

$$\begin{aligned} \Pr(Q_0 = 0) =& 1 - \Pr(Q_0 = 1) = \Pr(|s\rangle = |00\rangle) + \Pr(|s\rangle = |01\rangle) \\ \Pr(Q_1 = 0) =& 1 - \Pr(Q_1 = 1) = \Pr(|s\rangle = |00\rangle) + \Pr(|s\rangle = |10\rangle) \end{aligned} \tag{3.45}$$

which yields

$$\begin{aligned} \Pr(Q_0 = 0) = \Pr(Q_0 = 1) = \frac{1}{2} \\ \Pr(Q_1 = 0) = \Pr(Q_1 = 1) = \frac{1}{2} \end{aligned} \tag{3.46}$$

The binary entropy is therefore maximal, i.e., $H(Q_0) = H(Q_1) = 1$. This leads to a final expression for the addressability as

$$F_A = 1 - \frac{1+2u}{2}\log(1+2u) - \frac{1-2u}{2}\log(1-2u) \tag{3.47}$$

This model analyzed addressability for a simple Markov model as shown in Fig. 3.1. We show in Fig. 3.2 the variability of addressability using a numerical simulation of a correlated noise model. A similar spatial characterization (numerical simulation) is shown in Fig. 3.3. We quantify reliability using the distance between quantum noise densities at different times and register locations. For example, $H_{\mathrm{SPAM}}(t_1, t_2)$ measures the similarity in distributions of SPAM fidelity at different times. A reliable, but not necessarily ideal, device maintains the characteristic density for the fidelity at both times. By similar considerations, spatially-varying noise processes can be subjected to reliability analysis using this definition. We will call a device ε-reliable if:

$$H_X < \varepsilon \tag{3.48}$$

The normalized version in Eqn. 3.10 is more sensitive and useful for reliability testing.

3.2.4 Observable stability

Suppose that the distribution of quantum noise exhibits a time-dependence:

$$f_X(\mathrm{x}) = f(\mathrm{x}; t) \tag{3.49}$$

Observable stability studies bounds on the output of a noisy quantum circuit in presence of time-varying quantum noise[108].

Specifically, the instability between the results obtained at two different times t_1 and t_2 can be quantified by:

$$s(t_1, t_2) = |\, \langle O \rangle_{t_1} - \langle O \rangle_{t_2} \,| \tag{3.50}$$

where

$$\langle O \rangle_t = \int \langle O_{\mathrm{x}} \rangle \, f(\mathrm{x}; t) d\mathrm{x}. \tag{3.51}$$

Here, x denotes a specific realization of the random quantum noise parameter X, which has a density characterized by $f(\mathrm{x}; t)$. $\langle O_{\mathrm{x}} \rangle$ is the mean of the quantum observable $\hat{O}$, computed from results obtained from a noisy quantum circuit subject to noise x.

The presence of time-varying noise renders the observable (a random number) non-stationary stochastic process when observed as a time-series. Hence, the mean of the observable may exhibit symptoms characteristic of a non-stationary stochastic process such as drifting means or time-varying error bars.

As side note, stationary noise may produce irreproducible results because of large static variance. Also, stationary (stable) noise with small variance may produce reproducible results and yet be inaccurate because of bias. The point is that accuracy, reproducibility and stability of outcomes are all conceptually distinct (albeit related) concepts.

The $\epsilon-$stability condition can be stated as:

$$s(t_1, t_2) \leq \epsilon \tag{3.52}$$

This framework enables the question of time-scale for re-calibration. For example, consider a depolarizing channel characterized by the depolarizing parameter p. Suppose, p exhibits non-stationarity and its stochastic behavior can be modeled by an exponential distribution with a time-varying parameter $\lambda(t)$:

$$\lambda(t) = \lambda_0 - \chi\delta t. \tag{3.53}$$

This models the situation when error bars on p increases with time (as often happens in between calibrations) since the spread (or variance) of an exponential distribution $\propto \lambda^{-2}$. The time-scale at which the device transitions from stable to unstable (in the absence of re-calibration) can then be estimated from Eqn. 3.52 as:

$$\delta t_{\text{stable}} \leq 2\sqrt{2}\frac{\lambda_0}{\chi}\sqrt{1 - \sqrt{2\phi^2 - 1}} \tag{3.54}$$

where $\phi = \frac{3\sqrt{3}\epsilon}{14\sqrt{2}}$. Thus, in presence of this variance-drift model, we expect a stable device to remain stable until $t = \delta t_{stable}$. This gives an estimate of the frequency of device calibration required to meet the stability condition.

We end this section with a result for a special case. Assume that the quantum noise channel can be assumed to separable. Thus, the effect of noise on the quantum state can be understood by examining the impact of noise on individual qubits. Additionally, suppose that the channel is a first-order polynomial in x (for instance, a depolarizing channel). Also, assume that x is wide-sense stationary (WSS). Thus, its mean remains constant over time, and its standard deviation depends on

the time interval between observations. These are all reasonable assumptions that correspond to noise data collected from NISQ computers.

Under these assumptions, it can be shown that the mean of the observable $\langle O \rangle$ is invariant with respect to time:

$$\begin{aligned}
\langle O \rangle_t &= \int \langle O_X \rangle \, f_X(\mathrm{x}; t) d\mathrm{x} \\
&= \prod_{i=1}^{n} \int \langle r_i | \, \mathcal{E}_{\mathrm{x}_i}(|r_i\rangle \langle r_i|) \, |r_i\rangle \, f_{X_i}(\mathrm{x}_i; t) d\mathrm{x}_i \\
&= \prod_{i=1}^{n} \int \langle r_i | \, [|r_i\rangle \langle r_i| + \mathrm{x}_i g(|r_i\rangle \langle r_i|)] \, |r_i\rangle \, f_{X_i}(\mathrm{x}_i; t) d\mathrm{x}_i \qquad (3.55) \\
&\quad \text{(where } g(\cdot) \text{ is an arbitrary function that outputs a valid density matrix)} \\
&= \prod_{i=1}^{n} [1 + \mu_{\mathrm{x}_i} \langle r_i | \, g(|r_i\rangle \langle r_i|) \, |r_i\rangle]
\end{aligned}$$

which is independent of time because of the wide-sense stationarity assumption. Thus, $s(t_1, t_2) = 0$ when the quantum noise channel is separable, WSS and has noise terms till first-order only.

3.3 Test circuit used

We use the Bernstein-Vazirani [109] circuit as a test circuit for performance evaluation because it is a well-known example of quantum advantage, requiring only a modest number of gates, and is commonly used as a benchmarking tool for quantum computers. It was conceived by Bernstein and Vazirani in 1992 as an extension of Simon's algorithm. Using a quantum algorithm, it transforms a problem of $O(n)$ complexity to one of $O(1)$.

The algorithm is tasked with deciphering a n-bit secret string r, embedded in a black-box oracle function. The oracle responds with a yes/ no answer to the question: Is the secret string w? The algorithm locates the secret via just one query, irrespective of the value of n. The classical algorithm requires n queries (worst case).

Mathematically, the oracle function takes an n-bit string as input (w) and produces the following output:

$$f(w) = w \cdot r \bmod 2, \qquad (3.56)$$

where $(\cdot)$ represents bitwise multiplication and mod 2 ensures the output is either 0 or 1.

The quantum circuit for the secret string $r = 1000$ is shown in Fig. 3.4. In the first step, an equal superposition across all 2^n possible input bit strings is generated using a layer of Hadamard gates

acting on $|0\rangle^{\otimes n}$. The second layer has an implementation of the oracle function using a layer of CNOT gates. This is followed by another layer of Hadamard gates which yields the binary representation of the secret string r at the output.

Let us now understand precisely how this algorithm works. We will sometimes omit the normalization factors for clarity here. To begin, for 1-qubit, it is easily verified that:

$$H|\mathrm{x}\rangle = \sum_{w\in\{0,1\}} (-1)^{\mathrm{x}w} \frac{|w\rangle}{\sqrt{2}} \tag{3.57}$$

Thus,

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \tag{3.58}$$

$$H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \tag{3.59}$$

When the input register is a n-bit state $|w_0 \dots w_{n-1}\rangle$ and it is acted upon by a layer of Hadamard gates as shown in Fig. 3.4, then the output is:

$$\begin{aligned} H^{\otimes n}|w_0 \dots w_{n-1}\rangle &= H|w_0\rangle \otimes H|w_2\rangle \cdots \otimes H|w_{n-1}\rangle \\ &= \frac{1}{\sqrt{2^n}} \sum_{\mathrm{x}} (-1)^{w\cdot x} |\mathrm{x}\rangle \end{aligned} \tag{3.60}$$

Thus,

$$H^{\otimes n}|w\rangle = \frac{1}{\sqrt{2^n}} \sum_{\mathrm{x}} (-1)^{w\cdot \mathrm{x}} |\mathrm{x}\rangle \tag{3.61}$$

For our 5-qubit circuit, the initial input is $|0\rangle^{\otimes 5}$. After the first unitary layer, the output becomes:

$$|\psi_1\rangle = |++++\rangle\,|-\rangle\,. \tag{3.62}$$

Next comes the unitary layer U_f that encodes the oracle function $f(w)$. Specifically,

$$U_f(|xy\rangle) = |x\rangle\,|y + f(x)\rangle \tag{3.63}$$

When $f(x) = 0$:

$$\begin{aligned} U_f\left(|x\rangle|-\rangle\right) &= U_f\left(|x\rangle\left(|0\rangle - |1\rangle\right)\right) \\ &= U_f\left(|x\rangle|0\rangle\right) - U_f\left(|x\rangle|1\rangle\right) \\ &= |x\rangle|0+f(x)\rangle - |x\rangle|1+f(x)\rangle \\ &= |x\rangle|0+0\rangle - |x\rangle|1+0\rangle \\ &= |x\rangle|-\rangle \\ &= (-1)^{f(x)}|x\rangle|-\rangle \end{aligned} \tag{3.64}$$

When $f(x) = 1$:

$$\begin{aligned} U_f\left(|x\rangle|-\rangle\right) &= U_f\left(|x\rangle\left(|0\rangle - |1\rangle\right)\right) \\ &= U_f\left(|x\rangle|0\rangle\right) - U_f\left(|x\rangle|1\rangle\right) \\ &= |x\rangle|0+f(x)\rangle - |x\rangle|1+f(x)\rangle \\ &= |x\rangle|0+1\rangle - |x\rangle|1+1\rangle \\ &= -|x\rangle|-\rangle \\ &= (-1)^{f(x)}|x\rangle|-\rangle \end{aligned} \tag{3.65}$$

We shorten this to write:

$$\begin{aligned} U_f|x\rangle &= (-1)^{f(x)}|x\rangle \\ &= (-1)^{r'\cdot x}|x\rangle \end{aligned} \tag{3.66}$$

with the implicit assumption that the ancilla was set to $|-\rangle$. Here, r' is the 1-bit secret string. It follows then that,

$$\begin{aligned} \sum_{\mathrm{x}} U_f|x\rangle &= \sum_{\mathrm{x}}(-1)^{r\cdot x}|x\rangle \\ &= H^{\otimes n}|r\rangle \ \text{(from Eqn. 3.61)} \end{aligned} \tag{3.67}$$

where r is the n-bit secret string. Thus, at the end of the second layer of the circuit, the output is of the form:

$$|\psi_2\rangle = H^{\otimes n}|r\rangle|-\rangle\,. \tag{3.68}$$

So, to retrieve the secret string r, the third layer simply needs another layer of Hadamard gates. This yields the final output as:

$$|\psi_3\rangle = |r\rangle|1\rangle\,. \tag{3.69}$$

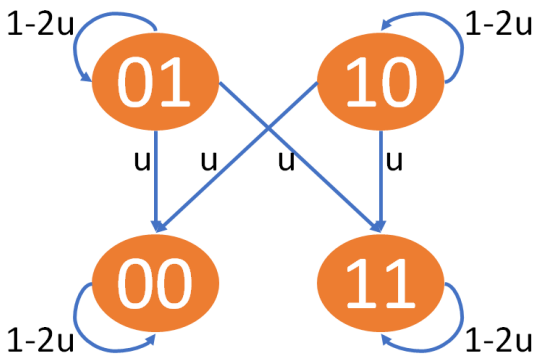


Figure 3.1: A classical Markov model for the correlated error process.

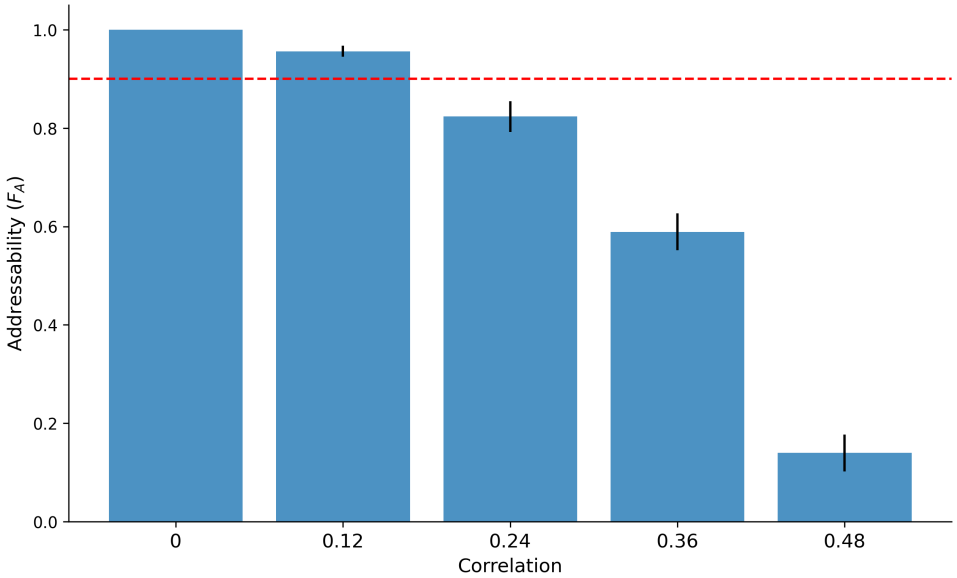


Figure 3.2: Simulation result for addressability F_A showing sensitivity to intra-register correlation.

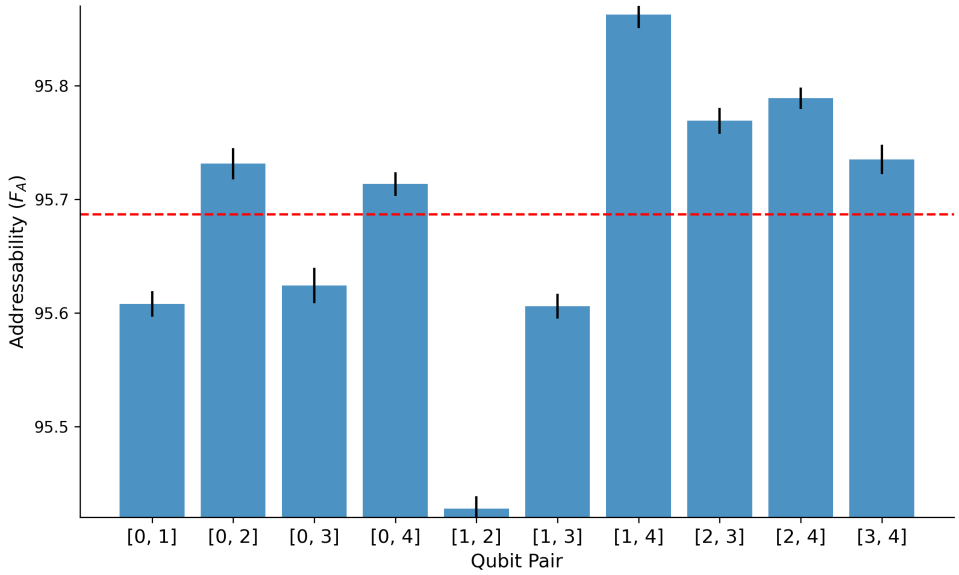


Figure 3.3: Addressability: pairwise comparison when correlation parameter $u = 0.12$. Ideally, there should have been zero spatial variation (as idealized simulations do not differentiate between qubits at different physical locations). However, small fluctuations are seen about a mean value of 95.80 with a standard deviation of 0.02. This arises due to the readout error fluctuations in the ibm_yorktown noise model.

Table 3.2: Transition Probabilities for Uncorrelated Noise

	$\|00\rangle$	$\|01\rangle$	$\|10\rangle$	$\|11\rangle$
$\|00\rangle$	$1-p$	$p/3$	$p/3$	$p/3$
$\|01\rangle$	$p/3$	$1-p$	$p/3$	$p/3$
$\|10\rangle$	$p/3$	$p/3$	$1-p$	$p/3$
$\|11\rangle$	$p/3$	$p/3$	$p/3$	$1-p$

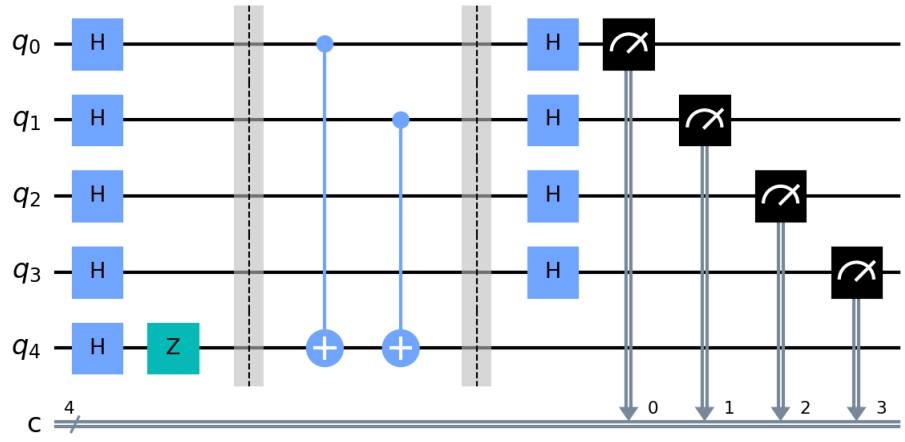


Figure 3.4: The Bernstein-Vazirani circuit for a 4-bit secret string $r = 1000$. H represents the Hadamard gate and Z represents the Z-gate. The meter symbols are measurement operations that project to the computational basis states. The measurement results are recorded in a classical register c.

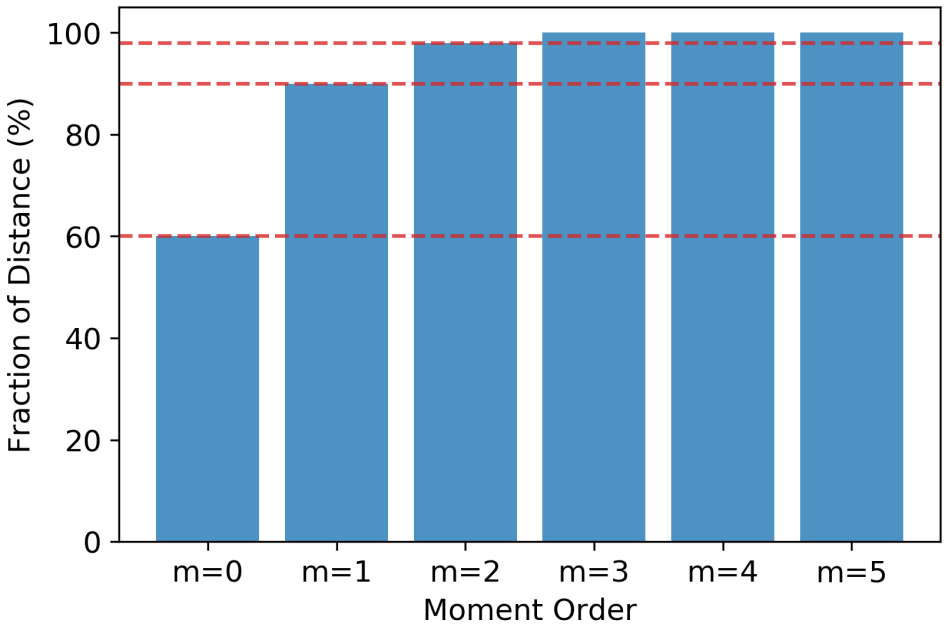


Figure 3.5: Contribution to moment based distance (d) from increasing moment orders. The graph shows the results of comparing two normal distributions: $\mathcal{N}_1(\mu = \mu_0, \sigma = \sigma_0)$ and $\mathcal{N}_2(\mu = 2\mu_0, \sigma = 2\sigma_0)$ where $\mu_0 = 40$ and $\sigma_0 = 4$.

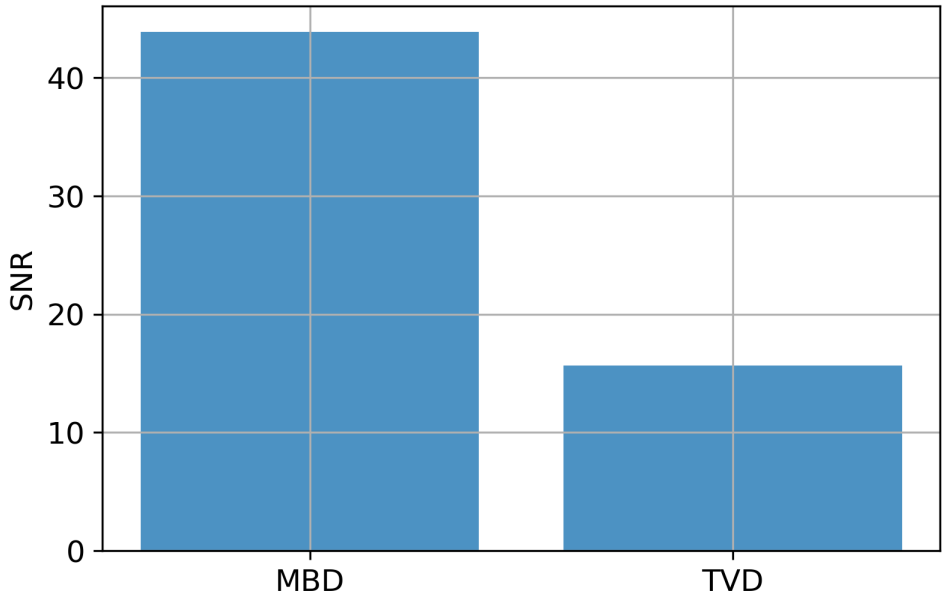


Figure 3.6: Simulated signal-to-noise ratios of the moment-based distance and total variation distance for two normal distributions of varying width.

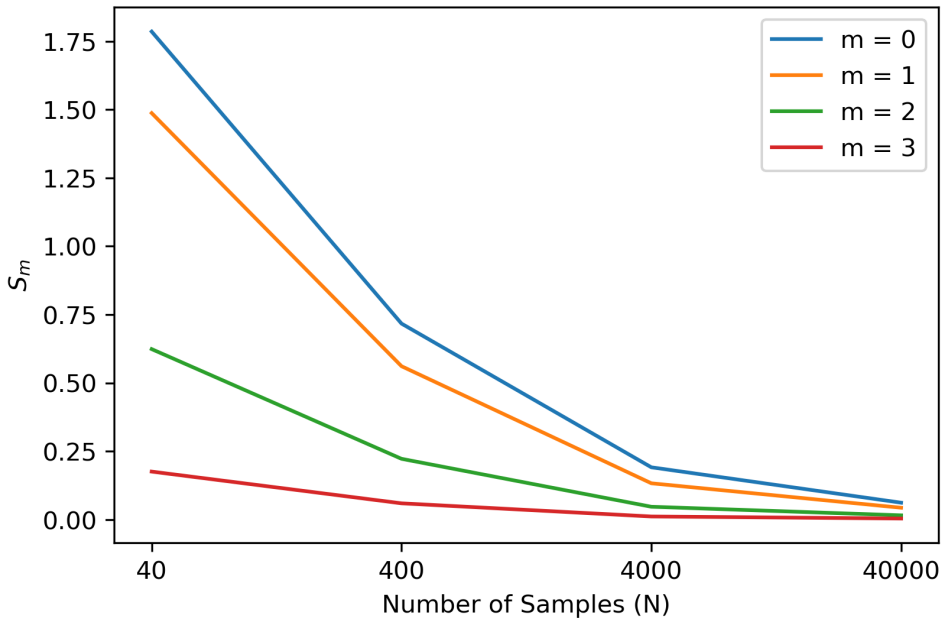


Figure 3.7: When the two distributions are similar, then we expect each S_m to be zero. Empirically, that happens as we increase the sample size. The lower order moments take longer to go to zero.

Chapter 4

Reliability of device characterization

In this chapter, we quantify and assess the reliability [110] of noisy quantum computers [34] using the performance evaluation framework introduced in the previous chapter. Our reliability testing will use experimental data from IBM [111] at various time-scales (monthly, daily, hourly and seconds level). Spatial reliability testing [112] will explore variations across different parts of the system while the temporal reliability testing will study changes over time. Both component level metrics (such as individual gates and qubits) as well as composite level metrics (such as circuits) will be discussed. The chapter will highlight the non-stationary nature of noise in contemporary quantum computers. It will discuss the implications of such unreliable devices on program outcomes [110]. This will in turn motivate the need for rigorous stability analyses (in subsequent chapters) to ensure confidence in results from noisy quantum computers. It will provide the setting for dynamic mitigation [74] to enhance reliability which will be discussed in later chapters.

4.1 Experimental data

4.1.1 Device

We used quantum computers provided by IBM, which are based on the transmon qubit architecture [30]. IBM has unveiled a series of processors over the past few years, with steadily increasing register size [55]. Some of the earliest processors were categorized as Canary, featuring 2-16 qubits. After that came the Falcon processors with 27 qubits, Egret processors with 33 qubits, Hummingbird processors with 65 qubits, Eagle processors with 127 qubits and Osprey processors with 433 qubits. As an aside, the classification of processors goes beyond mere qubit count and can encompass details like the connectivity graph [113]. Our research used data from ibm_yorktown (5 qubits), ibm_toronto (27

qubits), and ibm_washington (127 qubits), which belong to the Canary, Falcon and Eagle families respectively.

4.1.2 Data

We employed the Qiskit software library [114] to generate our data sets. The access to IBM quantum computers was provided by the Oak Ridge Leadership Computing Facility (OLCF) located at Oak Ridge National Laboratory, Tennessee [115].

The device characterization data (at daily time-scale) for the ibm_washington device spanned 16 months (starting from January 1, 2022, and ending on April 30, 2023). Examples of elements recorded include the date and time of the device calibration, the SPAM (State Preparation and Measurement) error rate for each individual qubit, the CNOT gate error rates (calculated via randomized benchmarking [116] employing varying lengths of two-qubit Clifford gates [117]), the duration of the CNOT gates [118], and the de-coherence times (T_1 and T_2) for each qubit [18]. Similar daily data was gathered for the quantum computer called ibm_yorktown (from March 1, 2019, to December 30, 2020).

A separate set of intra-day data was gathered from the 27-qubit ibm_toronto device on December 11, 2020, at the following time intervals: 8:00-8:30 am, 11:00-11:30 am, 2:00-2:30 pm, 5:00-5:30 pm, 8:00-8:30 pm, and 11:00-11:30 pm (UTC-5). In this dataset, each individual qubit was sequentially sampled a total of $212,992$ times. It had a total of $5,750,784$ recorded outcomes.

4.2 Component reliability

In this section, we evaluate the temporal and spatial reliability of the DiVincenzo metrics [69] that we discussed in Sec. 3.2.3. We use data from ibm_yorktown and ibm_toronto whose layouts are shown in Fig. 4.1 and 4.2.

4.2.1 Reliability of SPAM noise characterization

We find that the SPAM fidelity for all three IBM computers (ibm_yorktown, ibm_toronto, and ibm_washington) fluctuates significantly over time, even when the average fidelity is tightly controlled. This is indicative of poor device reliability. Fig. 4.3 shows an example of the SPAM noise density at a particular time. Fig. 4.4(a)-(e) show the temporal fidelity of Yorktown. The top panel has the time series for the mean and variance, while the bottom panel plots the distance between the time-varying densities. The red line is the median. Fig. 4.16 shows the results for washington (qubit 37). The

series starts in Dec-21 (when the device was commissioned) and ends in Oct-22. The red line is the median.

All these plots show that there can be long periods when the fidelity is tightly controlled, but there are also times when it fluctuates significantly. These fluctuations are not reflected in the variance and could be due to changes in the underlying physics [119].

On the other hand, spatial reliability refers to the similarity of the densities between different locations. We find that the SPAM fidelity can vary significantly depending on location, even on the same device! This could be due to variations in the physical properties of the individual qubits. Fig. 4.4(f) shows the distance between the spatial densities of ibm_toronto, Fig. 4.5 for ibm_yorktown and Fig. 4.20 for ibm_washington. The large distance measures show that spatial reliability is poor within the same device.

4.2.2 Reliability of CNOT noise characterization

We next analyze the reliability of CNOT gates by studying density similarity. The underlying random variable is the CNOT gate fidelity F_G [120].

Fig. 4.6(a)-(e) show the results for ibm_yorktown between Mar-19 and Dec-20. The reference density (for distance computation) is Mar-19. We see that the metric diverged sharply between Jun-19 and Dec-19, but fluctuated much less in the next 12 months. It stands to reason that the CNOT operations performed in Mar-19 are quite different from those performed in Dec-19! Fig. 4.17 shows similar results for ibm_washington for a CNOT between qubits 0 and 14. The distance reaches as high as 0.7 (max allowed is 1.0) in the second half of 2022.

The previous discussion was around temporal reliability of CNOT. The spatial reliability of ibm_yorktown is shown in Fig. 4.9. (Note: our spatial calculation used the entire temporal dataset.) The worst gates were found to be between qubits (1,2) and (3,2) which yielded a distance of 0.467. The worst densities are shown in the inset. Fig. 4.21 is the same plot but for ibm_washington for Sep-22. The most dissimilar pairs in this case were (11, 12) and (19, 20) for which the distance exceeded 0.99. The inset shows the worst case densities. The lookup table for the 144 connections (which specifies which qubit-pairs are being referred to in the CNOT gate) is provided in Table 4.3. One of the takeaways from these analyses is that CNOT fidelities may show misleadingly similar means despite having starkly dissimilar densities. Another takeaway is that there exist certain qubits (such as qubit 3 for yorktown) which can be associated with generally worse reliability outcomes.

4.2.3 Reliability of duty cycle characterization

The CNOT duty cycle [110] was defined in Sec. 3.2.3. It is a random variable with an observable density. An example density, shown in Fig. 4.14, uses data from ibm_washington between Dec-21 and May-22.

The temporal reliability analysis for ibm_yorktown is shown in Fig. 4.7. It has the time-series for the mean decoherence time T_2 (to be precise, this is the harmonic mean across the two qubits of the CNOT gate). It also has the time-series for the tunable gate duration and distance metric (that quantifies reliability). The latter uses densities based on running 3-month data. We can see, for example, in Fig. 4.7 (a) on July 24, 2020, the T_2 time decreased abruptly from 77 μs to 31 μs for register 0 and from 82 μs to 24 μs for register 1. A corresponding random sharp increase is seen in the gate duration from 370 ns to 441 ns. These random changes led to a sudden sharp decrease in the duty cycle from 107.2 to 30.9.

The spatial reliability of ibm_yorktown is depicted in Fig. 4.8 and for ibm_washington in Fig. 4.19. For ibm_washington, the most dissimilar duty cycles were CNOT$(46, 47)$ and CNOT$(96, 109)$. For ibm_yorktown, the most dissimilar duty cycles were CNOT$(0, 1)$ and CNOT$(3, 2)$.

In general, the temporal reliability of the duty cycle improved after July 2020, but the spatial reliability remained poor.

4.2.4 Reliability of addressability characterization

Addressability [121] (defined in Sec. 3.2.3) is a measure of how well register qubits can be addressed individually. The ideal addressability is 1, but in practice it can be lower due to hardware noise.

Fig. 4.22 plots the addressability of the ibm_toronto device when tested by encoding a fiducial separable state $|00\rangle$ in each register pair. The heatmap highlights how the addressability varies across the register. The inset compares the limits of this behavior by showing the worst case 0.89 for qubits $(23, 21)$ and the best ~ 1.00 for qubits $(11, 13)$.

An interesting extension can be performed by encoding a Bell-state within the register pair. As a maximally entangled state, the addressability should be 0. Fig. 4.13 shows the results for the 28 register pairs that support direct preparation of a Bell state based on the nearest-neighbor connections shown in Fig. 4.2. We found that the worst registers for Bell state information preservation were 12 and 15 with:

$$\eta = 1 - F_A = 0.14 \pm 0.014, \tag{4.1}$$

while the best registers were 25 and 26 with:

$$\eta = 1 - F_A = 0.84 \pm 0.023. \tag{4.2}$$

This unreliable device has now been retired.

4.3 Circuit reliability

So far, we focused on analyzing the device components at individual qubit and gate level. But how do we measure reliability more holistically at the circuit level? Examining thousands (if not millions) of qubits and gates may not help to drawn any conclusions at the application level.

To analyze circuit reliability, we study the stationarity of the multi-variate noise vector X associated with a quantum circuit. Conducting a holistic analysis requires studying the time-variation [46] of the joint distribution of the quantum noise. However, significant correlations exist amongst the noise parameters characterizing a quantum circuit, and these correlations can have a substantial impact on the performance of quantum error correction and validation methods. The correlation structure can also change over time (see Fig. 4.12 for example). For example, in [122], the authors examine the decoherence of a quantum computer in a temporally and spatially correlated environment, finding that minor adjustments to error correction codes can systematically reduce the impact of long-range correlations on the quantum system. In [123], researchers discuss the limitations of single-metric approaches for quantum characterization and study the influence of noise correlations on randomized benchmarking (RB). They demonstrate that temporal noise correlations affect the probability density function of RB outcomes, described by a gamma distribution with parameters dependent on the correlation structure, while also noting potential finite-sampling issues and deviations in mean RB outcomes from worst-case errors when noise correlations are present.

The selection of the noise metrics is based on the DiVincenzo criterion [69], like we discussed in Ch. 3 and it spans qubit-specific SPAM fidelity, gate fidelity, duty cycle, and addressability. A test circuit is used for reporting, specifically the Bernstein-Vazirani circuit [109], which is a canonical algorithm with proven quantum advantage in noiseless limit. These metrics are shown in Table 5.2.

Specifically, we compute the distance between the time-varying densities for the random variable:

$$X = (X_0, \cdots, X_{15}),$$

where $X_0, \cdots, X_{15}$ are described in Table 5.2. The distance is denoted by $H_X(t, t_0)$ where t denotes the months ranging from Jan-2022 to Apr-2023. The reference time t_0 is set to Jan-2022. Table 4.2 shows the distance from the Jan-2022 density. The first 16 columns present the distance data for univariate (marginal) distributions, while the last three columns contain the distances for the composite densities. Specifically, the H_r column shows the distance between the joint densities modeled using copulas [124], H_a represents the average over the marginals, and H_n represents the normalized distance (as discussed in Chapter 3).

Fig. 4.10 plots the the last three columns containing the distances for the composite densities. The unmodified distance per Eqn. 3.1 is less sensitive due to the curse of dimensionality [125]. The orange line represents the normalized distance per Eqn. 3.10, while the green line represents the distance averaged over all the marginal distributions per Eqn. 3.9. The contributions to the average distance for Apr-2023 distribution compared to that of Jan-2022 from various noise sources are compared and contrasted in Fig. 4.11. It is apparent that no single parameter dominates the average though SPAM noise accounts for the largest contribution to circuit non-stationarity.

It should be emphasized that the average measure does not consider correlations between parameters, and a specific correlation structure can cause the joint distance to increase by reducing overlap in a subset of dimensions. The normalized and average distances, represented by the orange and green lines respectively, demonstrate greater discriminatory power with the observed data ranges of 0.51 and 0.20, respectively. These ranges are considerably higher compared to the unmodified distance with an observed data range of 0.029.

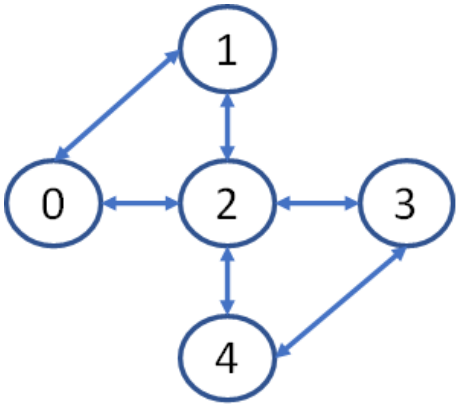


Figure 4.1: Schematic layout of the ibm_yorktown device produced by IBM. Circles denote register elements and edges denote connectivity of 2-qubit operations.

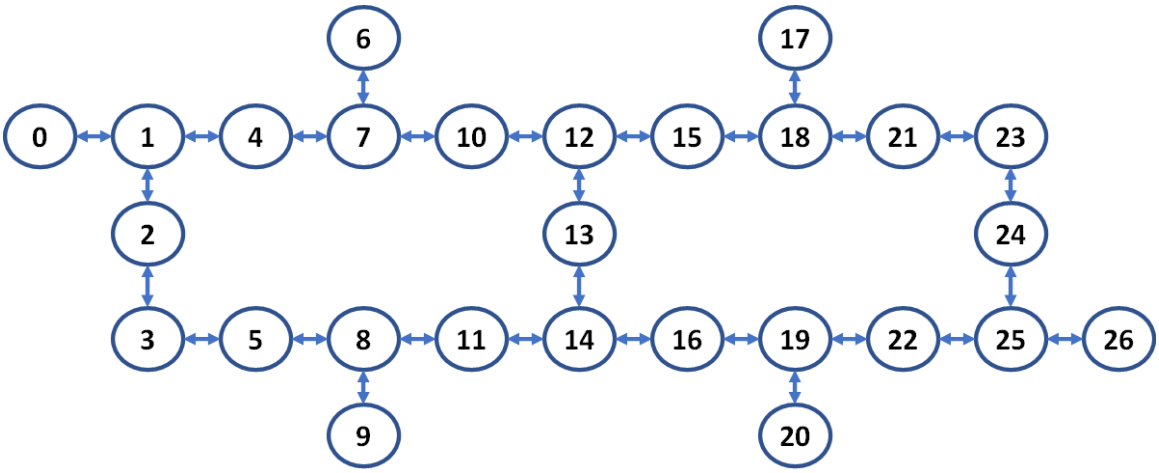


Figure 4.2: Schematic of the ibm_toronto device, produced by IBM. Circles represent register elements, while edges denote the connectivity for performing two qubit operations.

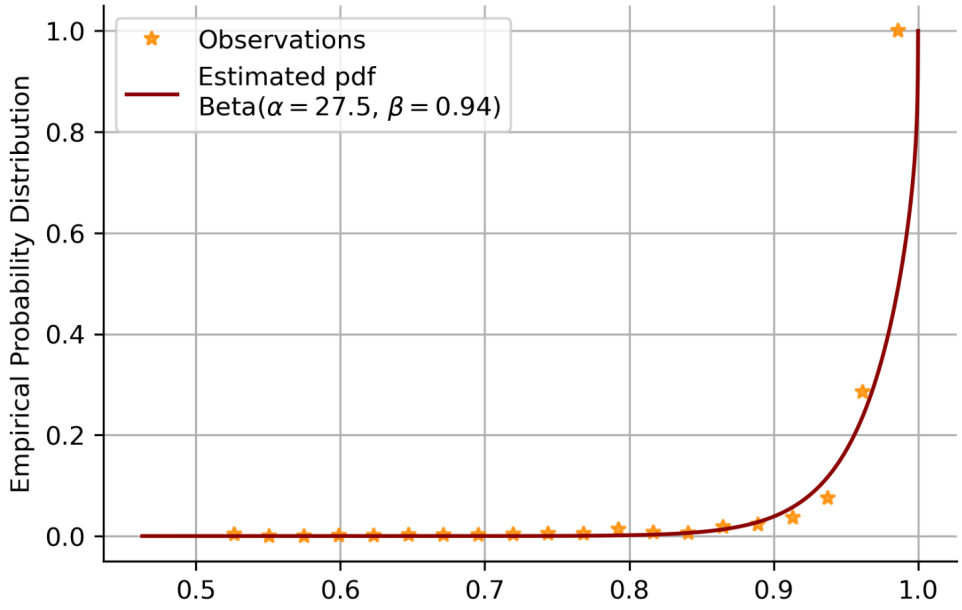
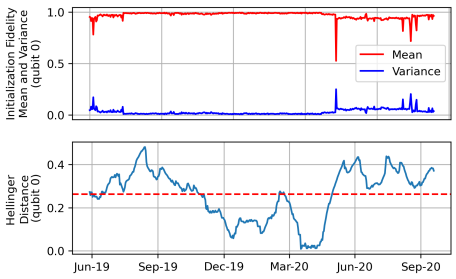


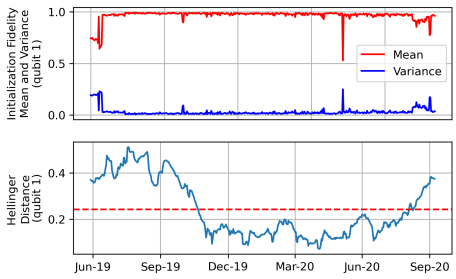
Figure 4.3: Experimentally observed probability density for SPAM (state preparation and measurement) fidelity of one of the register elements of the IBM transmon device named ibm_washington.

Parameter	*Description*	*Model*
x_0	SPAM fidelity, register 0	ABC
x_1	SPAM fidelity, register 1	ABC
x_2	SPAM fidelity, register 2	ABC
x_3	SPAM fidelity, register 3	ABC
x_4	CNOT fidelity, control 0, target 1	DP⊗DP
x_5	CNOT fidelity, control 2, target 1	DP⊗DP
x_6	T_2 time, register 0	TR
x_7	T_2 time, register 1	TR
x_8	T_2 time, register 2	TR
x_9	T_2 time, register 3	TR
x_{10}	T_2 time, register 4	TR
x_{11}	H fidelity, register 0	CP
x_{12}	H fidelity, register 1	CP
x_{13}	H fidelity, register 2	CP
x_{14}	H fidelity, register 3	CP
x_{15}	H fidelity, register 4	CP

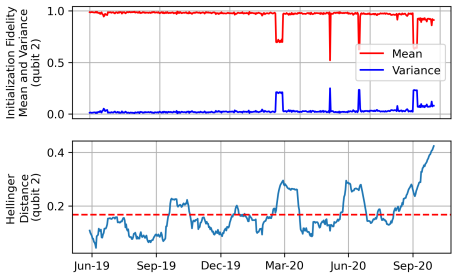
Table 4.1: The 16-parameter model derived from the ibm_washington data set has four types of quantum noise processes: (i) 'ABC': asymmetric binary channel model, (ii) 'CP': coherent phase error model, (iii) 'DP': depolarizing noise model and (iv) 'TR': thermal relaxation noise model. Note that the two-qubit model 'DP⊗DP' is a tensor product of depolarizing noise.



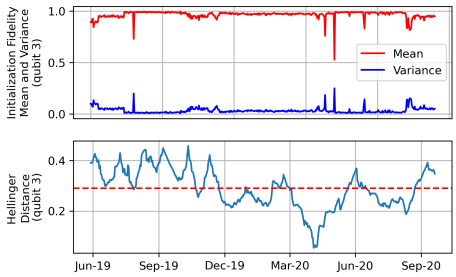
(a) ibm_yorktown qubit 0



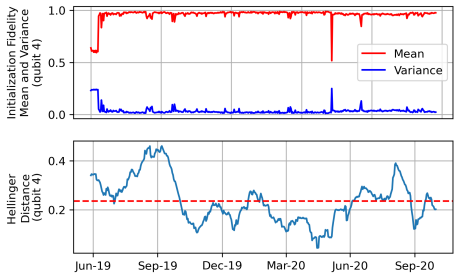
(b) ibm_yorktown qubit 1



(c) ibm_yorktown qubit 2



(d) ibm_yorktown qubit 3



(e) ibm_yorktown qubit 4

Figure 4.4: (a)-(e) Temporal stability of the SPAM fidelity F_{SPAM} of each register element in the ibm_yorktown device. The top panel shows the average F_{SPAM} of the register with associated variance, and the bottom panel shows a running calculation of the Hellinger distance using a one-month window. The dashed red line is the median value.

Table 4.2: Hellinger distance values for the device parameters.

Month	H_{X_0}	H_{X_1}	H_{X_2}	H_{X_3}	H_{X_4}	H_{X_5}	H_{X_6}	H_{X_7}	H_{X_8}	H_{X_9}	$H_{X_{10}}$	$H_{X_{11}}$	$H_{X_{12}}$	$H_{X_{13}}$	$H_{X_{14}}$	$H_{X_{15}}$	H_{n}	H_{a}	H_{r}
Jan-22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Feb-22	0.82	0.08	0.43	0.38	0.3	0.35	0.28	0.43	0.45	0.39	0.32	0.24	0.62	0.66	0.04	0.26	0.41	0.38	0.971439
Mar-22	0.97	0.22	0.31	0.3	0.07	0.05	0.31	0.17	0.6	0.22	0.32	0.31	0.11	0.36	0.48	0.17	0.47	0.31	0.99084
Apr-22	0.64	0.03	0.23	0.53	0.27	0.45	0.21	0.06	0.95	0.11	0.1	0.45	0.33	0.37	0.07	0.61	0.42	0.34	0.978632
May-22	0.8	0.27	0.61	0.77	0.11	0.4	0.64	0.21	0.38	0.36	0.26	0.65	0.21	0.34	0.34	0.49	0.57	0.43	0.99897
Jun-22	0.81	0.4	0.43	0.9	0.26	0.34	0.53	0.34	0.28	0.1	0.16	0.36	0.43	0.69	0.16	0.23	0.44	0.4	0.983197
Jul-22	0.74	0.42	0.96	1.0	0.3	0.44	0.15	0.25	0.48	0.37	0.2	0.16	0.38	0.17	0.14	0.32	0.79	0.41	1.0
Aug-22	0.89	0.5	0.9	1.0	0.26	0.41	0.53	0.35	0.97	0.31	0.21	0.55	0.22	0.21	0.27	0.3	0.85	0.49	1.0
Sep-22	0.82	0.48	0.93	1.0	0.45	0.22	0.44	0.34	0.91	0.1	0.08	0.46	0.27	0.31	0.31	0.13	0.79	0.45	1.0
Oct-22	0.72	0.55	0.9	1.0	0.05	0.42	0.32	0.18	0.95	0.07	0.27	0.43	0.36	0.66	0.74	0.25	0.77	0.49	1.0
Nov-22	0.36	0.63	0.65	1.0	0.4	0.13	0.55	0.53	0.98	0.22	0.14	0.4	0.29	0.39	0.29	0.3	0.61	0.45	0.999713
Dec-22	0.42	0.64	0.58	1.0	0.27	0.53	0.45	0.24	0.99	0.17	0.27	0.7	0.65	0.03	0.19	0.37	0.72	0.47	0.999995
Jan-23	0.46	0.59	0.46	1.0	0.06	0.5	0.31	0.3	0.91	0.53	0.26	0.69	0.55	0.34	0.53	0.46	0.68	0.5	0.999975
Feb-23	0.45	0.61	0.65	1.0	0.44	0.46	0.44	0.26	1.0	0.12	0.52	0.53	0.4	0.33	0.61	0.34	0.92	0.51	1.0
Mar-23	0.47	0.5	0.79	1.0	0.22	0.21	0.46	0.33	0.51	0.21	0.4	0.71	0.62	0.09	0.71	0.31	0.64	0.47	0.999876
Apr-23	0.43	0.55	0.65	1.0	0.16	0.61	0.37	0.25	0.12	0.14	0.15	0.55	0.26	0.42	0.36	0.17	0.61	0.39	0.999721

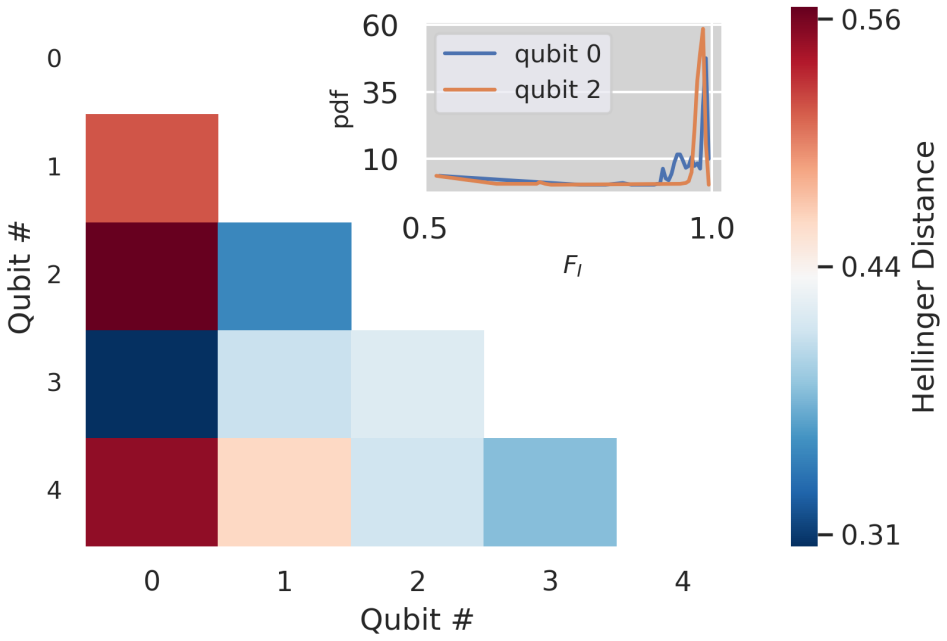
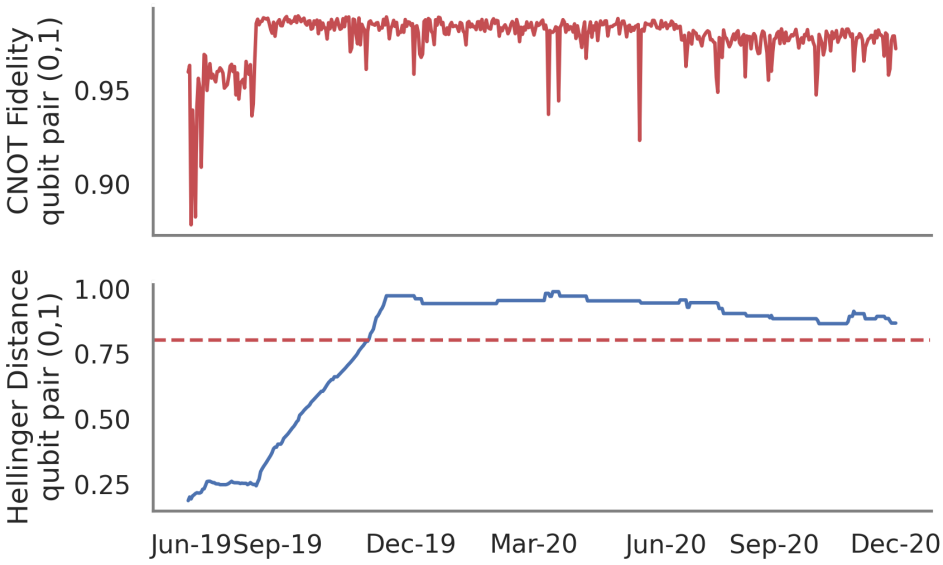
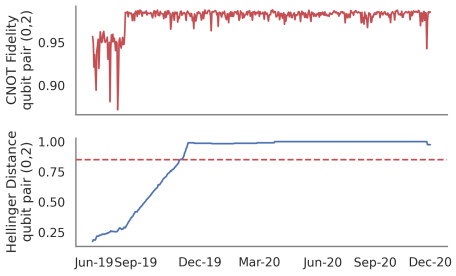


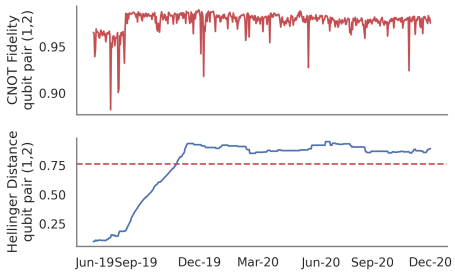
Figure 4.5: Spatial stability of the SPAM fidelity for the ibm_yorktown device from May 2019 to December 2020, where the inset highlights the registers with the maximum distance.



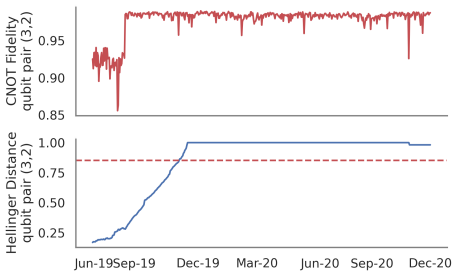
(a) CNOT (0, 1)



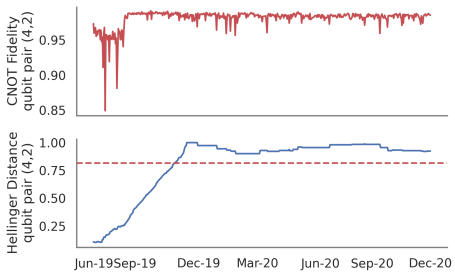
(b) CNOT (0, 2)



(c) CNOT (1, 2)

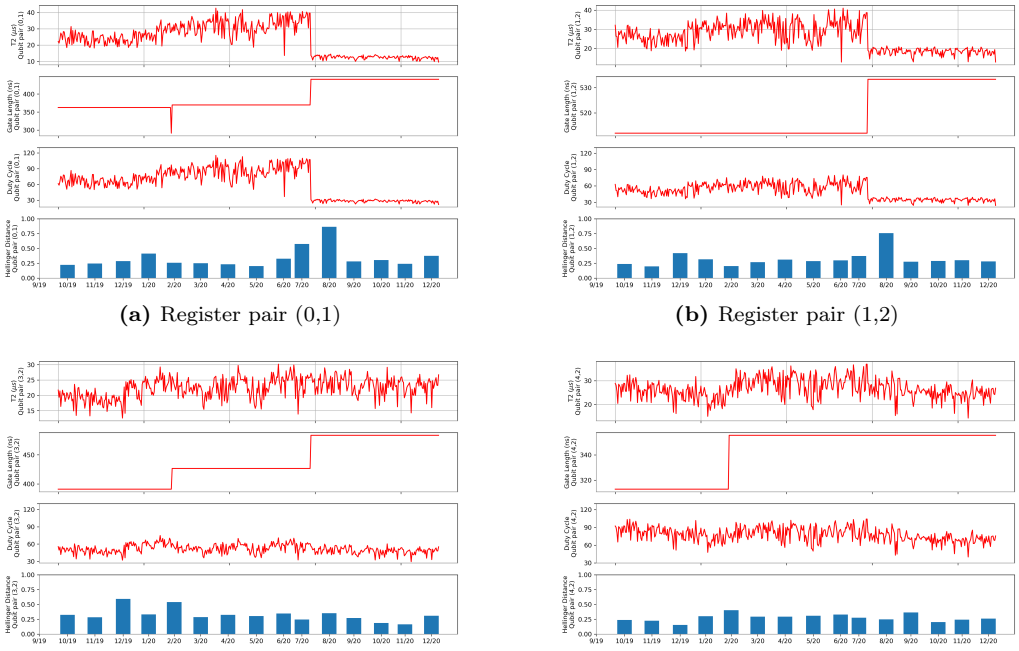


(d) CNOT (3, 2)



(e) CNOT (4, 2)

Figure 4.6: (a)-(e) Temporal stability of the gate fidelity F_G for the CNOT gate for sequential register pairs in the ibm_yorktown device from March 2019 to December 2020. The top panel shows the average F_G of the register pair and the bottom panel shows a running calculation of the Hellinger distance with respect to May 2019. The dashed red line is the median value.



(a) Register pair (0,1)

(b) Register pair (1,2)

(c) Register pair (2,3)

(d) Register pair (2,4)

Figure 4.7: (a)-(d) Temporal stability of the CNOT duty cycle for sequential register pairs in the yorktown device. The top panel shows the harmonic mean of the register decoherence time T_2 for the elements, the upper-middle panel shows the gate duration T_G, the lower-middle panel plots the corresponding duty cycle τ, and the bottom panel presents the Hellinger distance for the duty cycle averaged over a one-month window. The dashed red line is the median value.

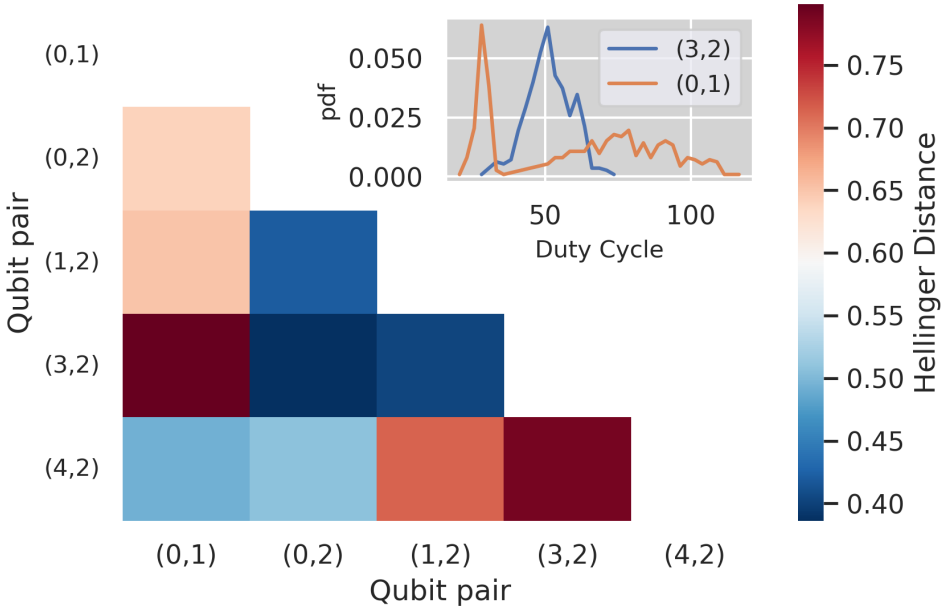


Figure 4.8: The spatial stability of the duty cycle τ for ibm_yorktown. The inset shows the experimental histograms for register pairs (0,1) and (2,3) which are separated by the largest Helligner distance of 0.789.

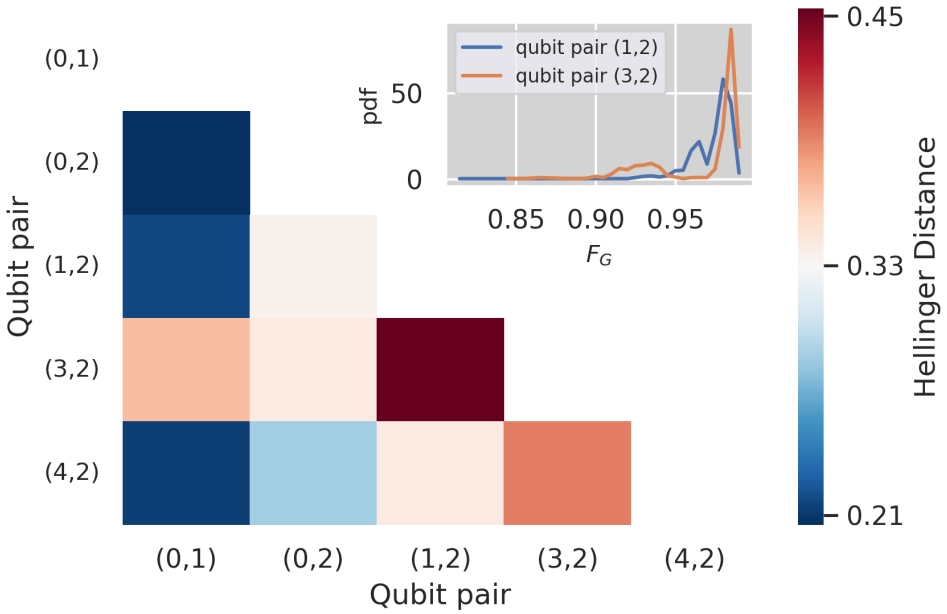


Figure 4.9: Spatial stability of the gate fidelity F_G for the CNOT gates of the ibm_yorktown device from March 2019 to December 2020. The inset shows the distribution of gate fidelities for pairs (1,2) and (3,2), which yield a Hellinger distance 0.467.

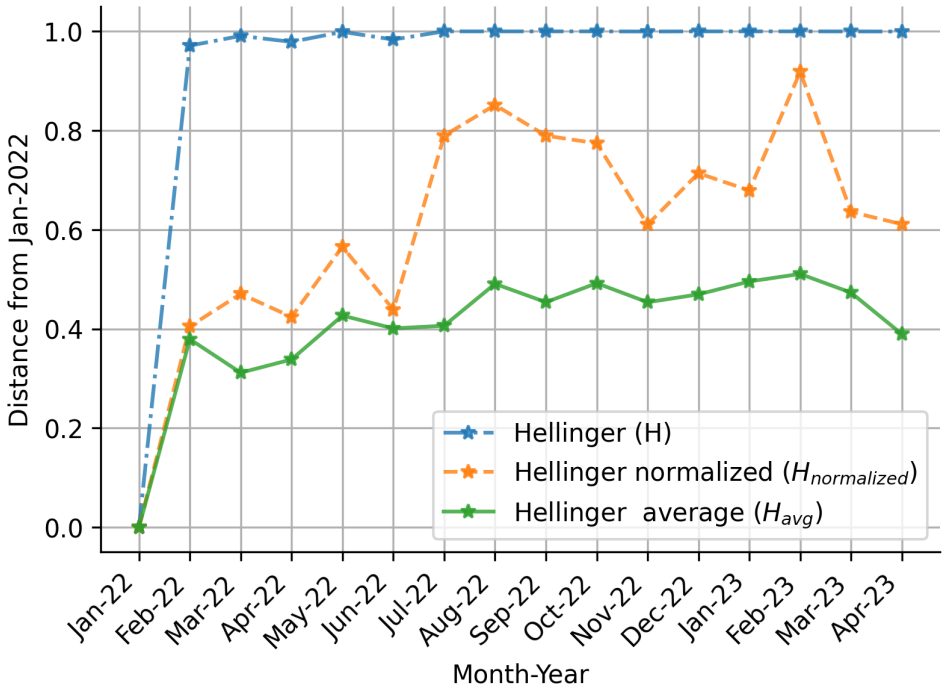


Figure 4.10: The figure presents results of reliability testing on a transmon platform over 16 months. Plots of the three Hellinger distance measures, with the unmodified measure H (blue line) being insensitive due to dimensionality issues. The blue and orange lines both capture the correlation structure of the joint distribution (with the orange line being normalized to enhance discrimination power), while the green line lacks correlation capture. The normalized measure $H_{\text{normalized}}$ ranges between 0.41 and 0.92, while the average measure H_{avg} varies between 0.431 and 0.51. The latter captures monthly variations in the marginal Hellinger distance for each of the 16 error parameters but fails to account for correlations.

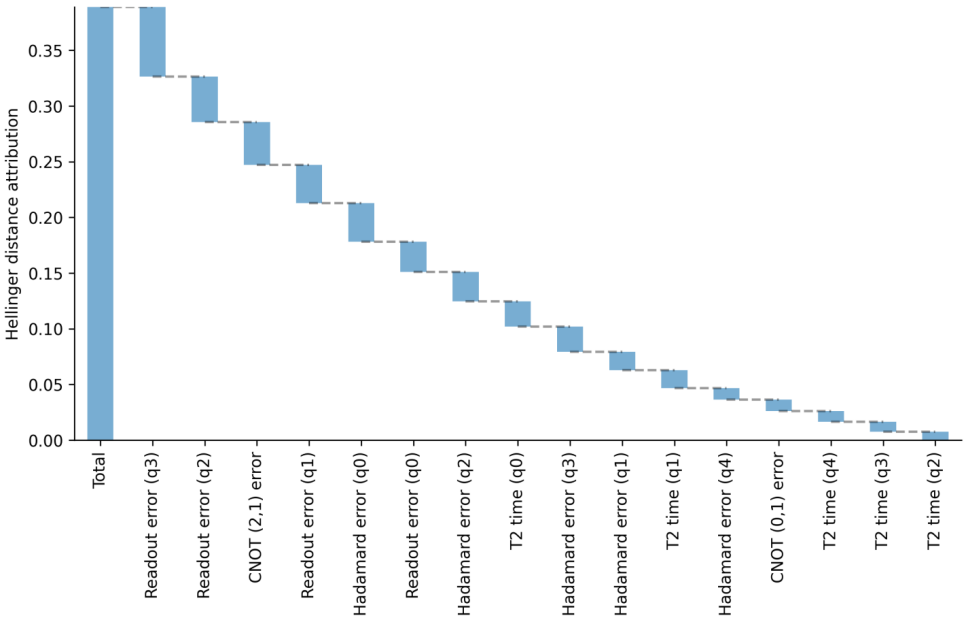


Figure 4.11: Decomposition of the sources of quantum noise non-stationarity. The degree of non-stationarity varies amongst the sources. The plot shows the contributions made by each noise type to the composite Hellinger distance (a measure of the degree of non-stationarity). The Hellinger distance measures the statistical distance between the joint distribution of the noise observed in Apr-2023 to the joint distribution of the noise observed in Jan-2022. The various noise types contribute varying percentages but no single term dominates the sum.

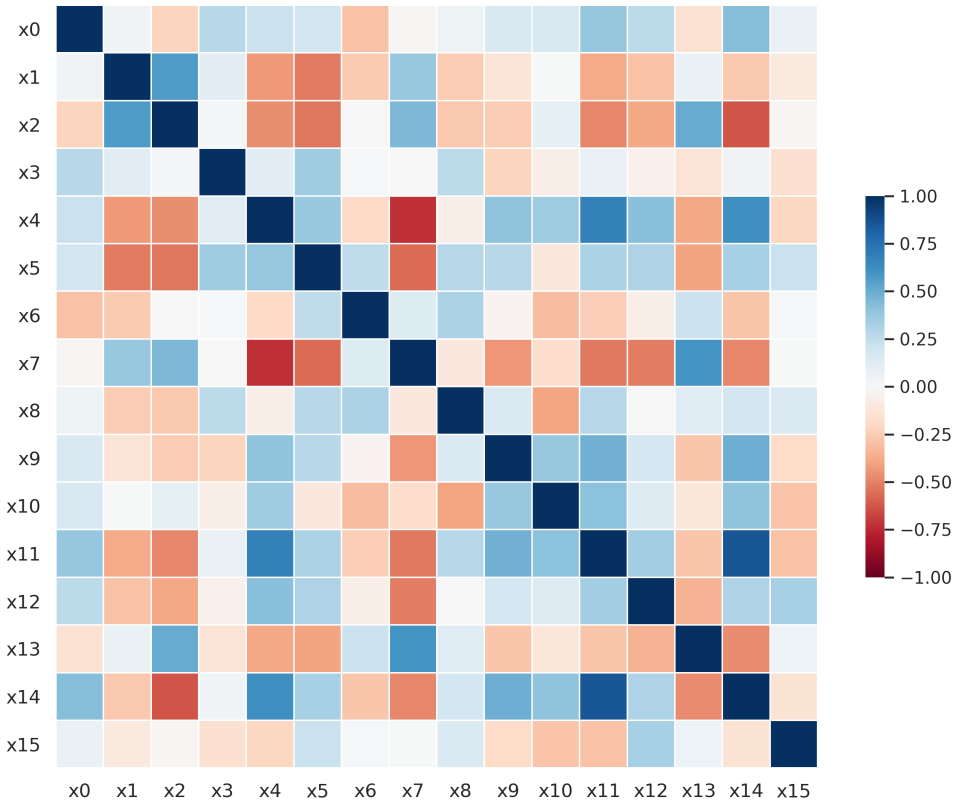


Figure 4.12: The Pearson correlation coefficients between the 16 characterization parameters as observed in April-2023. Dark blue and dark red colors represent the Pearson coefficients 1 and -1, respectively. The axes labels $(x0, x1, \cdots x15)$ correspond to various noise sources as listed in Table 5.2.

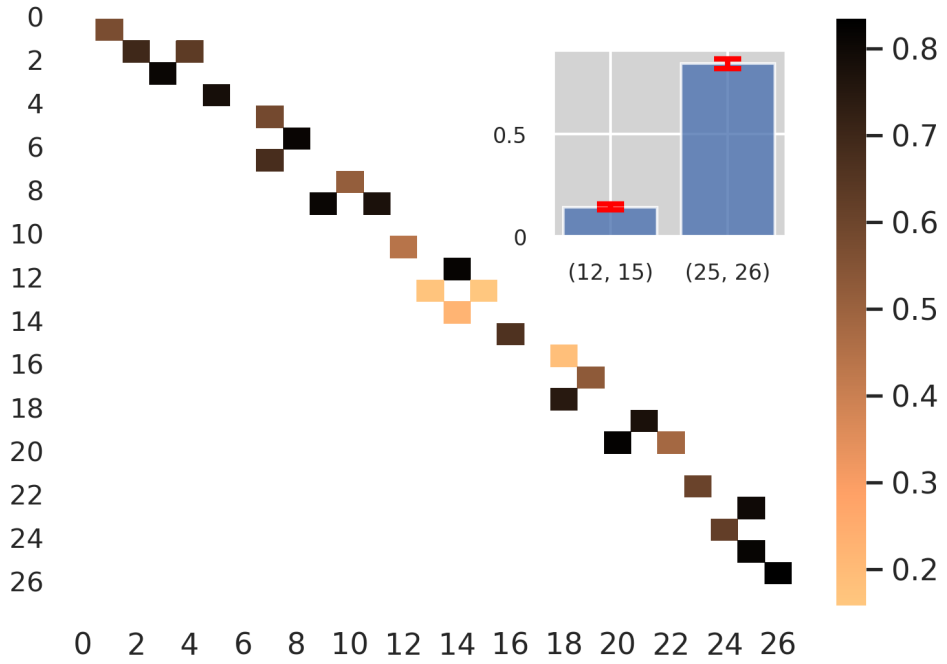


Figure 4.13: Normalized mutual information of register pairs in the toronto device sampled 11:00-11:30 PM (UTC-5) on 11 December 2020. Data corresponds to register pairs prepared in the Bell state and the inset shows the range of the lowest and highest values for the normalized mutual information.

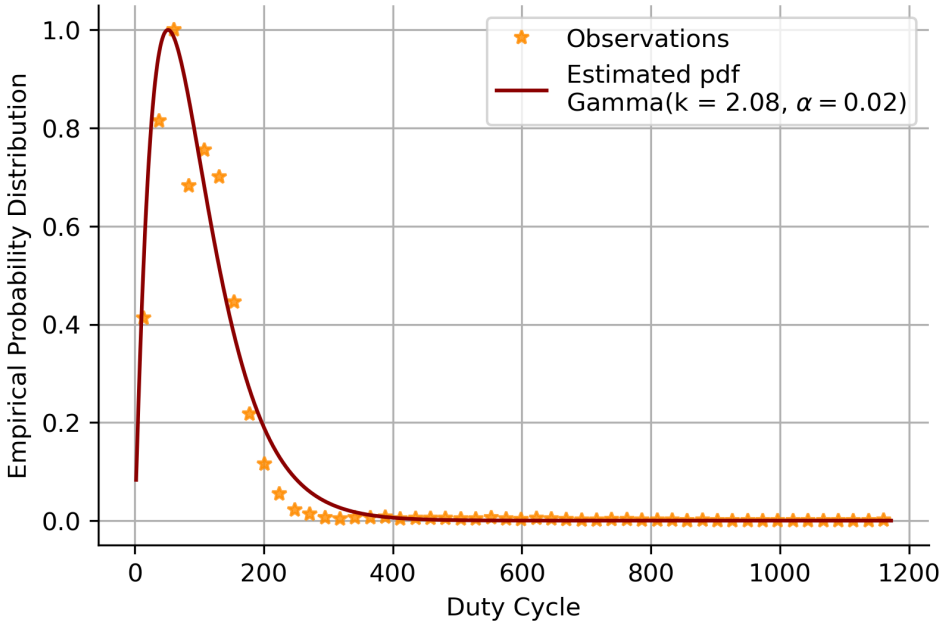


Figure 4.14: Experimentally observed duty cycle (τ) for the CNOT gate for washington.

Table 4.3: Gate lookup table

Label 0 = (0, 14)	Label 1 = (0, 1)	Label 2 = (100, 101)	Label 3 = (100, 110)	Label 4 = (100, 99)
Label 5 = (100, 101)	Label 6 = (101, 102)	Label 7 = (101, 102)	Label 8 = (102, 103)	Label 9 = (102, 92)
Label 10 = (102, 103)	Label 11 = (103, 104)	Label 12 = (103, 104)	Label 13 = (104, 105)	Label 14 = (104, 111)
Label 15 = (104, 105)	Label 16 = (105, 106)	Label 17 = (105, 106)	Label 18 = (106, 107)	Label 19 = (106, 93)
Label 20 = (106, 107)	Label 21 = (107, 108)	Label 22 = (107, 108)	Label 23 = (108, 112)	Label 24 = (109, 114)
Label 25 = (109, 96)	Label 26 = (10, 11)	Label 27 = (10, 9)	Label 28 = (100, 110)	Label 29 = (110, 118)
Label 30 = (104, 111)	Label 31 = (111, 122)	Label 32 = (108, 112)	Label 33 = (112, 126)	Label 34 = (113, 114)
Label 35 = (109, 114)	Label 36 = (113, 114)	Label 37 = (114, 115)	Label 38 = (114, 115)	Label 39 = (115, 116)
Label 40 = (115, 116)	Label 41 = (116, 117)	Label 42 = (116, 117)	Label 43 = (117, 118)	Label 44 = (110, 118)
Label 45 = (117, 118)	Label 46 = (118, 119)	Label 47 = (118, 119)	Label 48 = (119, 120)	Label 49 = (10, 11)
Label 50 = (11, 12)	Label 51 = (119, 120)	Label 52 = (120, 121)	Label 53 = (120, 121)	Label 54 = (121, 122)
Label 55 = (111, 122)	Label 56 = (121, 122)	Label 57 = (122, 123)	Label 58 = (122, 123)	Label 59 = (123, 124)
Label 60 = (123, 124)	Label 61 = (124, 125)	Label 62 = (124, 125)	Label 63 = (125, 126)	Label 64 = (112, 126)
Label 65 = (125, 126)	Label 66 = (11, 12)	Label 67 = (12, 13)	Label 68 = (12, 17)	Label 69 = (12, 13)
Label 70 = (0, 14)	Label 71 = (14, 18)	Label 72 = (15, 22)	Label 73 = (15, 4)	Label 74 = (16, 26)
Label 75 = (16, 8)	Label 76 = (12, 17)	Label 77 = (17, 30)	Label 78 = (14, 18)	Label 79 = (18, 19)
Label 80 = (18, 19)	Label 81 = (19, 20)	Label 82 = (0, 1)	Label 83 = (1, 2)	Label 84 = (19, 20)
Label 85 = (20, 21)	Label 86 = (20, 33)	Label 87 = (20, 21)	Label 88 = (21, 22)	Label 89 = (15, 22)
Label 90 = (21, 22)	Label 91 = (22, 23)	Label 92 = (22, 23)	Label 93 = (23, 24)	Label 94 = (23, 24)
Label 95 = (24, 25)	Label 96 = (24, 34)	Label 97 = (24, 25)	Label 98 = (25, 26)	Label 99 = (16, 26)
Label 100 = (25, 26)	Label 101 = (26, 27)	Label 102 = (26, 27)	Label 103 = (27, 28)	Label 104 = (27, 28)
Label 105 = (28, 29)	Label 106 = (28, 35)	Label 107 = (28, 29)	Label 108 = (29, 30)	Label 109 = (1, 2)
Label 110 = (2, 3)	Label 111 = (17, 30)	Label 112 = (29, 30)	Label 113 = (30, 31)	Label 114 = (30, 31)
Label 115 = (31, 32)	Label 116 = (31, 32)	Label 117 = (32, 36)	Label 118 = (20, 33)	Label 119 = (33, 39)
Label 120 = (24, 34)	Label 121 = (34, 43)	Label 122 = (28, 35)	Label 123 = (35, 47)	Label 124 = (32, 36)
Label 125 = (36, 51)	Label 126 = (37, 38)	Label 127 = (37, 52)	Label 128 = (37, 38)	Label 129 = (38, 39)
Label 130 = (33, 39)	Label 131 = (38, 39)	Label 132 = (39, 40)	Label 133 = (2, 3)	Label 134 = (3, 4)
Label 135 = (39, 40)	Label 136 = (40, 41)	Label 137 = (40, 41)	Label 138 = (41, 42)	Label 139 = (41, 53)
Label 140 = (41, 42)	Label 141 = (42, 43)	Label 142 = (34, 43)	Label 143 = (42, 43)	Label 144 = (43, 44)
Label 145 = (43, 44)	Label 146 = (44, 45)	Label 147 = (44, 45)	Label 148 = (45, 46)	Label 149 = (45, 54)
Label 150 = (45, 46)	Label 151 = (46, 47)	Label 152 = (35, 47)	Label 153 = (46, 47)	Label 154 = (47, 48)
Label 155 = (47, 48)	Label 156 = (48, 49)	Label 157 = (48, 49)	Label 158 = (49, 50)	Label 159 = (49, 55)
Label 160 = (15, 4)	Label 161 = (3, 4)	Label 162 = (4, 5)	Label 163 = (49, 50)	Label 164 = (50, 51)
Label 165 = (36, 51)	Label 166 = (50, 51)	Label 167 = (37, 52)	Label 168 = (52, 56)	Label 169 = (41, 53)
Label 170 = (53, 60)	Label 171 = (45, 54)	Label 172 = (54, 64)	Label 173 = (49, 55)	Label 174 = (55, 68)
Label 175 = (52, 56)	Label 176 = (56, 57)	Label 177 = (56, 57)	Label 178 = (57, 58)	Label 179 = (57, 58)
Label 180 = (58, 59)	Label 181 = (58, 71)	Label 182 = (58, 59)	Label 183 = (59, 60)	Label 184 = (4, 5)
Label 185 = (5, 6)	Label 186 = (53, 60)	Label 187 = (59, 60)	Label 188 = (60, 61)	Label 189 = (60, 61)
Label 190 = (61, 62)	Label 191 = (61, 62)	Label 192 = (62, 63)	Label 193 = (62, 72)	Label 194 = (62, 63)
Label 195 = (63, 64)	Label 196 = (54, 64)	Label 197 = (63, 64)	Label 198 = (64, 65)	Label 199 = (64, 65)
Label 200 = (65, 66)	Label 201 = (65, 66)	Label 202 = (66, 67)	Label 203 = (66, 73)	Label 204 = (66, 67)
Label 205 = (67, 68)	Label 206 = (55, 68)	Label 207 = (67, 68)	Label 208 = (68, 69)	Label 209 = (68, 69)
Label 210 = (69, 70)	Label 211 = (5, 6)	Label 212 = (6, 7)	Label 213 = (69, 70)	Label 214 = (70, 74)
Label 215 = (58, 71)	Label 216 = (71, 77)	Label 217 = (62, 72)	Label 218 = (72, 81)	Label 219 = (66, 73)
Label 220 = (73, 85)	Label 221 = (70, 74)	Label 222 = (74, 89)	Label 223 = (75, 76)	Label 224 = (75, 90)
Label 225 = (75, 76)	Label 226 = (76, 77)	Label 227 = (71, 77)	Label 228 = (76, 77)	Label 229 = (77, 78)
Label 230 = (77, 78)	Label 231 = (78, 79)	Label 232 = (78, 79)	Label 233 = (79, 80)	Label 234 = (79, 91)
Label 235 = (6, 7)	Label 236 = (7, 8)	Label 237 = (79, 80)	Label 238 = (80, 81)	Label 239 = (72, 81)
Label 240 = (80, 81)	Label 241 = (81, 82)	Label 242 = (81, 82)	Label 243 = (82, 83)	Label 244 = (82, 83)
Label 245 = (83, 84)	Label 246 = (83, 92)	Label 247 = (83, 84)	Label 248 = (84, 85)	Label 249 = (73, 85)
Label 250 = (84, 85)	Label 251 = (85, 86)	Label 252 = (85, 86)	Label 253 = (86, 87)	Label 254 = (86, 87)
Label 255 = (87, 88)	Label 256 = (87, 93)	Label 257 = (87, 88)	Label 258 = (88, 89)	Label 259 = (74, 89)
Label 260 = (88, 89)	Label 261 = (16, 8)	Label 262 = (7, 8)	Label 263 = (8, 9)	Label 264 = (75, 90)
Label 265 = (90, 94)	Label 266 = (79, 91)	Label 267 = (91, 98)	Label 268 = (102, 92)	Label 269 = (83, 92)
Label 270 = (106, 93)	Label 271 = (87, 93)	Label 272 = (90, 94)	Label 273 = (94, 95)	Label 274 = (94, 95)
Label 275 = (95, 96)	Label 276 = (109, 96)	Label 277 = (95, 96)	Label 278 = (96, 97)	Label 279 = (96, 97)
Label 280 = (97, 98)	Label 281 = (91, 98)	Label 282 = (97, 98)	Label 283 = (98, 99)	Label 284 = (100, 99)
Label 285 = (98, 99)	Label 286 = (10, 9)	Label 287 = (8, 9)		

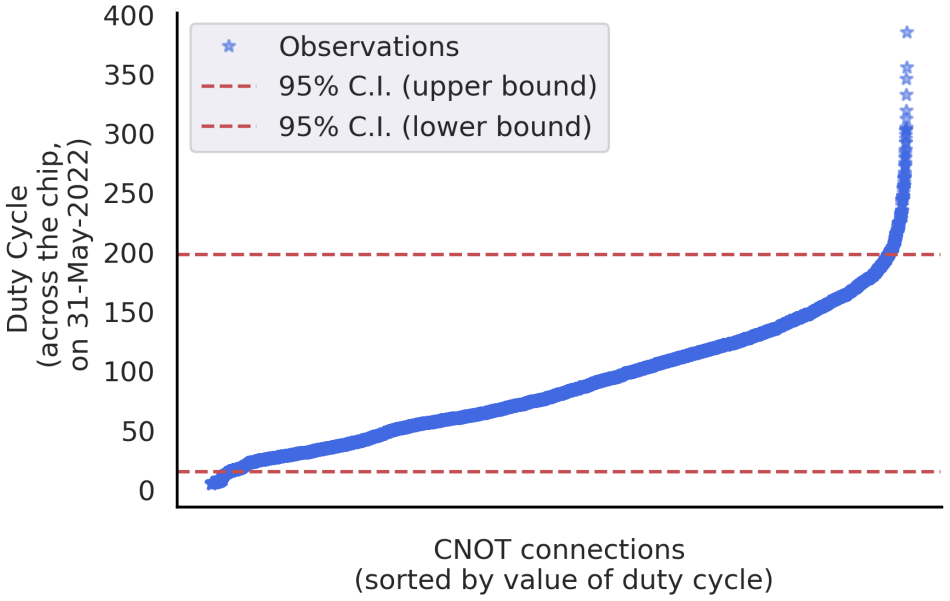


Figure 4.15: The 95% temporal confidence interval (vertical orange lines) for CNOT duty cycle (τ) for the physical nearest-neighbor connections of ibm_washington. The data-set contains the values for all the 144 physical CNOT gates of ibm_washington between 1-Dec-2021 to 31-May-2022. The dashed black line represents the mean.

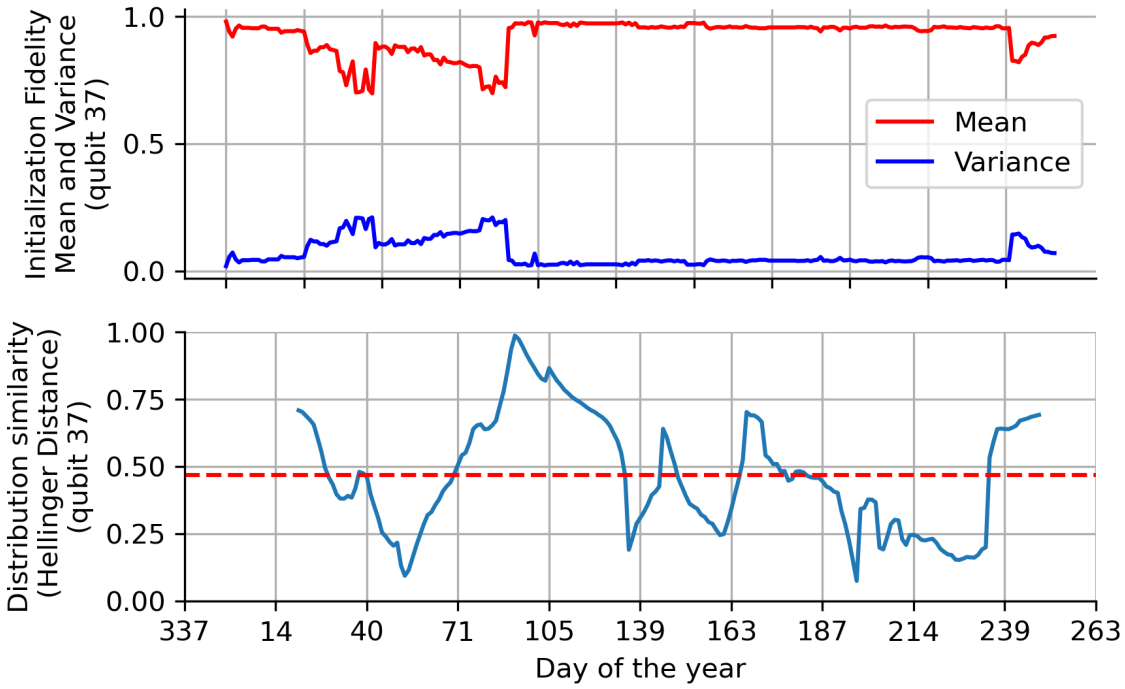


Figure 4.16: Temporal reliability of the SPAM fidelity F_{SPAM} of register element $q = 37$

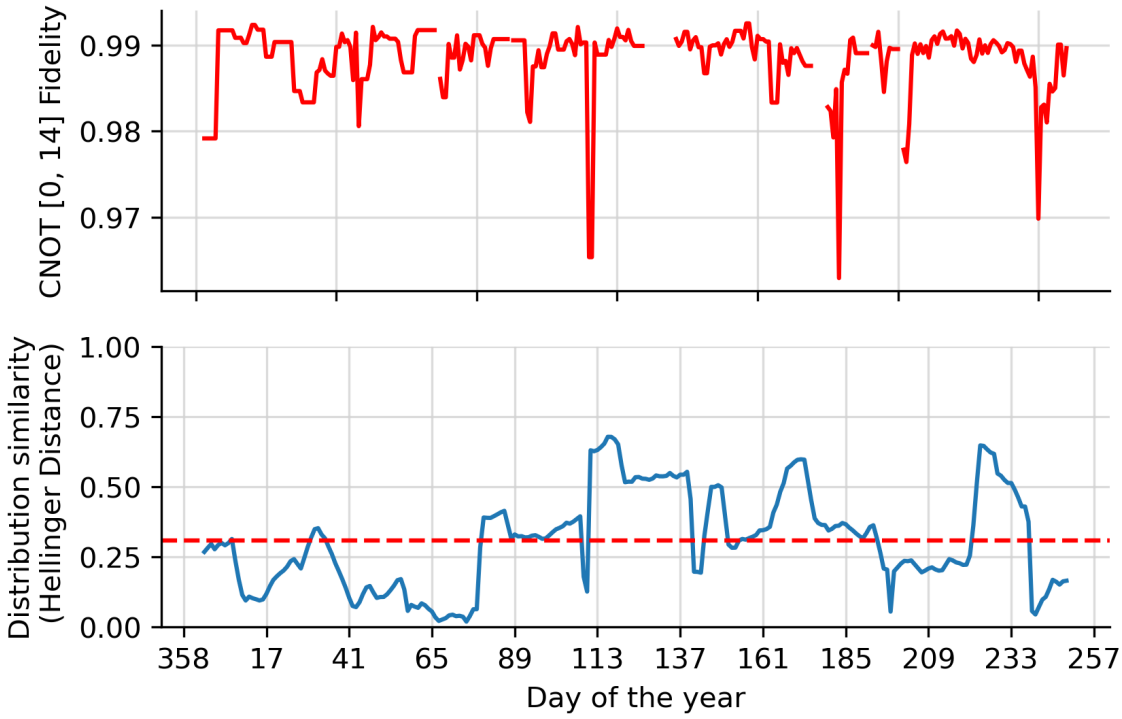


Figure 4.17: Temporal reliability of the gate fidelity F_G for the CNOT gate for register pairs 0 and 14.

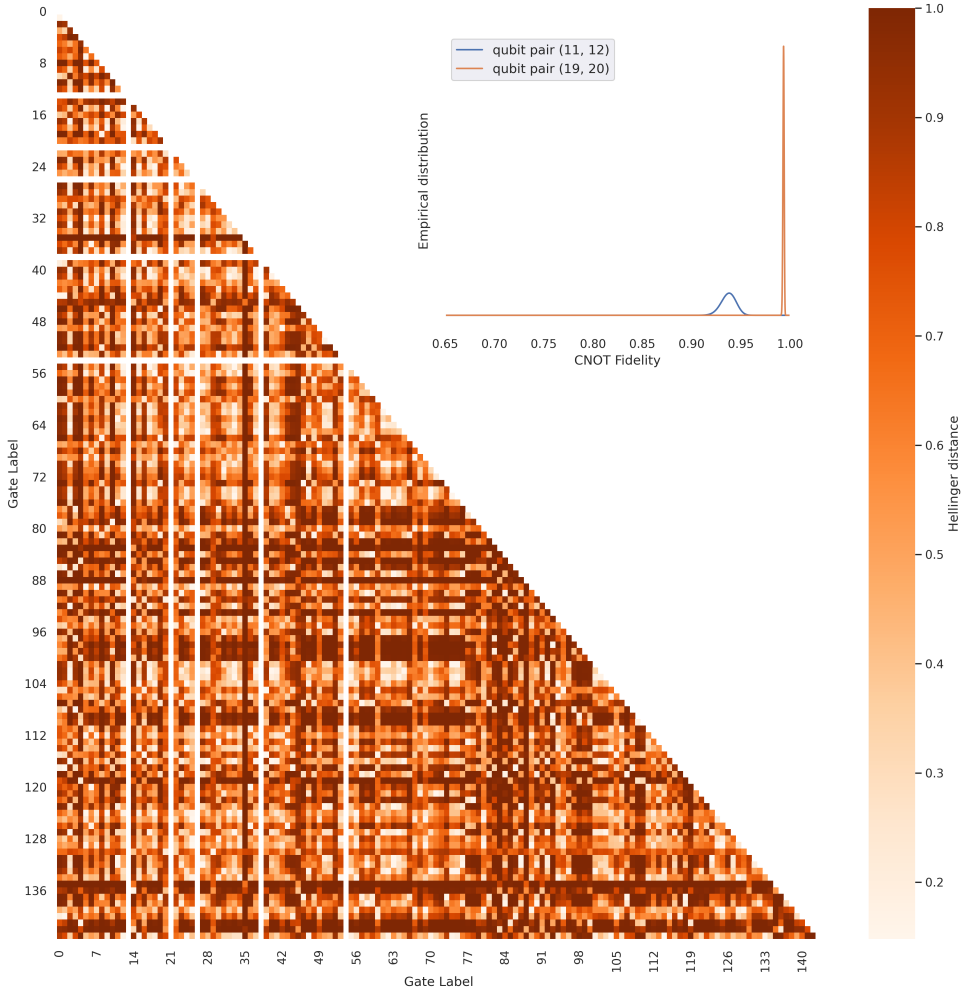


Figure 4.18: Spatial reliability of the gate fidelity F_G for the CNOT gates of the ibm_washington device. The heat map shows the Hellinger distance between the distributions for the gate fidelities of nearest-neighbor connections. Only the lower triangular matrix is shown to avoid redundancy. The upper triangular matrix as well as any data gaps in the lower triangular matrix are colored white. The inset shows the distribution of gate fidelities for pairs $(11, 12)$ and $(19, 20)$, which yield a Hellinger distance 0.99. This metric captures the probability that the quantum processor register is spatially dissimilar. The estimation of the distribution for CNOT-i utilizes data for Sep-2022.

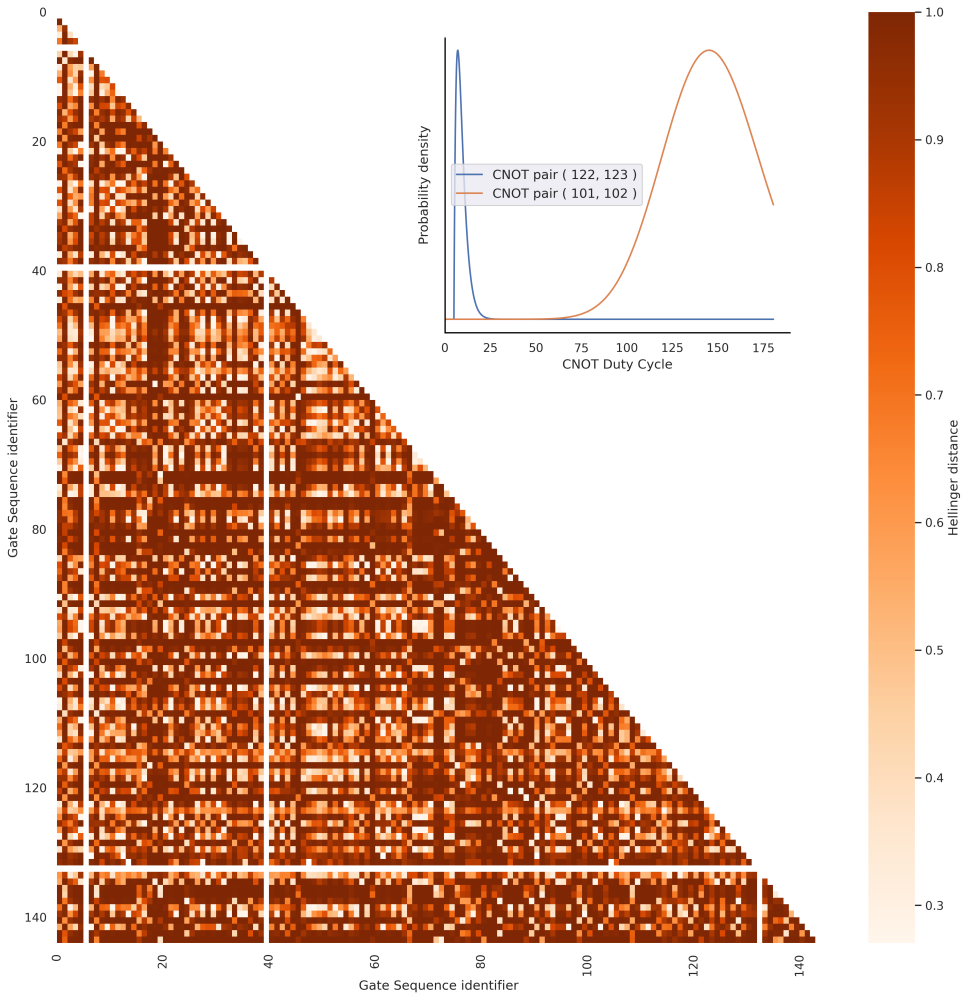


Figure 4.19: Spatial reliability of the duty cycle τ for the CNOT gates of the ibm_washington device. The heat map shows the Hellinger distance between the distributions for the duty cycles of nearest-neighbor connections. The inset shows the distribution of duty cycle for pairs (46, 47) and (96, 109), which yield a Hellinger distance 0.99. This metric captures the probability that the quantum processor register is spatially dissimilar. The estimation of the distribution for CNOT-i utilizes duty-cycle data for Sep-2022.

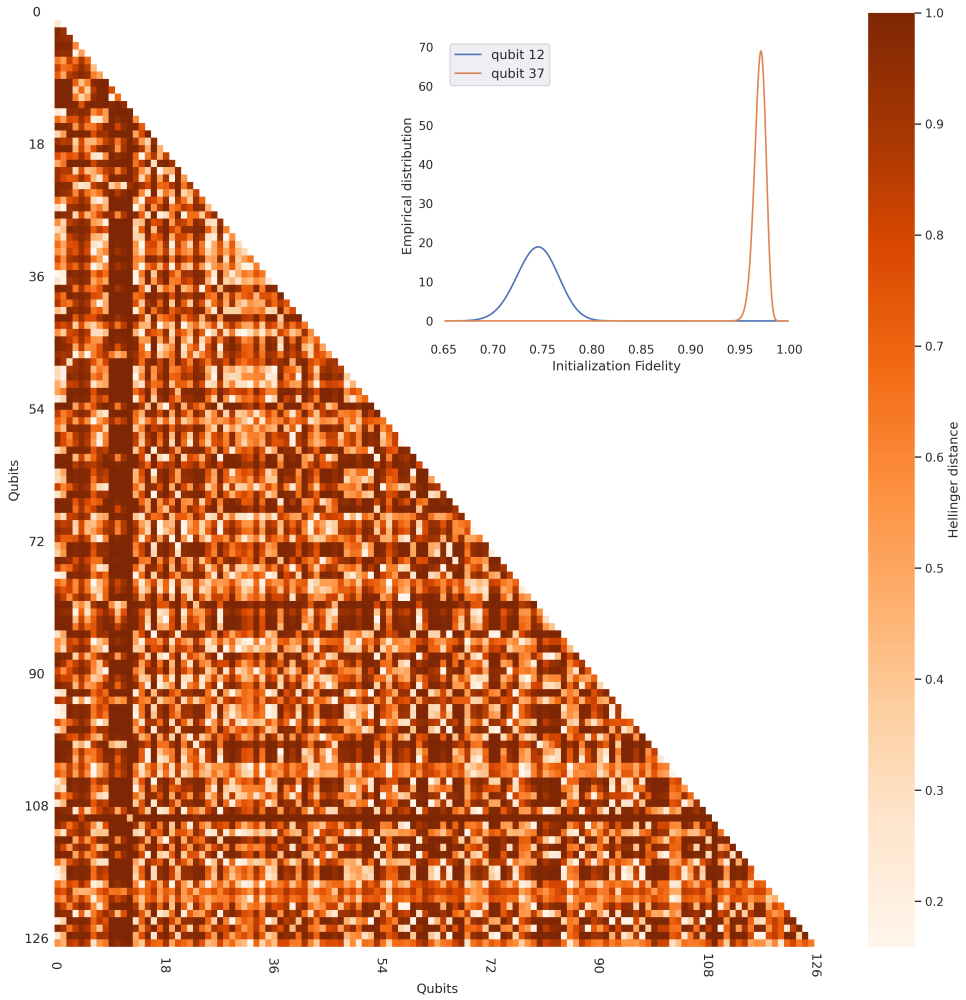


Figure 4.20: Spatial reliability of the SPAM fidelity F_{SPAM} for the ibm_washington device. The heat map shows the Hellinger distance between the distributions for each register pair. This metric captures the probability that the quantum processor register is spatially dissimilar. The inset shows the distributions of F_{SPAM} for registers 12 and 37, which represent distance=0.99 due to minimal overlap. The estimation of the distribution for qubit-i utilizes data for Sep-2022.

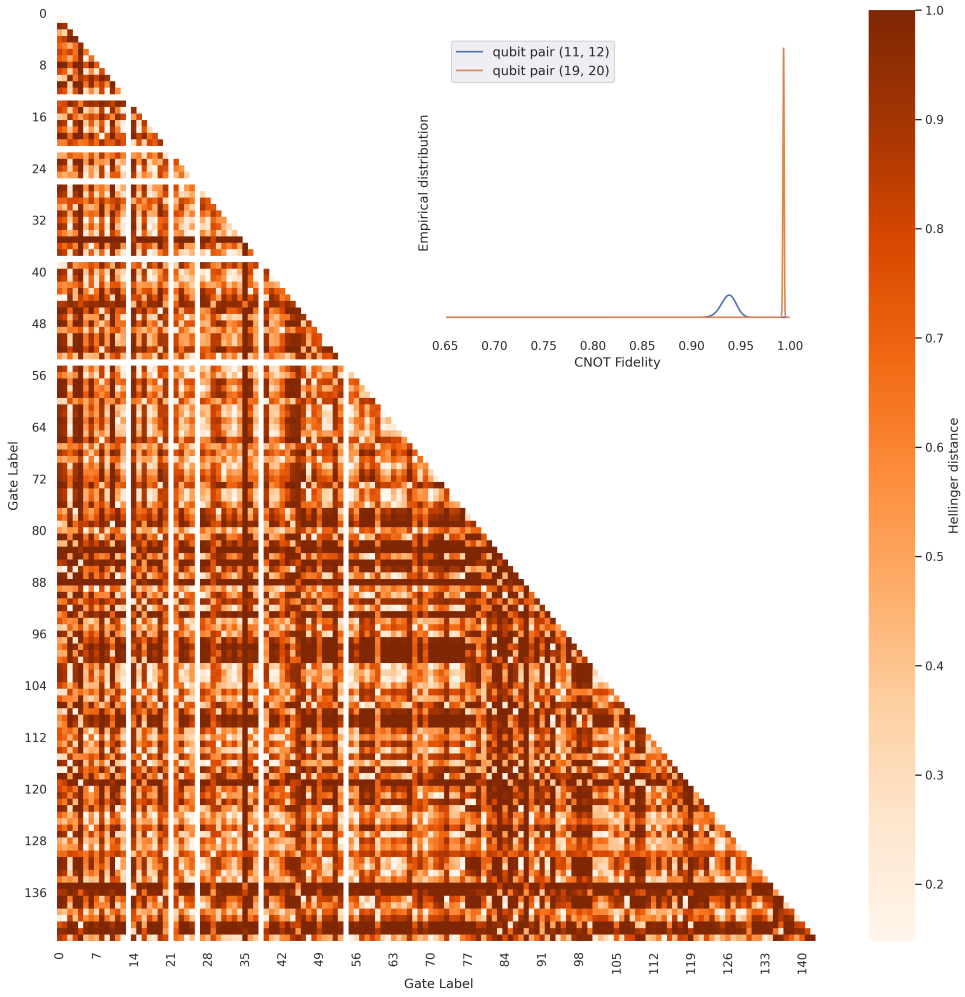


Figure 4.21: Spatial reliability of the gate fidelity F_G for the CNOT gates of the ibm_washington device. The heat map shows the Hellinger distance between the distributions for the gate fidelities of nearest-neighbor connections. Only the lower triangular matrix is shown to avoid redundancy. The upper triangular matrix as well as any data gaps in the lower triangular matrix are colored white. The inset shows the distribution of gate fidelities for pairs $(11, 12)$ and $(19, 20)$, which yield a Hellinger distance 0.99. This metric captures the probability that the quantum processor register is spatially dissimilar. The estimation of the distribution for CNOT-i utilizes data for Sep-2022.

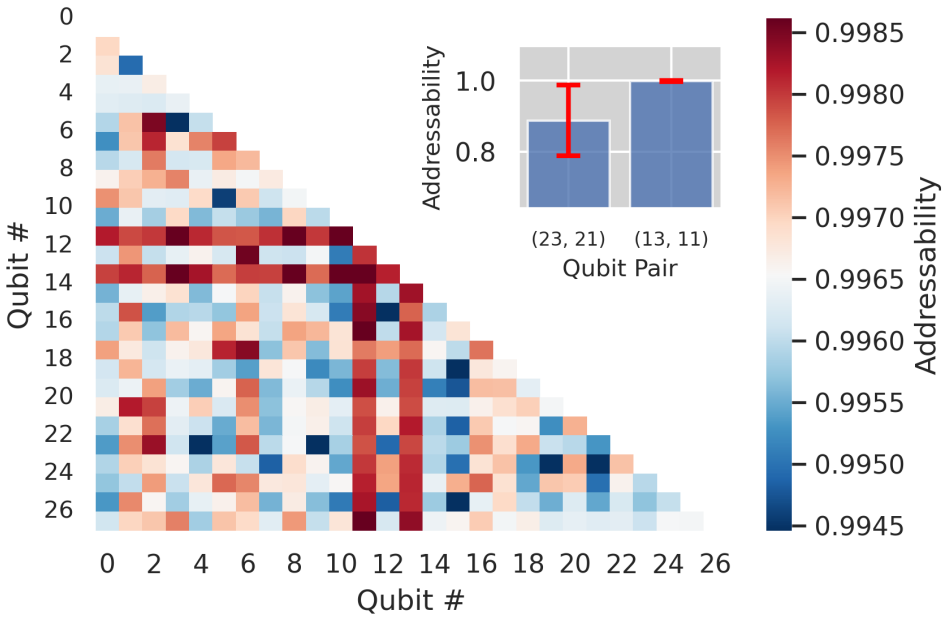


Figure 4.22: Addressability of register pairs in the ibm_toronto device sampled 08:00-08:30 AM (UTC-5) on 11 December 2020. This data corresponds to the register prepared in the separable fiducial state. The inset shows the range i.e. the lowest and highest values for addressability. The average of (23,21) is the lowest value at 0.887 while all other values lie in the range [0.992, 1). The outlier is the only value that does not appear in the plot.

Chapter 5

Bounds on stability of program outcomes

In Chapter 3, we developed a performance assessment framework for noisy quantum outputs, using accuracy, reproducibility, device reliability, and outcome stability. In this chapter, we aim to establish bounds for this framework using available device characterization data. Section 1 asks, given an ϵ-bound on histogram accuracy (measured in computational basis), how to bound a proxy parameter that encapsulates multiple device noise parameters. Section 2 asks, what is the minimum sample size required to ensure histogram reproducibility with $1 - \delta$ confidence. Section 3 bounds outcome stability in terms of the distance between time-varying noise densities. Section 4 bounds device reliability metric to attain an ϵ-stable outcome.

5.1 Sample bounds on accuracy

In this section, our aim is to address the following challenge posed by an ϵ-bound on the accuracy of digital histograms obtained from a noisy test circuit (measured in the computational basis). Can we formulate and bound a proxy parameter γ_D that encompasses the array of noise parameters characterizing a noisy device D?

$$\gamma_D \leq \gamma_{\max}. \tag{5.1}$$

The importance of such a proxy lies in its ability to streamline high-dimensional noise analysis.

Consider an n qubit state $|\psi\rangle$ prepared as a uniform superposition across the 2^n computational basis states $\{|v\rangle\}$ as:

$$|\psi\rangle = 2^{-n/2} \sum_{v\in\{0,1\}^n} |v\rangle\,, \tag{5.2}$$

which is the output (in the noiseless limit) from the circuit shown in Fig. 5.1 i.e. $|\psi\rangle = \mathbb{H}^{\otimes n}\,|0\rangle^{\otimes n}$. For our experiments, we use a $n = 27$ qubit register. The noiseless distribution for this circuit is: $p_v^{\text{noiseless}} = 2^{-n} \quad \forall v \in \{0, \cdots, 2^n - 1\}$. We assume that gate errors and SPAM noise capture the principal sources of noise in this circuit and ignore inter-qubit cross-talk. Let $\mathbb{I}, \mathbb{X}, \mathbb{Y}$, and $\mathbb{Z}$ denote the 2×2 identity matrix, Pauli-X matrix, Pauli-Y matrix, and the Pauli-Z matrix, respectively:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbb{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbb{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.3}$$

Let $R_Y(\theta)$ denote rotation by an angle (θ) about the Y-axis on the Bloch sphere[11]:

$$R_Y(\theta) = e^{-i\frac{\theta}{2}Y} = \cos\frac{\theta}{2}\mathbb{I} - i\sin\frac{\theta}{2}\mathbb{Y} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \tag{5.4}$$

A noiseless Hadamard gate is given by:

$$H = R_Y\left(\frac{\pi}{2}\right)\mathbb{Z} = \begin{pmatrix} \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & -\cos\frac{\pi}{4} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{5.5}$$

We model an over- or under-rotated Hadamard gate ($\tilde{\mathbb{H}}$) by the unitary:

$$\tilde{\mathbb{H}} = \begin{pmatrix} \cos\left(\frac{\pi}{4} + \mathrm{x}\right) & \sin\left(\frac{\pi}{4} + \mathrm{x}\right) \\ \sin\left(\frac{\pi}{4} + \mathrm{x}\right) & -\cos\left(\frac{\pi}{4} + \mathrm{x}\right) \end{pmatrix} \tag{5.6}$$

where x is a small implementation error in radians ($\mathrm{x} \ll \pi/4$). The operator representation for a unitary control error [126] has only one term which can be seen as follows. Write the noisy unitary $\tilde{\mathcal{U}}$ as: $\tilde{\mathcal{U}} = \tilde{\mathcal{U}}\mathcal{U}^\dagger\mathcal{U}$ where $\mathcal{U}$ is the noiseless unitary. Thus, $\rho' = \tilde{\mathcal{U}}\rho\tilde{\mathcal{U}}^\dagger = (\tilde{\mathcal{U}}\mathcal{U}^\dagger)(\mathcal{U}\rho\mathcal{U}^\dagger)(\mathcal{U}\tilde{\mathcal{U}}^\dagger) = M(\mathcal{U}\rho\mathcal{U}^\dagger)M^\dagger$ where M is the operator representing the unitary control error that arises due to imperfections in the control system. For our circuit,

$$M = \tilde{\mathbb{H}}\mathbb{H}^\dagger = \begin{pmatrix} \cos\mathrm{x} & -\sin\mathrm{x} \\ \sin\mathrm{x} & \cos\mathrm{x} \end{pmatrix}, \quad \tilde{\mathbb{H}} = \frac{1}{\sqrt{2}}\begin{pmatrix} \cos\mathrm{x} - \sin\mathrm{x} & \cos\mathrm{x} + \sin\mathrm{x} \\ \cos\mathrm{x} + \sin\mathrm{x} & -\cos\mathrm{x} + \sin\mathrm{x} \end{pmatrix} \tag{5.7}$$

In the absence of SPAM noise, when we initialize a qubit in the ground state, subject it to a noisy Hadamard gate, and measure in the $\mathbb{Z}$-basis, we get the probabilities for observing the outputs 0 and 1 as:

$$\Pr\left(|v\rangle = |i\rangle\right) = \left(1 - (-1)^i \sin 2\mathrm{x}\right)/2, \quad i \in \{0, 1\} \tag{5.8}$$

We next consider what happens when the Hadamard gate is followed by a noisy measurement. The SPAM channel can be characterized as a quantum channel[127, 28, 128] using two parameters f_0 and f_1 for each qubit. The first parameter (f_0) defines the probability of observing 0 post readout when the channel input state is $|0\rangle$, and the second (f_1) defines the probability of observing 1 post readout when the channel input state is $|1\rangle$.

A classical representation for the single qubit SPAM channel is:

$$\mathbb{P}^{\text{out}} = \Lambda \mathbb{P}^{\text{in}} \tag{5.9}$$

where Λ is the SPAM error matrix with elements Λ_{ij} = probability of observing $|i\rangle$ when the input to channel is $|j\rangle$ $(i, j \in \{0, 1\})$:

$$\Lambda = \begin{pmatrix} f_0 & 1 - f_1 \\ f_1 & 1 - f_0 \end{pmatrix} \tag{5.10}$$

Equivalently, the quantum channel representation for a single qubit noisy measurement has two Kraus operators, M_0 and M_1, and can be specified by a super-operator ($\mathcal{E}$), whose action on the quantum state is as follows [105]:

$$\begin{aligned} \mathcal{E}(\rho) =& M_0 \rho M_0^\dagger + M_1 \rho M_1^\dagger \\ M_0 =& \sqrt{f_0}\,|0\rangle\langle 0| + \sqrt{1 - f_1}\,|1\rangle\langle 1| \\ M_1 =& \sqrt{1 - f_0}\,|0\rangle\langle 0| + \sqrt{f_1}\,|1\rangle\langle 1| \end{aligned} \tag{5.11}$$

This is equivalent to Eqn. 5.10, when you consider the action of the measurement operators $\{|i\rangle\langle i|\}$, given by:

$$Pr(|v\rangle = |i\rangle) = \mathrm{Tr}\left[|i\rangle\langle i|\,\mathcal{E}(\rho)\right]. \tag{5.12}$$

Let Pr(0) be the probability of observing 0 when we prepare a qubit in the ground state, subject it to a Hadamard gate, and then measure it. Let Pr(1) be the corresponding probability of observing 1 for the same experiment (i.e., we prepare a qubit in the ground state, subject it to a Hadamard gate, and then measure it). Additionally, let f be the average SPAM fidelity and $\varepsilon^{\text{SPAM}}$ be the SPAM

fidelity asymmetry:

$$f = \frac{f_0 + f_1}{2} \tag{5.13}$$

$$\varepsilon^{\text{SPAM}} = f_0 - f_1. \tag{5.14}$$

Thus, in the presence of SPAM noise, the probability of observing 0 and 1 for each qubit in Fig. 5.1 is given by [105]:

$$Pr(|v\rangle = |i\rangle) = \text{Tr}\left[|i\rangle\langle i|\,\mathcal{E}(\rho)\right] = \frac{1 + (-1)^i \gamma}{2} \tag{5.15}$$

where

$$\gamma = \varepsilon^{\text{SPAM}} - 2\sin 2\text{x}\left(f - \frac{1}{2}\right). \tag{5.16}$$

Let $(v_{n-1}v_{n-2}\cdots v_0)$ represent the n-bit string with $v_i \in \{0,1\}$, and let $v = \sum_{i=0}^{n-1} 2^i v_i$ be the decimal integer equivalent. In the absence of cross-talk between gates, $\mathbb{P}^{\text{noisy}} = \{p_v^{\text{noisy}}\}$ where:

$$p_v^{\text{noisy}} = \prod_{i=0}^{n-1} \left(\frac{1+\gamma_i}{2}\right)^{1-v_i} \left(\frac{1-\gamma_i}{2}\right)^{v_i} \tag{5.17}$$

and γ_i refers to the γ-parameter from Eqn. 5.16 for the i-th register element.

It follows then that the Bhattacharya coefficient[91] is given by:

$$\begin{aligned} BC(\mathbb{P}^{\text{noiseless}}, \mathbb{P}^{\text{noisy}}) &= \sum_{v=0}^{2^n-1} \sqrt{\frac{1}{2^n} \prod_{i=0}^{n-1} \left(\frac{1+\gamma_i}{2}\right)^{1-v_i} \left(\frac{1-\gamma_i}{2}\right)^{v_i}} \\ &= \frac{1}{2^n} \sum_{v=0}^{2^n-1} \prod_{i=0}^{n-1} \sqrt{1+\gamma_i}^{1-v_i} \sqrt{1-\gamma_i}^{v_i} \end{aligned} \tag{5.18}$$

For meeting an ϵ-accuracy constraint, we need:

$$\frac{1}{2^n} \sum_{v=0}^{2^n-1} \prod_{i=0}^{n-1} \sqrt{1+\gamma_i}^{1-v_i} \sqrt{1-\gamma_i}^{v_i} \leq \epsilon^2. \tag{5.19}$$

When assuming that ϵ is small, this yields:

$$\left|\varepsilon^{\text{SPAM}} - 2\sin 2\text{x}\left(f - \frac{1}{2}\right)\right| \leq \gamma_{\max} = 2(1-\epsilon^2)^{1/n}\sqrt{1-(1-\epsilon^2)^{2/n}} \tag{5.20}$$

The left side is the proxy γ_D and the right side is the upper bound under ϵ-accuracy constraint for the digital histogram, in the context of the superposition circuit.

Conversely, when armed with noise characterization data, we can estimate the best possible accuracy as:

$$\epsilon_{\min} \geq \frac{1}{2}\sqrt{\frac{n}{2}}\gamma_D. \tag{5.21}$$

5.1.1 Validation using device characterization data

We validated our bound using the device ibm_toronto, whose schematic is shown in Fig. 4.2. The test circuit shown in Fig. 5.1 was programmed using the IBM Qiskit toolkit [111] and compiled and executed remotely on 8 April, 2021.

To estimate the noise parameters, we repeated our experiments L times. Let l denote the index of the l-th experiment. For any instance of circuit execution, the device prepared an ensemble of N_s identical circuits, where N_s denotes the number of shots and k denotes the k-th shot in a particular experiment. In the tests reported below, $L = 203$ was the number of repetitions successfully executed during a 30-min reservation-window, and $N_s = 8,192$ was the number of shots, the maximum allowed by the device. We separately analyzed the results for the case $n = 1$ using each of the 27 register elements available.

In this section, we use the convention that a caret sign denotes a particular realization of a random variable. We first characterized SPAM fidelity, in which SPAM(0) denotes an experiment with a register element, prepared as $|0\rangle$ and measured. Similarly, SPAM(1) denotes an experiment in which a register element, prepared as $|1\rangle$ is measured.

Let $b_{l,k,q}^{SPAM(0)}$ denote the binary outcome of measuring in the computational basis, when collecting the k-th shot of the l-th experiment of the SPAM(0) circuit on the q-th register element. Additionally, let $f_1^q(l)$ denote the initialization fidelity observed in the l-th SPAM(1) experiment for the q-th register element. Similarly, let $f_0^q(l)$ denote the initialization fidelity observed in the l-th SPAM(0) experiment for the q-th register element. Thus:

$$f_1^q(l) = \frac{\sum_{k=1}^{N_s} b_{l,k,q}^{SPAM(1)}}{N_s}, \quad \hat{f}_1^q(l) = \frac{\sum_{k=1}^{N_s} \hat{b}_{l,k,q}^{SPAM(1)}}{N_s} \tag{5.22}$$

$$f_0^q(l) = 1 - \frac{\sum_{k=1}^{N_s} b_{l,k,q}^{SPAM(0)}}{N_s}, \quad \hat{f}_0^q(l) = 1 - \frac{\sum_{k=1}^{N_s} \hat{b}_{l,k,q}^{SPAM(0)}}{N_s}. \tag{5.23}$$

Let $\varepsilon_{q,l}^{\text{SPAM}}$ denote the realized fidelity asymmetry of the q-th register element in the l-th experiment. Thus:

$$\varepsilon_{q,l}^{\text{SPAM}} = f_0^q(l) - f_1^q(l). \tag{5.24}$$

Let $\bar{\varepsilon}_q^{\text{SPAM}}$ denote the mean of the fidelity asymmetry for the q-th register element over the L experiments, and let $\hat{\bar{\varepsilon}}_q^{\text{SPAM}}$ be the corresponding observed value. Thus:

$$\bar{\varepsilon}_q^{\text{SPAM}} = \frac{\sum_{l=1}^{L} \varepsilon_{q,l}^{\text{SPAM}}}{L}, \quad \hat{\bar{\varepsilon}}_q^{\text{SPAM}} = \frac{\sum_{l=1}^{L} \hat{\varepsilon}_{q,l}^{\text{SPAM}}}{L}. \tag{5.25}$$

To quantify the error on these measurements, we define $\sigma(\bar{\varepsilon}_q^{\text{SPAM}})$ as the standard deviation of population mean $\bar{\varepsilon}_q^{\text{SPAM}}$, such that:

$$\hat{\sigma}^2(\bar{\varepsilon}_q^{\text{SPAM}}) = \frac{\hat{\sigma}^2(\varepsilon_q^{\text{SPAM}})}{L} = \frac{1}{L(L-1)} \sum_{l=1}^{L} \left(\hat{\varepsilon}_{q,l}^{\text{SPAM}} - \hat{\bar{\varepsilon}}_{q,l}^{\text{SPAM}}\right)^2. \tag{5.26}$$

The average SPAM fidelity f^q for each qubit q is then calculated using Eqn. 5.13. The initialization fidelities of the computational states are not the same. The asymmetric nature of the single qubit noise channel is brought out starkly by the negligible overlap between the distributions of f_0 and f_1 for qubit 5. Additionally, observe the significant spread in values in qubit 3, relative to the others. These results show that the naive approach of assuming a single value for SPAM error for a qubit is fallacious. Not only do we have to characterize f_0 and f_1 separately, our work must also take into account the significant dispersion around the mean.

The register-wise variation of the SPAM asymmetry is depicted in Fig. 5.2. The plot illustrates the SPAM asymmetry $(f_0 - f_1)$ for ibm_toronto, revealing substantial spatial non-stationarity across the register. The y-axis arranges individual qubits in ascending order of SPAM asymmetry magnitude, while the x-axis represents the mean SPAM fidelity asymmetry expressed as a percentage. Evidently, qubit 0 demonstrates the best performance in this regard, while qubit 24 exhibits the least favorable outcome. Given that SPAM errors are a predominant source of quantum computer errors, this variation is particularly concerning for mitigation procedures[59], as precise knowledge of SPAM noise parameters f_0 and f_1 is essential. The figure features 27 qubits (register elements), with the dashed red line indicating the mean SPAM asymmetry value (averaged across all qubits). The error bars represent the standard deviation of population means across $L = 203$ experiments. Consistency in the register's color scheme is maintained across Figs. 5.2, 5.3, 5.4, and 5.5.

The probability $\text{Pr}_q(0)$, for each qubit, as defined by Eqn. 5.15, was estimated by executing the circuit $\mathcal{C}$ in Fig. 5.1 and counting the fraction of zeros in the 8192-bit long binary string, returned by the remote server. Let $b_{l,k,q}^{\mathcal{C}}$ denote the random measurement outcome (a classical bit) when we conduct an experiment and measure the q-th register element in the l-th experiment's k-th shot (measurement done in the computational Z-basis). Let $\hat{b}_{l,k,q}^{\mathcal{C}}$ denote the corresponding observed value.

Similarly, let $\Pr_{l,q}(1)$ denote the probability of observing 1 as the outcome in the l-th experiment for the q-th register element. Similarly, let $\Pr_{l,q}(0)$ denote the probability of observing 0 as the outcome in the l-th experiment for the q-th register element. Thus:

$$\Pr_{l,q}(1) = \frac{\sum_{k=1}^{N_s} b^{\mathcal{C}}_{l,k,q}}{N_s}, \quad \hat{\Pr}_{l,q}(1) = \frac{\sum_{k=1}^{N_s} \hat{b}^{\mathcal{C}}_{l,k,q}}{N_s} \tag{5.27}$$

and

$$\Pr_{l,q}(0) = 1 - \Pr_{l,q}(1), \quad \hat{\Pr}_{l,q}(0) = 1 - \hat{\Pr}_{l,q}(1). \tag{5.28}$$

Let $H_{l,q}$ denote the Hellinger distance[91] between the noisy and noiseless outcomes in the l-th experiment for the q-th register element. Let $\hat{H}_{l,q}$ be the corresponding observed value (or realization). Let $\bar{H}^q$ denote the mean (a random variable) of the distance for the q-th register element over L experiments. Let $\hat{\bar{H}}^q$ be the corresponding observed realization. Thus:

$$H_{l,q} = \sqrt{1 - \sqrt{\frac{1}{2}\Pr_{l,q}(0)} - \sqrt{\frac{1}{2}\Pr_{l,q}(1)}}, \quad \hat{H}_{l,q} = \sqrt{1 - \sqrt{\frac{1}{2}\hat{\Pr}_{l,q}(0)} - \sqrt{\frac{1}{2}\hat{\Pr}_{l,q}(1)}} \tag{5.29}$$

$$\bar{H}^q = \frac{\sum_{l=1}^{L} H_{l,q}}{L}, \quad \hat{\bar{H}}^q = \frac{\sum_{l=1}^{L} \hat{H}_{l,q}}{L} \tag{5.30}$$

To quantify the error on these measurements, define $\sigma(\bar{H}^q)$ as the standard deviation of population mean $\bar{H}^q$. Thus:

$$\hat{\sigma}^2(\bar{H}^q) = \frac{\hat{\sigma}^2(H^q)}{L} = \frac{1}{L(L-1)} \sum_{l=1}^{L} \left(\hat{H}_{l,q} - \hat{\bar{H}}_q\right)^2. \tag{5.31}$$

Fig. 5.3 pertains to the quantum register of ibm_toronto, comprising 27 register elements. It shows the experimentally derived distance variation across the register. The dashed red line represents the mean distance across the register (averaged over all qubits). Among the qubits, qubit 19 demonstrates the closest proximity to the noiseless state, while qubit 24 exhibits the greatest deviation. The error bars depict the standard deviation of the population mean from 203 experiments. This graph highlights the impact of spatially non-stationary noise on program outcomes. The y-axis arranges qubits in ascending order of the distance between the obtained and noiseless outputs while the x-axis denotes the mean distance, computed as the population mean across $L = 203$ experiments.

The Hadamard gate angle error was subsequently estimated using Eqn. 5.16. The register variation of the the Hadamard gate angle error (in degrees) is shown in Fig. 5.4. Among the qubits, qubit 21 demonstrates the closest proximity to the noiseless, while qubit 24 exhibits the greatest deviation.

This graph serves to highlight spatial non-stationarity in Hadamard gate noise, revealing the inappropriateness of averaging qubit values for coherent noise error mitigation. The y-axis arranges qubits in ascending order of Hadamard gate error, while the x-axis quantifies the Hadamard gate error in degrees.

Fig. 5.5 displays the register-wise variation of the proxy γ_D. Qubit 16 outperforms others, while qubit 24 fares the worst. Fig. 5.6 shows the values for $\gamma_{\max}$ and γ_D for ibm_toronto on 8 April, 2021, when ϵ is set to be the observed distance. The blue dots are experimentally-observed data for each register element (see Table 5.1 for the full list), using the characterization data versus the actual observed distance. It validates our noise model as Eqn. 5.1 holds. The dashed line in Fig. 5.6 provides the decision boundary to test circuit accuracy, using characterization data as a proxy. Given an ϵ-accuracy bound on the statistical distance between the observed distribution and reference to be accurately generated, the plot provides an upper bound for the proxy γ_D (the register variation of γ_D is shown in Fig. 5.5. The latter must lie below this boundary for accuracy by the device. We conjecture that the magnitude of $|\gamma_{\max} - \gamma_D|$ serves as a reliability indicator, i.e., higher value provides greater cushion against temporal fluctuations. Table 5.1 can serve as a basis for register selection.

5.2 Sample bounds on reproducibility

This section aims to establish the minimum sample size (L) required to achieve reproducibility in generated outputs with a specified statistical confidence level ($1 - \delta$), using the reproducibility condition:

$$\Pr(\hat{H} \leq \epsilon) \geq 1 - \delta, \tag{5.32}$$

Consider an n-qubit Bernstein-Vazirani problem (Sec. 3.3) where the secret string is denoted as r. In a noiseless, the probability of obtaining the string r is certain:

$$p_r^{\text{noiseless}} = 1. \tag{5.33}$$

When a circuit is executed once, it yields a single classical bit string v. We introduce the indicator variable Y_r, assigned the value 1 if v matches r, and 0 otherwise.

Let $Y_r(l)$ indicate the outcome of Y_r from the l-th execution. Upon conducting the L runs of the circuit, the experimental sample estimate for the success probability:

$$\hat{p}_r = \frac{\sum_{l=0}^{L-1} Y_r(l)}{L}, \tag{5.34}$$

The mean of $\hat{p}_r$ equals p_r, while its variance is $\frac{p_r(1-p_r)}{L}$. The experimental distance obtained from the noiseless histogram is:

$$\hat{H} = 1 - \sqrt{\hat{p_r}} \tag{5.35}$$

Then, the δ-reproducibility condition translates to:

$$\begin{aligned} &\Pr(\hat{H} \leq \epsilon) \geq 1 - \delta \\ \Rightarrow &\Pr\left[\hat{p_r} \geq (1-\epsilon)^2\right] \geq 1 - \delta \end{aligned} \tag{5.36}$$

This can be reformulated as:

$$\Pr\left[\frac{\hat{p_r} - p_r}{\sqrt{\frac{p_r(1-p_r)}{L}}} \geq \frac{(1-\epsilon)^2 - p_r}{\sqrt{\frac{p_r(1-p_r)}{L}}}\right] \geq 1 - \delta \tag{5.37}$$

Defining the standard normal variable as z with $z \sim \mathcal{N}(0,1)$, the central limit theorem establishes that $\frac{\hat{p_r} - p_r}{\sqrt{\frac{p_r(1-p_r)}{L}}}$ follows the standard normal distribution. Given that $\Pr(z \geq z_\delta) = 1 - \delta$, where z_δ corresponds to a constant dependent on δ for the one-sided confidence interval, we can set:

$$\frac{\hat{p_r} - p_r}{\sqrt{\frac{p_r(1-p_r)}{L}}} = z_\delta \tag{5.38}$$

This satisfies Eqn.5.37. Solving for L yields the minimum bound as:

$$L_{\min} = z_\delta^2 \frac{p_r^{-2} - 1}{p_r^{-2}(1-\epsilon)^2 - 1} \tag{5.39}$$

Hence, the minimum sample size exhibits an inverse relation with the accuracy threshold ϵ and a direct, non-linear correlation with the confidence level $1 - \delta$.

5.3 General bounds on stability

The goal of this section is understanding the extent to which non-stationary noise[42, 36, 46] affects the stability of outcomes generated from a noisy quantum device.

In this section, $\langle O_{\mathrm{x}} \rangle$ denotes the mean of a quantum observable O in presence of a sample realization of the circuit noise x. It is well know that $f_X(\mathrm{x}; t)$ varies with time. We define the average of $\langle O_{\mathrm{x}} \rangle$ with respect to the $f_X(\mathrm{x}; t)$ as:

$$\langle O \rangle_t = \int \langle O_{\mathrm{x}} \rangle \, f_X(x; t) d\mathrm{x}. \tag{5.40}$$

In the absence of knowledge of the exact realization of x at time t, $\langle O \rangle_t$ is an estimate for the mean of the observable in presence of time-varying noise channels. Let $s(t_1, t_2)$ be the absolute difference in the mean of the observable obtained from the noisy quantum device at times t_1 and t_2:

$$s(t_1, t_2) = |\, \langle O \rangle_{t_1} - \langle O \rangle_{t_2} \,|. \tag{5.41}$$

We will refer to $s(t_1, t_2)$ as the stability of the observable [112]. Now,

$$\begin{aligned} s^2(t_1, t_2) &= \left(\langle O \rangle_{t_1} - \langle O \rangle_{t_2} \right)^2 \\ &= \left(\int \langle O_{\mathrm{x}} \rangle \, f_X(\mathrm{x}; t_1) \mathrm{dx} - \int \langle O_{\mathrm{x}} \rangle \, f_X(\mathrm{x}; t_2) \mathrm{dx} \right)^2 \\ &= \left(\int \langle O_{\mathrm{x}} \rangle \, \{ f_X(\mathrm{x}; t_1) \mathrm{dx} - f_X(\mathrm{x}; t_2) \} \mathrm{dx} \right)^2 \\ &\leq \left(\int |\langle O_{\mathrm{x}} \rangle \, \{ f_X(\mathrm{x}; t_1) \mathrm{dx} - f_X(\mathrm{x}; t_2) \}| \, \mathrm{dx} \right)^2. \end{aligned} \tag{5.42}$$

In the last step, the inequality stems from the absolute value on the integrand. Now, per Hölder's inequality, if $m, n \in [1, \infty)$ and $1/m + 1/n = 1$, then:

$$\int |f(x) g(x)| \, dx \leq \left(\int |f(x)|^m dx \right)^{1/m} \left(\int |g(x)|^n dx \right)^{1/n}. \tag{5.43}$$

Thus, our inequality becomes:

$$\begin{aligned} &\left(\int |\langle O_{\mathrm{x}} \rangle \, \{ f_X(\mathrm{x}; t_1) \mathrm{dx} - f_X(\mathrm{x}; t_2) \}| \, \mathrm{dx} \right)^2 \\ &\leq \left[\left(\int |\, \langle O_{\mathrm{x}} \rangle \,|^m \mathrm{dx} \right)^{1/m} \left(\int |f_X(\mathrm{x}; t_1) - f_X(\mathrm{x}; t_2)|^n \mathrm{dx} \right)^{1/n} \right]^2. \end{aligned} \tag{5.44}$$

Now, let $m \to \infty, n = 1$ and define:

$$c = \sup_{\mathrm{x}} |\, \langle O_{\mathrm{x}} \rangle \,|. \tag{5.45}$$

Clearly,

$$\lim_{m \to \infty} \left(\int |\, \langle O_{\mathrm{x}} \rangle \,|^m \mathrm{dx} \right)^{1/m} \leq \lim_{m \to \infty} \left(\int c^m \mathrm{dx} \right)^{1/m} = c \left(\lim_{m \to \infty} \left(\int \mathrm{dx} \right)^{1/m} \right) = c. \tag{5.46}$$

Thus, we have

$$
\begin{aligned}
s(t_1,t_2)^2 &\leq \lim_{m\to\infty,n=1}\left(\left(\int |\langle O_{\mathrm{x}}\rangle|^m \mathrm{dx}\right)^{1/m}\left(\int |\{f_X(\mathrm{x};t_1)\mathrm{dx} - f_X(\mathrm{x};t_2)\}|^n \mathrm{dx}\right)^{1/n}\right)^2 \\
&= c^2\left(\int \left|\sqrt{f_X(\mathrm{x};t_1)} - \sqrt{f_X(\mathrm{x};t_2)}\right|\left(\sqrt{f_X(\mathrm{x};t_1)} + \sqrt{f_X(\mathrm{x};t_2)}\right)\mathrm{dx}\right)^2 \\
&= c^2\int\left(\sqrt{f_X(\mathrm{x};t_1)} - \sqrt{f_X(\mathrm{x};t_2)}\right)^2\mathrm{dx}\int\left(\sqrt{f_X(\mathrm{x};t_1)} + \sqrt{f_X(\mathrm{x};t_2)}\right)^2\mathrm{dx} \\
&\text{(applying Hölder's inequality with m=n=2)} \\
&= c^2\int\left(f_X(\mathrm{x};t_1) + f_X(\mathrm{x};t_2) - 2\sqrt{f_X(\mathrm{x};t_1)}\sqrt{f_X(\mathrm{x};t_1)}\right)\mathrm{dx} \\
&\int\left(f_X(\mathrm{x};t_1) + f_X(\mathrm{x};t_2) + 2\sqrt{f_X(\mathrm{x};t_1)}\sqrt{f_X(\mathrm{x};t_1)}\right)\mathrm{dx} \\
&= c^2\left(1+1-2\int\sqrt{f_X(\mathrm{x};t_1)}\sqrt{f_X(\mathrm{x};t_1)}\mathrm{dx}\right)\left(1+1+2\int\sqrt{f_X(\mathrm{x};t_1)}\sqrt{f_X(\mathrm{x};t_1)}\mathrm{dx}\right) \\
&= 4c^2H_X^2(2-H_X^2),
\end{aligned}
\tag{5.47}
$$

where, for clarity, we use:

$$
H_X^2 = H_X^2(t_1,t_2) = 1 - \int\sqrt{f_X(\mathrm{x};t_1)}\sqrt{f_X(\mathrm{x};t_2)}dx \tag{5.48}
$$

Thus the observable stability s is always upper bounded by

$$
s_{\max}^2 = 4c^2H_X^2(2-H_X^2), \tag{5.49}
$$

an upper bound determined by the degree of time-variation of the device parameters. Thus,

$$
\frac{s}{s_{\max}} \leq 1. \tag{5.50}
$$

The upper bound on the observable stability can also be expressed as:

$$
s_{\max} = 2c\sqrt{1-(1-H^2_{\mathrm{normalized}})^{2d}}. \tag{5.51}
$$

using Eq. 3.10.

5.3.1 Validation using synthetic data

Our synthetic example evaluates the bound on a noisy Bernstein-Vazirani [109] (Sec. 3.3). We are interested in the probability of success to compute the secret bit string r where

$$|r\rangle = \bigotimes_{q=0}^{n-1} |r_q\rangle \tag{5.52}$$

with $r_q \in \{0, 1\}$. The observable for the problem is:

$$O = \Pi_r = |r\rangle \langle r| . \tag{5.53}$$

The state for the noiseless, noiseless circuit is $\rho_{\text{out}}^{\text{noiseless}} = |r\rangle \langle r|$ and, hence, the corresponding probability of success for the noisy circuit describe in Fig. 5.8 is

$$\Pr(\mathrm{r}) = \mathrm{Tr}\left[\Pi_r \mathcal{E}_{\mathrm{x}}(|r\rangle \langle r|)\right] . \tag{5.54}$$

This synthetic simulation, our noise model assumes each register element is acted upon by depolarizing noise, such that the super-operator $\mathcal{E}_{\mathrm{x}}(\cdot)$ represents the tensor product of independent single-qubit depolarizing channels. The i-th qubit is acted upon by the de-polarizing noise channel:

$$\mathcal{E}_{\mathrm{x}_i}(\rho) = \left(1 - \frac{3\mathrm{x}_i}{4}\right)\rho + \frac{\mathrm{x}_i}{4}(\mathbb{X}_i \rho \mathbb{X}_i + \mathbb{Y}_i \rho \mathbb{Y}_i + \mathbb{Z}_i \rho \mathbb{Z}_i) \tag{5.55}$$

where x_i denotes the depolarizing parameter for the i-th qubit's noise channel and $\mathbb{X}_i$, $\mathbb{Y}_i$, and $\mathbb{Z}_i$ are the Pauli matrices. Further, let x_i be a particular realization of the random variable X_i, sampled from the multi-variate joint distribution $f_X(\mathrm{x}; t)$ which has d random variables characterizing the noise in circuit C. We will further assume that the $\{X_i\}$ can have correlations in their values[32]. The univariate marginal distribution for the random variable X_i is denoted by $f_{X_i}(\mathrm{x}_i; t)$ where $i \in (1, \cdots, d)$. In this specific example, $d = n$.

Assuming the noise channel is separable but correlated:

$$\begin{aligned}
\langle O_{\mathrm{x}}\rangle &= \mathrm{Tr}\left[O\mathcal{E}_{\mathrm{x}}\left(\rho^{\mathrm{out}}_{\mathrm{noiseless}}\right)\right] \\
&= \mathrm{Tr}\left[O\mathcal{E}_{\mathrm{x}}\left(|r\rangle\langle r|\right)\right] \text{ (for Bernstein-Vazirani)} \\
&= \mathrm{Tr}\left[|r\rangle\langle r|\,\mathcal{E}_{\mathrm{x}}\left(|r\rangle\langle r|\right)\right] \\
&= \mathrm{Tr}\left[|r\rangle\langle r|\,\mathcal{E}_{\mathrm{x}}\left(|r_1\rangle\langle r_1|\otimes\cdots|r_n\rangle\langle r_n|\right)\right] \\
&= \prod_{i=1}^{n}\mathrm{Tr}\left[|r_i\rangle\langle r_i|\,\mathcal{E}_{\mathrm{x}_i}(|r_i\rangle\langle r_i|)\right] \\
&= \prod_{i=1}^{n}\mathrm{Tr}\left[|r_i\rangle\langle r_i|\left[(1-x_i)\,|r_i\rangle\langle r_i| + x_i\frac{\mathbb{I}}{2}\right]\right] \\
\langle O_{\mathrm{x}}\rangle &= \prod_{i=1}^{d}\left(1-\frac{\mathrm{x}_i}{2}\right).
\end{aligned} \tag{5.56}$$

As a specific instance of a time-varying depolarizing channel, suppose the noise marginals stay constant in the mean while the variance increases linearly with time:

$$\begin{aligned}
\mathbb{E}(\mathrm{x}_i) =& \mu_0 \quad \forall i,t \\
\mathrm{Var}(\mathrm{x}_i) = \sigma_t^2 =& \sigma_0^2\left(1+(\omega-1)\frac{t}{T}\right) \quad \forall i,
\end{aligned} \tag{5.57}$$

where $\omega = \sigma_T^2/\sigma_0^2$ is a constant capturing how volatile the distribution becomes at time T compared to initial time t=0 and i denotes the register number. Classical correlation in the noise is modeled by the correlation matrix Σ where Σ_{ij} represents the correlation coefficient between the depolarizing parameter X_i acting on register element i and X_j acting on register element j.

We use a beta distribution to represent the marginal distribution of the depolarizing parameter x_i as

$$f_{X_i}(\mathrm{x}_i;t) = \frac{\mathrm{x}_i^{\alpha_t-1}(1-\mathrm{x}_i)^{\beta_t-1}}{\mathrm{Beta}(\alpha_t,\beta_t)}, 0\le \mathrm{x}_i \le 1, \tag{5.58}$$

with time-varying parameters α_t and β_t:

$$\alpha_t = \frac{\alpha_0}{k_0+t}, \quad \beta_t = \frac{\beta_0}{k_0+t}, \tag{5.59}$$

and the Beta function, by definition:

$$\mathrm{Beta}(\alpha_t,\beta_t) = \int_0^1 y^{\alpha_t-1}(1-y)^{\beta_t-1}dy, \quad y\in[0,1]. \tag{5.60}$$

We will show later how to estimate the constants α_0, β_0, k_0 from observed data. This choice of model is appropriate if the parameter value ranges between 0 and 1, and the observed data follows a skewed bell-shaped distribution. For simplicity, we will assume the distribution parameters do not vary with register location and the constants α_0, β_0 and k_0 can be estimated from the model requirements in Eqn. 5.57 as:

$$\begin{aligned} k_0 =& T\left(\omega\left(1+(\omega-1)\left(1-\frac{\mu_0(1-\mu_0)}{\sigma_0^2}\right)^{-1}\right)^{-1}-1\right)^{-1} \\ \alpha_0 =& \frac{\mu_0(\mu_0-\mu_0^2-\sigma_0^2)}{\sigma_0^2}k_0 \\ \beta_0 =& \frac{(1-\mu_0)(\mu_0-\mu_0^2-\sigma_0^2)}{\sigma_0^2}k_0. \end{aligned} \tag{5.61}$$

It is verified by substitution that this model satisfies the requirements of Eqn. 5.57. The higher moments of the depolarizing parameter under the beta distribution given by:

$$\mathbb{E}(\mathrm{x}_i^k) = \frac{\mathrm{Beta}(\alpha_t+k,\beta_t)}{\mathrm{Beta}(\alpha_t,\beta_t)} = \prod_{n=0}^{k-1}\frac{\alpha_t+n}{\alpha_t+\beta_t+n} \tag{5.62}$$

We next construct a joint distribution for the d-dimensional distribution using a copula structure, a direct application of Sklar's theorem [129], to model the correlation Σ between the register elements. The use of copulas to study empirical correlation is well-established[130, 124, 131, 132]. Various choices for copulas exist including the Gaussian copula, elliptical copulas, Archimedean copulas, Ali-Mikhail-Haq copula, Clayton copula, Gumbel copula, Independence copula, and Joe copula [133, 134, 79, 135, 136, 137]. They offer different types of modeling capabilities for tail-risk correlations. We chose the Gaussian copula for its simplicity and ease of interpretation.

$$f_X(\mathrm{x};t) = \Theta\left[F_{X_1}(\mathrm{x}_1;t),\cdots F_{X_d}(\mathrm{x}_d;t)\right]\prod_{j=1}^{d} f_{X_j}(\mathrm{x}_j;t), \tag{5.63}$$

where $\Theta(\cdot)$ is the copula function. We use $F_X(\mathrm{x};t)$ as the joint cumulative distribution function for the multi-variate random variable X at time t. Thus,

$$F_X(\mathrm{x};t) = \int_{-\infty}^{\mathrm{x}} f_X(y;t)dy. \tag{5.64}$$

Also, $F_{X_i}(\mathrm{x}_i;t)$ is the cumulative distribution function for the univariate random variable X_i at time t. The Gaussian copula is simply the standard multi-variate normal distribution with correlation matrix Σ:

$$\Theta(y) = \Theta(y_1,\cdots,y_d) = \frac{\exp\left(-\frac{1}{2}(y-\mu_y)^T\Sigma^{-1}(y-\mu_y)\right)}{(2\pi)^{n/2}|\Sigma|^{1/2}}, \tag{5.65}$$

where the vector μ_y is the mean of y.

Having specified the statistics of the time-evolution of the depolarizing noise, we now turn to the task of estimating the distance of the distribution at time t relative to a distribution at time 0. The univariate case has an analytical solution:

$$H_{\mathrm{x}_i}(0,t) = \sqrt{1 - \frac{\mathrm{Beta}\left(\frac{\alpha_0}{2}\left(\frac{1}{k_0} + \frac{1}{k_0+t}\right), \frac{\beta_0}{2}\left(\frac{1}{k_0} + \frac{1}{k_0+t}\right)\right)}{\sqrt{\mathrm{Beta}(\frac{\alpha_0}{k_0}, \frac{\beta_0}{k_0})}\sqrt{\mathrm{Beta}(\frac{\alpha_0}{k_0+t}, \frac{\beta_0}{k_0+t})}}}, \tag{5.66}$$

while the general multi-variate correlated case is analytically intractable. However, the distance can also be computed using Monte Carlo methods. Let $H_X(t_1, t_2)$ be the distance between the d-dimensional multi-variate correlated distributions. Drawing N samples from the distribution $f_X(\mathrm{x}; t_1)$ yields $\{\mathrm{x}^j\}_{j=1}^N$ and, assuming N is large enough to ensure convergence, we numerically approximate the integral as:

$$\frac{1}{N}\sum_{j=1}^{N}\sqrt{\frac{f_X(\mathrm{x}^j_{1,\cdots,d}; t_2)}{f_X(\mathrm{x}^j_{1,\cdots,d}; t_1)}} \approx \mathbb{E}\left(\sqrt{\frac{f_X(\mathrm{x}; t_2)}{f_X(\mathrm{x}; t_1)}}\right) = \int \sqrt{f_X(\mathrm{x}; t_1) f_X(\mathrm{x}; t_2)} d\mathrm{x} = 1 - H_X^2. \tag{5.67}$$

We now demonstrate the validity of Eqn. 5.72 using simulations of the noisy quantum circuit under the correlated depolarizing channel, for which the constant

$$c = \sup_{\mathrm{x}} |\langle O_{\mathrm{x}} \rangle| = \sup_{(\mathrm{x}_1, \cdots, \mathrm{x}_n)} \prod_{i=1}^{d}\left(1 - \frac{\mathrm{x}_i}{2}\right) = 1, \tag{5.68}$$

is maximal in the absence of noise and the noisy, time-dependent observable is modeled as

$$\langle O \rangle_t = \int \langle O_{\mathrm{x}} \rangle \Theta\left[F_1(\mathrm{x}_1; t), \cdots F_d(\mathrm{x}_d; t)\right] f_1(\mathrm{x}_1; t) \cdots f_d(\mathrm{x}_d; t) d\mathrm{x}_1 \cdots d\mathrm{x}_d. \tag{5.69}$$

We estimate this observable through Monte Carlo sample of numerical simulations of the noisy quantum circuit. Our correlated depolarizing noise model assumes the variance of the univariate noise distribution increases linearly each month while the correlation between the isotropic single-qubit depolarizing coefficients is fixed as $\Sigma_{i,j} = 0.80$ for $i \neq j$.

Fig. 5.9 plots the ratio of the simulated stability $s(t)$ to the upper bound $s_{\max}$ with respect to the simulated month for the cases of 4, 8, and 12-bit secret-strings. The results confirm the analytical upper bound.

5.3.2 Validation using device data

We now verify the analytical upper bound using data from the ibm_washington device. The register elements 0, 1, 2, 3, and 4 in the algorithm are mapped to the physical qubits 4, 3, 2, 1, and 0, respectively, as shown in Fig. 5.12. The CNOT gates used in the circuit connect the physical qubits $(0, 1)$ and $(2, 1)$. The data spanned from 1-Jan-2022 to 30-Apr-2023.

Fig. 4.12 shows the correlation between the 16 device parameters taken from Table 5.2. Axes index the corresponding parameters. The data correspond to daily observations made in Apr-2023. The figure presents the Pearson coefficients with blue shades indicative of positive correlation and red shades indicative of negative correlation. Our estimate for the error bars on these coefficients is approximately $1/\sqrt{30-1} \approx 0.18$.

We constructed the joint density using the copulas method [129] discussed in Eqn. 5.63. The full 16-dimensional distribution cannot be visualized but the significance of these correlations is apparent from the example of a bi-variate marginal distribution shown in Fig. 5.10, which compares the constructed probability distribution with and without correlation. Importantly, the correlation structure itself changes monthly with the characterization data.

The full 16-dimensional problem requires a high Monte Carlo sampling overhead for convergence as per Eqn. 5.67. To address this issue, we determined that our machine's configuration allows for a Monte Carlo sampling size of 100,000, which corresponds to a program runtime of approximately six hours including IBM Qiskit [23] simulations and Monte Carlo sampling overhead[138]. Introducing a threshold for correlation enables the clustering of variables and reduces the effective problem dimensionality[52].

As the correlations between device parameters varies each month, the number of clusters and their composition also changes. For example, in May 2022, our method identified 13 clusters with the biggest cluster comprising 3 device parameters, while in April 2023, we found 16 independent clusters.

Generally, given d device parameters that form K independent clusters at time t, denote the i-th cluster as $\mathcal{B}_i(t)$. The cardinality of $\mathcal{B}_i(t)$ is denoted by $m_i(t)$, such that $\sum_i m_i(t) = d$ for all t. Let the elements of $\mathcal{B}_i(t)$ be given by $\{\mathrm{x}_{(1,i)}, \cdots, \mathrm{x}_{(m_i(t),i)}\}$ and let $\Theta_i(t)$ from Eqn. 5.63 denote the copula function for cluster $\mathcal{B}_i(t)$, i.e.,

$$\Theta_i(t) = \Theta\left[F_{X_{(1,i)}}\left(\mathrm{x}_{(1,i)}; t\right), \cdots, F_{X_{(m_i(t),i)}}\left(\mathrm{x}_{(m_i(t),i)}; t\right)\right]. \tag{5.70}$$

Then, Eqn. 5.67 becomes:

$$1 - H_X^2 = \mathbb{E}\left(\sqrt{\frac{\prod_{i\in t_1 \text{ clusters}} \Theta_i(t_2)}{\prod_{j\in t_2 \text{ clusters}} \Theta_j(t_1)} \prod_{k=1}^{d} \frac{f_{X_k}(\mathrm{x}_k; t_2)}{f_{X_k}(\mathrm{x}_k; t_1)}}\right), \tag{5.71}$$

which we will approximate through Monte Carlo sampling.

We use 100,000 Markov-Chain Monte Carlo simulations to estimate $\langle O \rangle_t$ for a given month. This sample size was chosen based on numerical convergence by using the Qiskit Aer numerical simulator to calculate noisy simulations of the circuit. From these estimates, we then calculated the monthly average observable value, $\langle O \rangle_t$ and the observable stability, $s = |\langle O \rangle_t - \langle O \rangle_0|$. Moreover, we performed these simulation 100 times for each month to estimate the underlying distribution for the stability itself.

Fig. 5.11 presents the the observable stability s to $\mathrm{s}_{\max}$ ratio from these simulations for each month. In this box-and-whisker plot, the central box at each point signifies the interquartile range (IQR), with its lower and upper edges representing the first (Q1) and third quartiles (Q3), respectively. The median is indicated by a line within the box. Notably, all ratios remain well below unity and verify that the upper bound calculated from characterization data is never surpassed.

From Fig. 5.11, we also see that our upper bound for the temporal variations of the quantum observable is 100 times higher than the experimentally observed values. Although looser bounds are symptomatic of an overestimation of the device noise, that is acceptable because underestimating the noise is not an option for performance improvement roadmap[55, 56, 57] and exact bounds is impossible.

Note that Eqn. 5.72 does not provide a tight bound due to three sources. Firstly, we can make Eqn. 5.42 tighter by restricting ourselves to scenarios where the noise distribution function at a later time is consistently lower than at an earlier time, which often occurs in between calibrations. In fact, the reason that Fig. 5.9 was able to achieve a more accurate estimate of the temporal variations of the quantum observable is because we had modeled an in-between calibrations scenario. Secondly, the use of Hölder's inequality introduces additional loss of tightness, since the equality holds only when the two functions are linearly dependent. Thirdly, Eqn. 5.46 leads to a looser bound for observables that heavily fluctuate with platform characterization metrics. This approximation, found in the appendix, employs the maximum value of $\langle O_{\mathrm{x}} \rangle$ to set the integral's bound. The accuracy of this approximation diminishes as $\langle O_{\mathrm{x}} \rangle$ fluctuates more with x, while it improves with less variation in

x. In our Bernstein-Vazirani application, where $\langle O_{\mathrm{x}} \rangle$ ranges from 0 to 1, this introduces significant approximation, contributing to the observed loose bound.

Despite not being very tight, our bound in Eqn. 5.72 is still useful for several reasons. Firstly, it helps estimate the maximum temporal variations and ensures result reproducibility. Secondly, if the platform noise stays within the bounds determined by the analysis, experimental reproducibility can be guaranteed with a high degree of certainty. Finally, numerical simulations using real data allow us to scale down the requirements to be less restrictive.

5.4 General bounds on reliability

The purpose of this section is to determine the bound on the reliability metric for a noisy quantum device in achieving an ϵ-stable outcome.

Eqn. 5.51 can be re-arranged to yield the upper bound on distance:

$$H_{\max}(t_1, t_2) = \sqrt{1 - \sqrt{1 - \phi}} \tag{5.72}$$

with $\phi = s_{\mathrm{tol}}^2/(4c^2)$.

We validate the bound using a numerical simulation of of the Bernstein-Vazirani circuit like before. To validate this bound, first, we mapped the 16 noise parameters essential for our simulation of the 5-qubit Bernstein-Vazirani circuit shown Fig. 3.4 to specific independent noise processes. The parameters mapped to gate and register specific noise model in Table 5.2. For example, the asymmetric binary channel for register 0 flips the measured output bit b_0 to $b_0 \oplus 1$ with probability x_0, while the coherent phase error channel[58] for the Hadamard gate H applied to register 0 transforms the underlying quantum state as $CP(H\rho H) = R_z(\theta)H\rho H R_z^{\dagger}(\theta)$. Thermal relaxation[139] is modeled by an exponential dephasing process that depends on the T_2 time and the duration of the underlying gate not shown here. While the 16 noise processes above act independently, the underlying noise parameters are assumed to be correlated. We construct a joint distribution of to describe these parameters using the method of Gaussian copula[79].

We generate an ensemble of noisy simulations by drawing samples from the multi-parameter noise distribution. We initially establish a joint distribution from the daily data gathered in January 2022 for the ibm_washington device, utilizing copulas. Over the next 15 months, we introduce minor perturbations to this distribution, ensuring that the distance never exceeds $H_{\max}$ between the perturbed and original January 2022 distributions. In this perturbation scheme, the marginal distribution of the CNOT gate error between qubits 1 and 2 is modeled using a beta distribution,

which is based on the aforementioned January 2022 daily data. Small, random perturbations to the beta distribution parameters are incorporated over 15 months for the CNOT error, with distance constraint maintained. For each perturbed distribution, we generate 100,000 noise metric samples, and execute 100 Qiskit simulations (each with 8192 shots). The stability metric is then computed from the obtained output.

Figure 5.7 presents the simulation results illustrating the relationship between the stability metric (s) and the reliability of a quantum device characterized by the distance (H). The results demonstrate that when $H \leq H_{\max}$ the device is reliable such that the temporal difference of the observable (s) remains within the specified upper bound ($s \leq s_{\max}$). In our simulations, we set the tolerance threshold $s_{\text{tol}} = 20\%$, which limits the maximum acceptable deviation in the expectation value over time. According to Eqn. 5.72, this results in an upper limit of 7.1% for the device reliability metric $H_{\max}$ for the Bernstein-Vazirani circuit. The lower panel presents the distance between the noise processes. These calculations show how noise can fluctuate on a monthly basis while still respecting the $H_{\max}$ constraint. While time varying, these process emulate the behavior of a reliable device. The upper panel of Fig. 5.7 presents the corresponding stability metric, which never exceeds the 20% tolerance. Moreover, we find the stability is nearly two orders of magnitude smaller than the tolerance, with an average of 0.6%. By selecting a reliable device, we can ensure the stability of quantum output.

Table 5.1: Register values for $\gamma_D(\tau)$ and $\gamma_{\max}$

Register No.	$\gamma_{\mathbf{max}}$	$\gamma_D(\tau)$
0	1.4590	1.3040
1	1.1365	0.6755
2	2.7284	2.7118
3	6.9946	6.9931
4	4.3229	4.3226
5	5.8171	5.8157
6	4.5425	4.5325
7	2.6946	2.6649
8	8066	5.4724
9	8.9672	8.9666
10	2.7272	2.7231
11	11.5502	11.5486
12	3.2212	3.0797
13	1.7818	0.6460
14	11.9104	11.9038
15	2.0713	2.0228
16	1.3392	0.2359
17	4.8557	4.8553
18	1.5986	1.4980
19	1.0322	0.4378
20	9.0893	9.0886
21	1.2259	1.0620
22	10.9146	10.9136
23	3.0018	3.0017
24	14.1254	14.1241
25	1.4325	1.2624
26	1.3103	0.9567

Table 5.2: Device parameters

Parameter	*Description*
x_0	SPAM fidelity for register element 0
x_1	SPAM fidelity for register element 1
x_2	SPAM fidelity for register element 2
x_3	SPAM fidelity for register element 3
x_4	CNOT gate fidelity for control 0, target 1
x_5	CNOT gate fidelity for control 2, target 1
x_6	T_2 de-coherence time for register element 0
x_7	T_2 de-coherence time for register element 1
x_8	T_2 de-coherence time for register element 2
x_9	T_2 de-coherence time for register element 3
x_{10}	T_2 de-coherence time for register element 4
x_{11}	Hadamard gate fidelity for register element 0
x_{12}	Hadamard gate fidelity for register element 1
x_{13}	Hadamard gate fidelity for register element 2
x_{14}	Hadamard gate fidelity for register element 3
x_{15}	Hadamard gate fidelity for register element 4

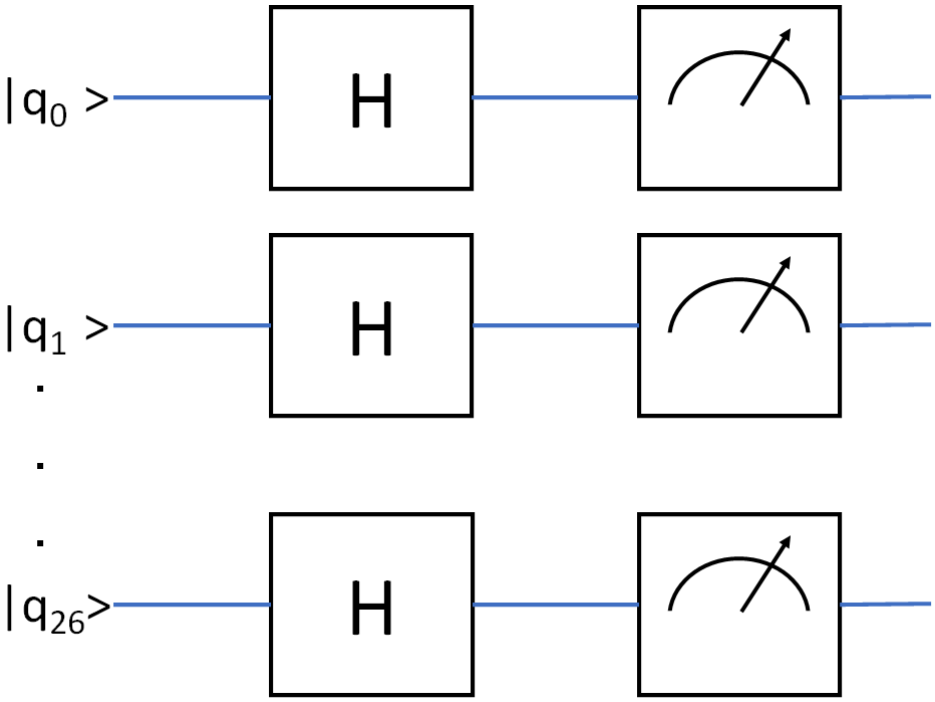


Figure 5.1: Circuit used for our experiment. In this figure, H represents the Hadamard gate. The meter symbol denotes measurement gate.

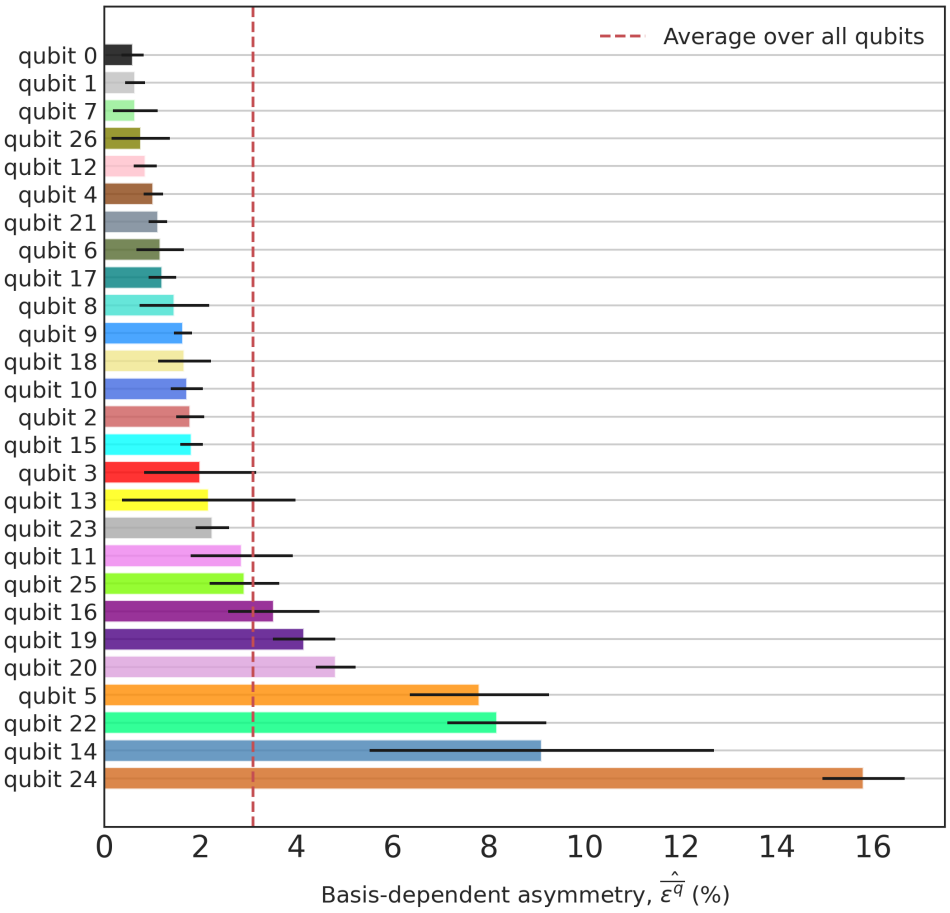


Figure 5.2: Plot illustrating significant spatial non-stationarity in the register-wise variation of the SPAM asymmetry for ibm_toronto.

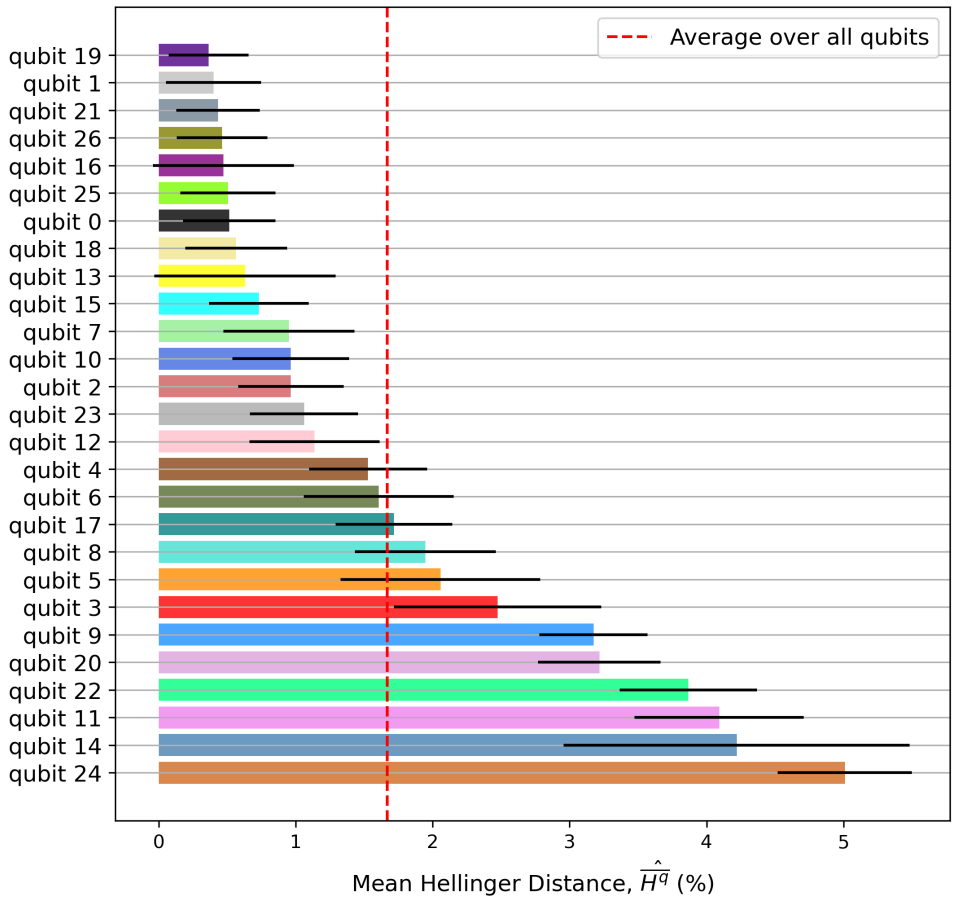


Figure 5.3: The plot depicts distance variation across the 27 register elements of ibm_toronto device, illustrating the impact of spatially non-stationary noise on program outcomes and the dependence of output on register choice.

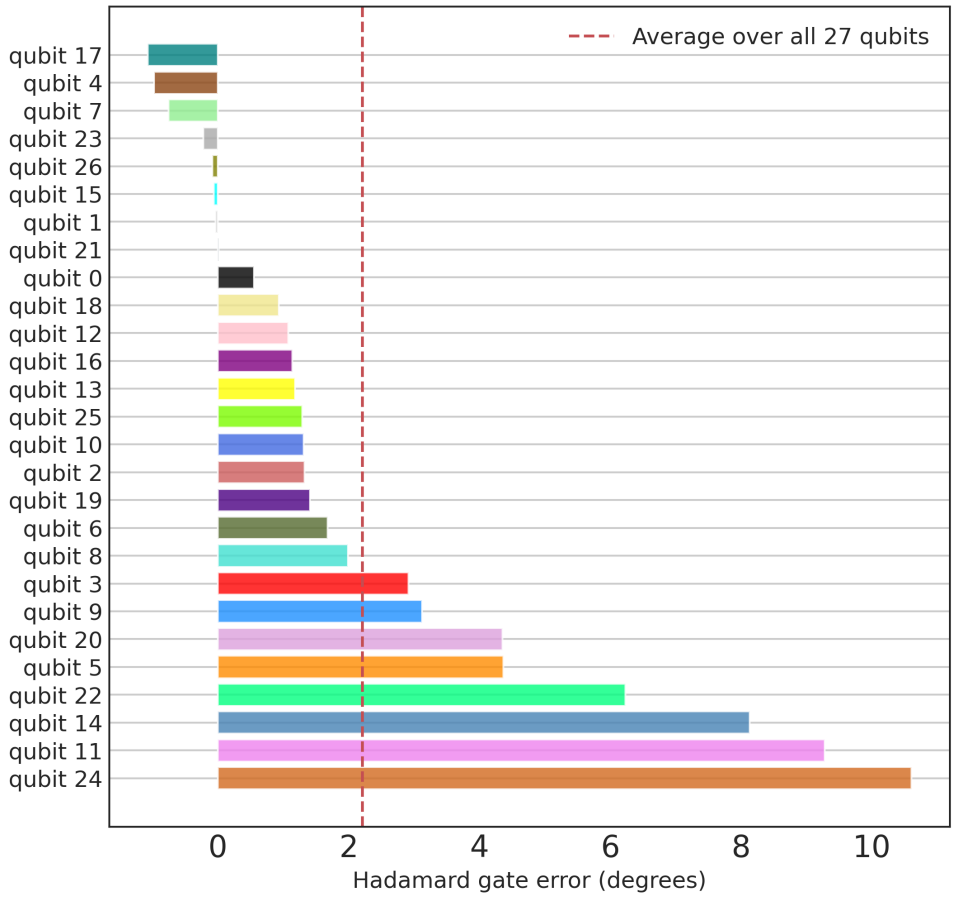


Figure 5.4: The plot illustrates the register-wise Hadamard gate angle error (in degrees) within ibm_toronto's quantum register of 27 elements, emphasizing spatial non-stationarity and cautioning against averaging qubit values for coherent noise mitigation.

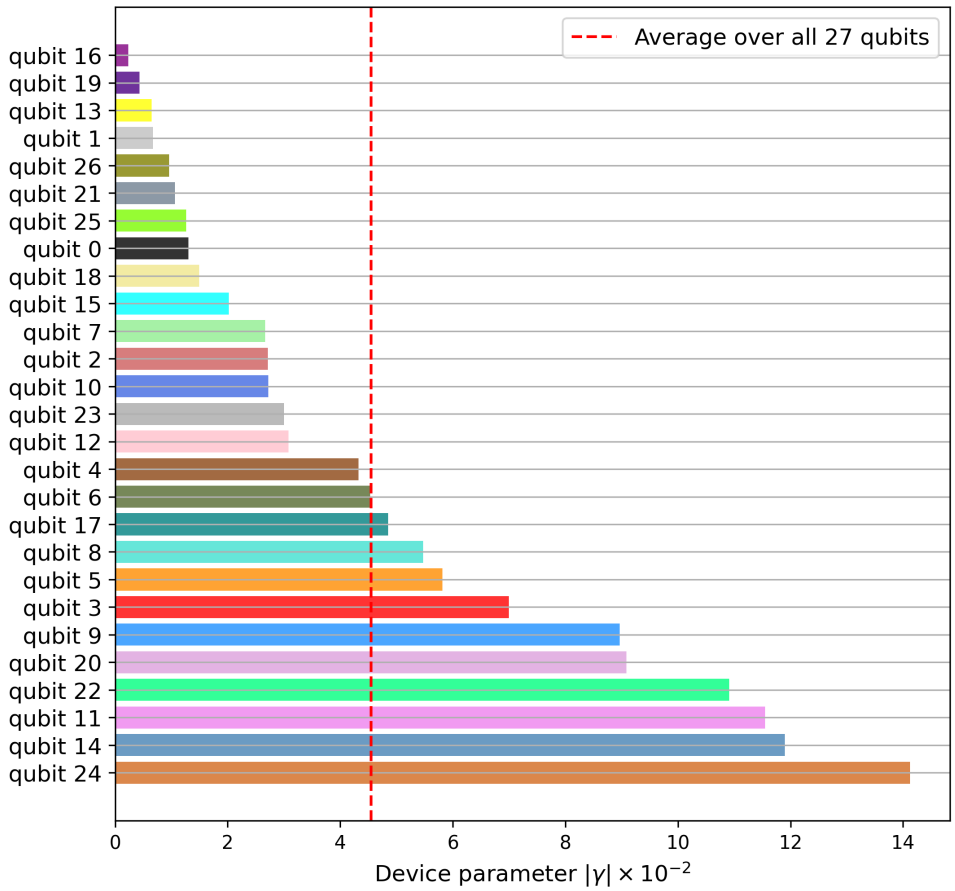


Figure 5.5: This plot shows the register-wise variation of the composite accuracy metric γ_D for ibm_toronto device with 27 qubits, where a higher value significantly impacts program precision. The graph highlights the necessity of re-estimating the metric due to temporal non-stationary noise in unreliable devices, emphasizing the crucial role of analyzing noise parameter interactions for desired accuracy.

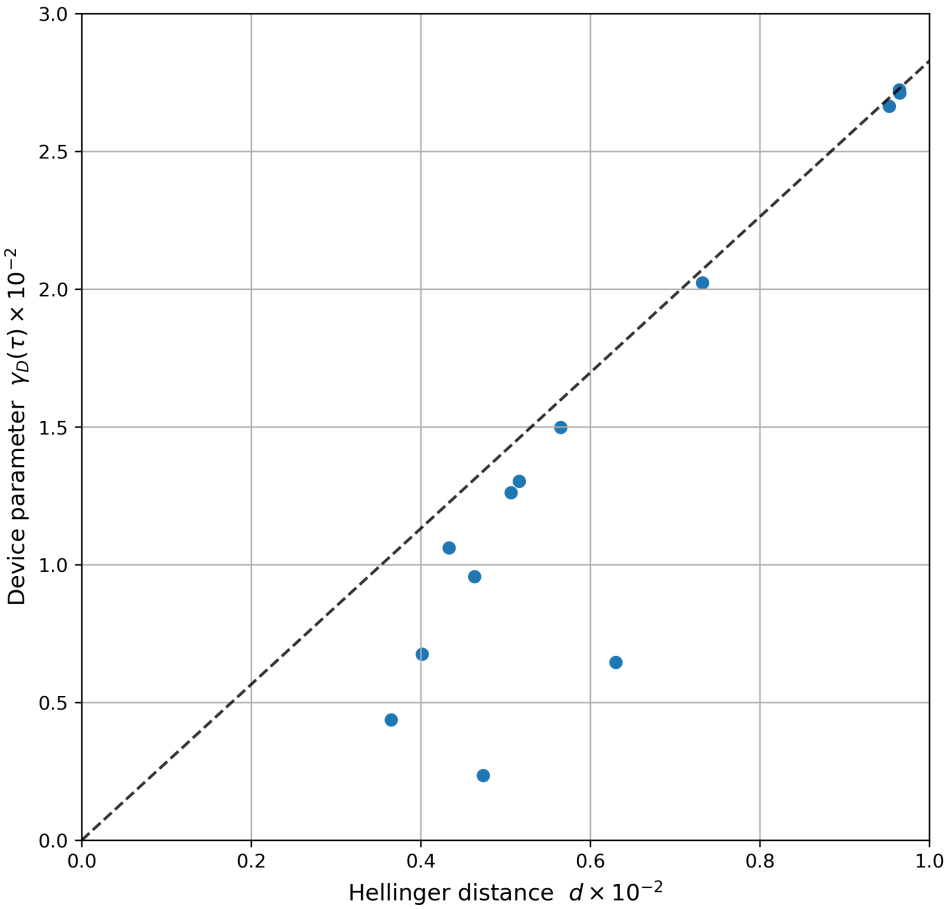


Figure 5.6: Characterizing circuit accuracy on toronto. Plot of $\gamma_{\max}$ (dashed line) and γ_D (blue dots) for ibm_toronto on 8 April 2021. The blue dots are experimentally-observed data plotted using the characterization data versus the actual observed distance (d) for each register element. Only a subset of qubits are shown.

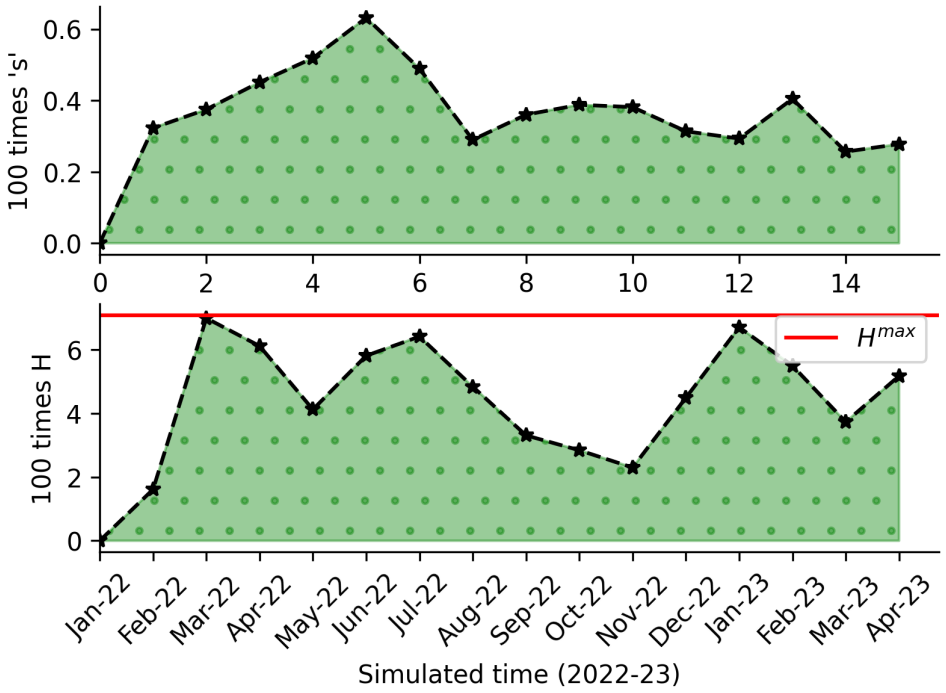


Figure 5.7: Simulation demonstrating that when $H \leq H_{\max}$ (i.e. a reliable, slowly varying noise platform), then $s \leq s_{\max}$ (i.e. the temporal difference of the observable stays within the predicted upper bound).

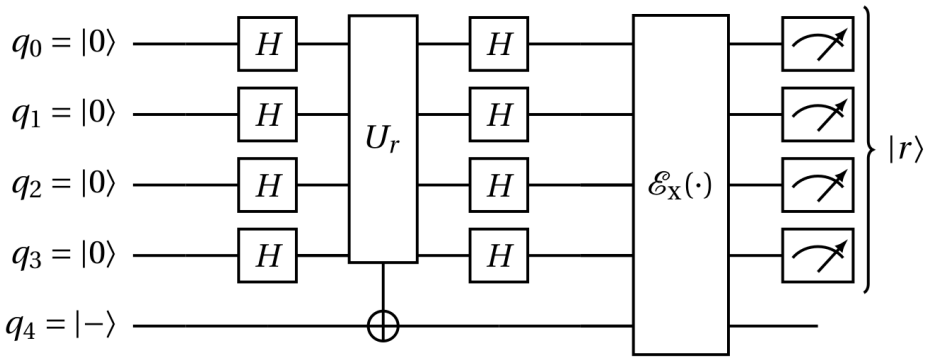


Figure 5.8: A quantum circuit implementation of the Bernstein-Vazirani algorithm that employs 5 qubits, denoted q_0 to q_4. The first four qubits are used to compute the 4-bit secret string, while the fifth qubit serves as an ancilla and initially resides in the $|-\rangle$ superposition state. The symbol H denotes the Hadamard gate while the oracle unitary (U_r) implements the secret string (r). The depolarizing noise channel is denoted by $\mathcal{E}_x(\cdot)$. A quantum measurement operation is represented by the meter box symbol at the circuit's end.

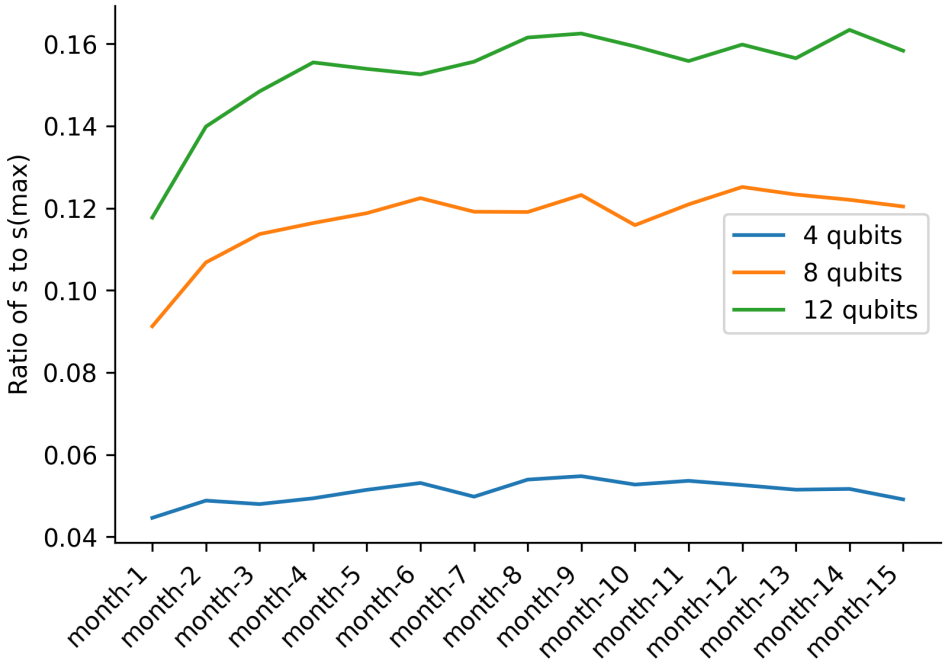
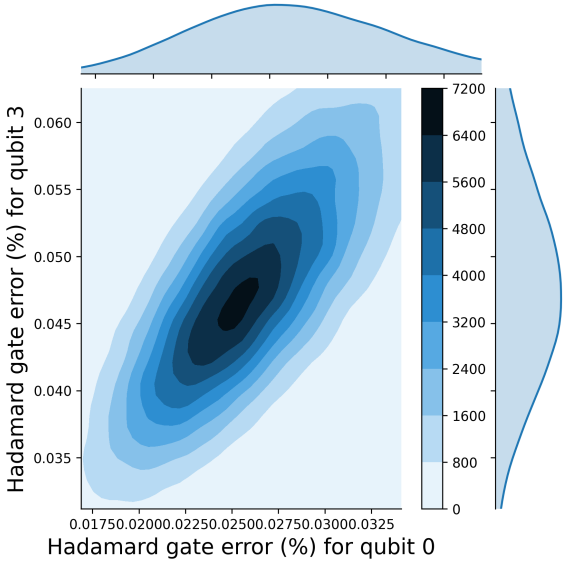
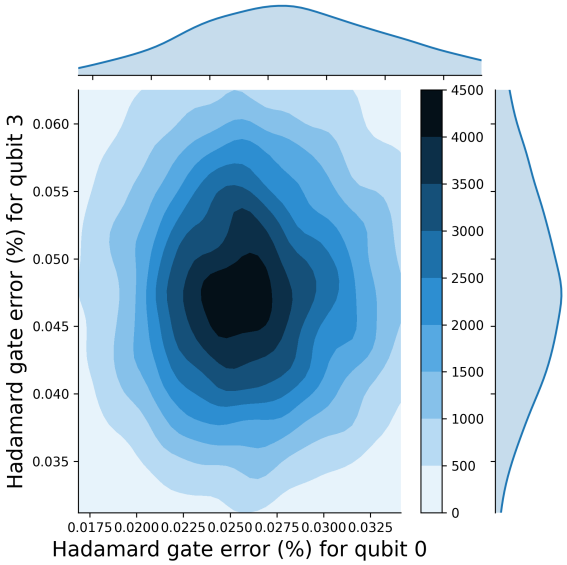


Figure 5.9: The ratio $s/s_{\max}$ for a simulated time-varying noisy circuit is plotted with respect to the increasing noise variance across 15 months. The results from noisy simulations of the Bernstein-Vazirani circuit with a secret string of 4, 8, or 12 bits validate the bound expected.



(a)



(b)

Figure 5.10: Contour plots to compare the probability densities of a two-dimensional subset of Hadamard gate errors for qubit 0 and 3 in Apr-2023, (a) with and (b) without correlation modeling using a copula function.

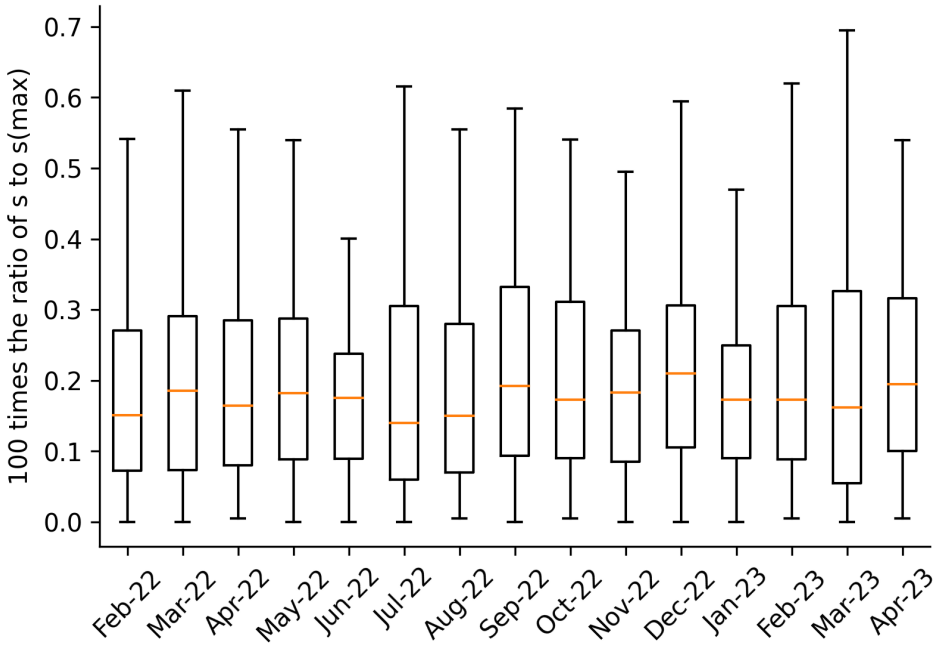


Figure 5.11: Simulations of the stability ratio s/s_{max} times 100 for a 4-qubit Bernstein-Vazirani circuit using the noise characterization from the ibm_washington platform. The box-and-whisker plot of the monthly statistics are based on noisy circuit simulations using the joint probability distribution derived from data from 1-Jan-2022 to 30-Apr-2023. Ratio values below unity confirm that the upper bound is never exceeded.

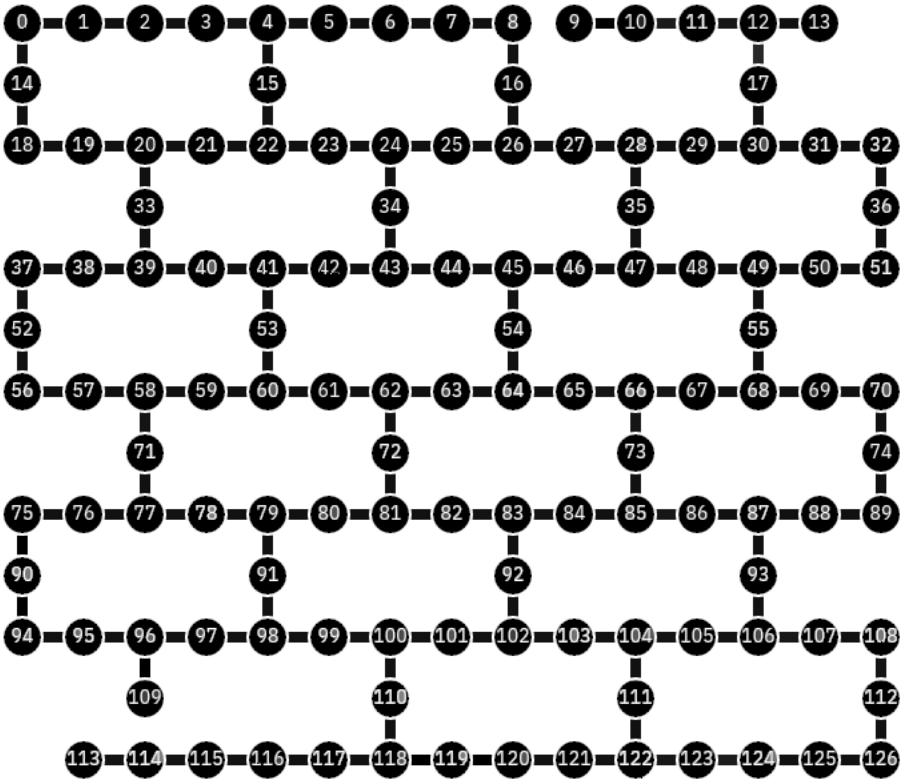


Figure 5.12: Schematic layout of the 127-qubit washington device produced by IBM. Circles denote register elements and edges denote connectivity of 2-qubit operations. The register elements 0, 1, 2, 3, and 4 are mapped to the physical qubits 4, 3, 2, 1, and 0, respectively, in the diagram above. The CNOT gates used in the circuit connect the physical qubits $(0, 1)$ and $(2, 1)$, in the diagram above, where the first number represents the control qubit and the second one represents the target qubit.

Chapter 6

Enhancing histogram accuracy

The previous chapter focused on modeling the stability of outcomes [110] from noisy quantum computers in presence of non-stationary quantum noise [46]. It developed analytical bounds leveraging device characterization data to enable estimation of outcome stability. The bounds were validated using noise characterization data collected from IBM transmon processors [30].

In this chapter, we study Bayesian techniques [73, 74, 75, 76] to improve the accuracy of histograms obtained from a noisy quantum computer using a uniform superposition circuit [140] is utilized as a test case. The performance metric in this chapter is the Hellinger distance (H_{x}) [78] between the noisy histogram in the computational $\mathbb{Z}$-basis observed at time t and the noiseless histogram. The chapter is divided into two sections: the first section focuses on improving histogram accuracy in presence of uncorrelated noise, while the second deals with correlated noise.

6.1 Uncorrelated noise

Suppose we want to execute a noisy quantum circuit L times, indexed by l. In each execution, the number of repetitions allowed by a remote cloud computer [141] (also called n-shots) is N_s and is indexed by k. Let us call the n-bit digital output of the noisy quantum computer as v. This n-bit digital output is measured in the computational basis. To be precise, $v(k, l)$ denotes the output of the k-th shot for the l-th circuit execution. Thus,

$$v(k, l) = [v_{n-1}(k, l) \cdots v_0(k, l)] \tag{6.1}$$

When represented as a decimal integer, $v(k,l)$ takes values in $\{0,\cdots,2^n-1\}$. When represented in binary, each $v_i(k,l)$ denotes a classical bit $\in\{0,1\}$. In Dirac notation, the classical bit v_i can be written as as $|v_i\rangle \in \{|0\rangle, |1\rangle$ and the output can be written as:

$$v(k,l) = |v_{n-1}\rangle(k,l) \otimes \cdots \otimes |v_0\rangle(k,l) \tag{6.2}$$

Note that l can also be thought of as a proxy for a short time-window during which the noise can be assumed to stay constant. During this short time-window, we are able to collect N_s outcomes (each of length n bits). These N_s observations are denoted by $\{v(k,l)\}_{k=0}^{N_s-1}$ and they correspond to the l-th3 circuit execution instantiation.

The noise parameter x characterizes a quantum noise channel $\mathcal{E}_{\mathrm{x}}(\cdot)$. For instance, for a single-qubit depolarizing channel, where $\mathrm{x}\in[0,1]$, the effect of the noise on an input density matrix ρ is represented by:

$$\mathcal{E}_{\mathrm{x}}(\rho) = (1-\mathrm{x})\rho + \frac{\mathrm{x}}{3}(\mathbb{X}\rho\mathbb{X} + \mathbb{Y}\rho\mathbb{Y} + \mathbb{Z}\rho\mathbb{Z}) \tag{6.3}$$

where $\mathbb{X},\mathbb{Y},\mathbb{Z}$ are the Pauli matrices. Also, let the probability density for x at the l-th instant be denoted by $f_X(\mathrm{x};l)$.

Using Bayes' theorem [142]:

$$\Pr[\mathrm{x};l \mid \{v(k,l)\}_{k=0}^{N_s-1}] \propto \Pr[\{v(k,l)\}_{k=0}^{N_s-1} \mid \mathrm{x}] f_X(\mathrm{x};l) \tag{6.4}$$

where $\Pr[\{v(k,l)\}_{k=0}^{N_s-1} \mid \mathrm{x}]$ is the likelihood, $f_X(\mathrm{x};l)$ is the prior, and $\Pr[\mathrm{x};l \mid \{v(k,l)\}_{k=0}^{N_s-1}]$ is the posterior. The prior can be assumed from available old device characterization data. In absence of available old data, it is also okay to assume that the prior is a uniform distribution (to indicate a lack of information about the quantum channel) [143]. Note that the missing proportionality constant $\tilde{c}$ in Eqn. 6.4 is given by:

$$\frac{1}{c} = \int_{\mathrm{x}} f_X(\{v(k,l)\}_{k=0}^{N_s-1} \mid \mathrm{x}) f_X(\mathrm{x};l) d\mathrm{x} \tag{6.5}$$

Only in rare cases, this proportionality constant can be calculated analytically. It is usually computationally intractable. However, in Markov Chain Monte Carlo based methods [144], $\tilde{c}$ is often not required.

The next step in the algorithm is to find the maximum-a-posteriori (MAP) [145] estimate for x:

$$\hat{\mathrm{x}} = \underset{\mathrm{x}}{\operatorname{argmax}}\ f_X(\mathrm{x} \mid \{v(k,l)\}_{k=0}^{N_s-1}) \tag{6.6}$$

For purposes of quantum error mitgation, $\hat{\mathrm{x}}$ is our best guess for x in a non-stationary noise environment (such as depicted in Fig. 6.1 and Fig. 6.2).

An advantage of this method is that it helps to mitigate noise in gate operations from the software interface without having to do pulse-level programming [23]. However, Bayesian methods are notorious for not being rapidly scaleable. The scalability depends on the noise model granularity. Using numerous noise parameters that exponentially increase with register elements might not aid in efficient statistical estimation and can lead to poorer outcomes. Usually, embracing simpler models reduces bias and prevents over-fitting [146].

To illustrate, consider a 4-qubit register initialized to $|0000\rangle$. Each register element is subjected to a Hadamard gate to produce four-qubits in uniform superposition. The noise parameter x for this circuit has 12 elements:

- the SPAM error, $\mathrm{x}_{q,0}$, characterizing the SPAM noise for register element q, when the input state is $|0\rangle$ ($q \in \{0,1,2,3\}$).
- the SPAM error, $\mathrm{x}_{q,1}$, characterizing the SPAM noise for register element q, when the input state is $|1\rangle$ ($q \in \{0,1,2,3\}$).
- the Hadamard gate error, $\mathrm{x}_{q,2}$, characterizing the single-qubit rotation error for register element q ($q \in \{0,1,2,3\}$).

As a side note, in other chapters, we have denoted SPAM fidelity with the letter f. However, in this chapter, we use a slightly different notation for the sake of clarity.

Next, we will specify the Bayesian model [144]. Since this section addresses the case of independent noise sources, it implies no entanglement between register elements. This enables us to use a separable noise channel [16]:

$$\mathcal{E}_{\mathrm{x}}(|v\rangle) = \bigotimes_{i=n-1}^{0} \mathcal{E}_{\mathrm{x}_i}(|v_i\rangle) \tag{6.7}$$

Note that i is in descending order to reflect register endianness [147]. The prior density is:

$$f_X(\mathrm{x}) = \prod_{j=0}^{2} \prod_{i=0}^{n-1} f_X(\mathrm{x}_{i,j}) \tag{6.8}$$

where j iterates through the circuit noise types (SPAM errors, gate error) and i through the qubits.

Each of the independent univariate parameters $\mathrm{x}_{i,j}$ is modeled using a time-varying beta distribution [148]:

$$f(\mathrm{x}_{i,j};t) = \frac{\mathrm{x}_{i,j}^{\alpha_{i,j}(t)-1}(1-\mathrm{x}_{i,j})^{\beta_{i,j}(t)-1}}{\mathrm{Beta}(\alpha_{i,j}(t),\beta_{i,j}(t))} \tag{6.9}$$

where $\mathrm{Beta}(\cdot,\cdot)$ is the Beta function:

$$\mathrm{Beta}\left[\alpha_{i,j}(t),\beta_{i,j}(t)\right] = \int_0^1 t^{\alpha_{i,j}(t)-1}(1-t)^{\beta_{i,j}(t)-1}dt, \quad \forall\alpha_{i,j}(t),\beta_{i,j}(t) > 0 \tag{6.10}$$

The choice of the beta distribution is motivated by the bounded nature of the experimental data (upper bounded at 1 for SPAM noise and $\pi/4$ for hadamard gate noise) and its skewed characteristics, often exhibiting a peak. We remind the reader that the discrete letter l (describing the quantum circuit execution instant) will be used interchangeably with the continuous time t.

The time-varying likelihood function [149] is given by:

$$\mathcal{L} = \Pr[\{v(k,l)\}_{k=0}^{N_s-1} \mid \mathrm{x}] \tag{6.11}$$

$$= \prod_{k=0}^{N_s-1} \Pr(v(k,l) \mid \mathrm{x}) \tag{6.12}$$

$$= \prod_{k=0}^{N_s-1}\prod_{i=0}^{n-1}[\pi_i(l)]^{1-v_i(k,l)}[1-\pi_i(l)]^{v_i(k,l)} \tag{6.13}$$

where,

$$\pi_i(l) = \frac{1-\mathrm{x}_{i,0}(l)+\mathrm{x}_{i,1}(l)}{2} + \sin\left[2\mathrm{x}_{i,2}(l)\right]\frac{1-\mathrm{x}_{i,0}(l)-\mathrm{x}_{i,1}(l)}{2} \tag{6.14}$$

is the probability of observing $|0\rangle$ for qubit i when measured in the computational basis after the l-th execution.

The posterior distribution is estimated (using Metropolis-Hastings algorithm [138]) per:

$$\text{Posterior} \;\propto\; \text{Likelihood} \;\times\; \text{Prior}$$

$$\Rightarrow \Pr[\mathrm{x}; l \mid \{v(k,l)\}_{k=0}^{N_s-1}] \propto \prod_{i=0}^{n-1}\left[\pi_i(l)^{N_s-\sum_k v_i(k,l)}(1-\pi_i(l))^{\sum_k v_i(k,l)}\prod_{j=0}^{2}\frac{\mathrm{x}_{i,j}^{\alpha_{i,j}(l-1)-1}(1-\mathrm{x}_{i,j})^{\beta_{i,j}(l-1)-1}}{\mathrm{Beta}\left[\alpha_{i,j}(l-1),\beta_{i,j}(l-1)\right]}\right] \tag{6.15}$$

Lastly, the maximum-a-posteriori (MAP) estimate is obtained using a log maximization:

$$\hat{\mathrm{x}}(l) = \underset{\mathrm{x}}{\operatorname{argmax}} \ \log \Pr[\mathrm{x}; l \mid \{v(k,l)\}_{k=0}^{N_s-1}] \tag{6.16}$$

$\hat{\mathrm{x}}(l)$ is our best estimate for the time-varying quantum noise when the noise terms are independent. This time-varying noise estimate is then used for quantum error mitigation at time t. Specifically, $\hat{\mathrm{x}}_{i,0}(l)$ and $\hat{\mathrm{x}}_{i,1}(l)$ define the time-varying SPAM noise matrix for qubit i for readout mitigation using matrix inversion [150]. $\hat{\mathrm{x}}_{i,2}(l)$ is our best estimate for Hadamard calibration noise [11]. We mitigate this noise by using $\pi/4 - \hat{\mathrm{x}}_{i,2}$ as the input in the software for the single qubit rotation. This helps us to avoid pulse level programming [23].

We used Qiskit [111] for our simulations. The circuit layout is depicted in Fig. 6.5. For simulating the SPAM noise channel, we employed a binary asymmetric channel for each register element. The SPAM parameters were drawn from a beta distribution. For the initial state $|0\rangle$, the mean of the SPAM fidelity distributions were $(0.9, 0.8, 0.85, 0.75)$ for the respective qubits, with the standard deviation being one-tenth of the mean. Similarly, for the initial state $|1\rangle$, the mean of the SPAM fidelity distributions were $(0.85, 0.75, 0.80, 0.70)$ for each register element, with the standard deviation remaining one-tenth of the mean. For simulating the Hadamard noise, we used Qiskit's $U3$ gate [111] which is parameterized by three angles: θ, ϕ and λ. For our noise simulation, we used $\theta = \frac{\pi}{2} + \mathrm{x}_{i,2}$, $\phi = 0$, and $\lambda = \pi$. Here, $\mathrm{x}_{i,2}$ is stochastic Hadamard noise drawn from a beta distribution with mean (in degrees) given by: $3.1, 4.1, 4.9, 2.9$ for the four qubits respectively. The standard deviation remained one-tenth of the mean, similar to previous cases. The circuit was repeated $L = 10$ times to obtain acceptable error bars. A sample execution is shown in Fig. 6.3. Each execution comprised $N_s = 8192$ shots. Three scenarios were investigated:

- Unmitigated: Raw results with no mitigation.
- Static Mitigation: Traditional method disregarding time-varying noise (uses average numbers).
- Adaptive Mitigation: Using Bayesian optimization for inferring time-varying estimates (as detailed in this section).

The result of our experiment is shown in Fig. 6.4. It clearly demonstrates a decrease in error (quantified by the Hellinger distance) with an adaptive approach. Importantly, it also illustrates that mitigation using average noise parameters can sometimes increase the error in presence of time-varying quantum noise, compared to raw results with no mitigation.

6.2 Correlated noise

In this section, we will explore a correlated multi-qubit noise model. We will use a n-qubit Pauli noise channel. The Pauli oise model, although not a completely general noise model, still manages to model many practical situations. It is widely used because of two reasons: (a) it is efficiently simulatable on a classical computer (per the Gottesman-Knill theorem) and (b) when used as a proxy for physically accurate noise models (such as the amplitude and phase damping noise) which are not efficiently simulatable on a classical computer, it still manages to preserve important properties like entanglement fidelity [85]. The channel coefficients constitute a probability simplex [151] (i.e. they add up to 1 and remain positive at all times). Thus, these coefficients are not independent and introduce correlations between the terms.

We will use a Dirichlet distribution [152] to model a stochastic Pauli noise channel. We will deploy Bayesian techniques to improve the accuracy of probabilistic error cancellation (PEC) [72, 153] under time-varying noise. Our results will show that Bayesian PEC can outperform non-adaptive approaches by a factor of 4.5x when measured using Hellinger distance from the ideal distribution.

Consider a single-qubit amplitude and phase damping channel (APD) [58]. Upon Pauli twirling [89]:

$$\mathcal{E}_{\text{twirl}}(\rho) = \frac{1}{4} \sum_{A \in \{I,X,Y,Z\}} A^\dagger \mathcal{E}_{\text{APD}} \left(A \rho A^\dagger \right) A \tag{6.17}$$

$$= \sum_{k=0}^{3} c_k \sigma_k \rho \sigma_k \tag{6.18}$$

an APD channel becomes a Pauli noise channel. Here, $\{\sigma_k\}_{k=0}^3 = \{I, X, Y, Z\}$ are the Pauli matrices. Thus, the coefficients of the Pauli noise channel are functions of the coefficients of the original APD channel, which in turn are functions of the decoherence times T_1 and T_2 [46]:

$$c_1 = c_2 = \frac{1}{4} \left[1 - \exp\left(-t/T_1 \right) \right] \tag{6.19}$$

$$c_3 = \frac{1}{4} \left[1 - \exp\left(-t/T_2 \right) \right] \tag{6.20}$$

$$c_0 = 1 - (c_1 + c_2 + c_3) \tag{6.21}$$

We model the decoherence times T_1 and T_2 as random variables dependent on register location i and time τ:

$$T_1 = T_1(i, \tau) \quad \text{and} \quad T_2 = T_2(i, \tau) \, . \tag{6.22}$$

It follows from Eq. (6.21) that the coefficients of the single-qubit Pauli noise channel are temporally and spatially varying stochastic processes [46] which also depend on register location i and time τ:

$$c_0 = c_0(i,\tau),\;\; c_1 = c_1(i,\tau),\;\; c_2 = c_2(i,\tau),\;\; c_3 = c_3(i,\tau)\;. \tag{6.23}$$

Having modeled the stochasticity of the single-qubit Pauli noise channel, let us generalize to the n-qubit case. The quantum noise channel model for an n-qubit register is given by:

$$\mathcal{E}_{\mathrm{x}}(\rho) = \sum_{i=0}^{N_p-1} \mathrm{x}_i P_i(n)\rho P_i(n)^\dagger\;, \tag{6.24}$$

where n is the register size, $N_p = 4^n$ is the total number of Pauli coefficients and $P_i(n)$ are the n-qubit Pauli operators. The channel coefficients are subject to the conditions:

$$\sum_{i=0}^{N_p-1} \mathrm{x}_i = 1, \quad \mathrm{x}_i \geq 0\;. \tag{6.25}$$

The N_p coefficients of the Pauli operators are then:

$$\mathrm{x}_k = \mathrm{x}_k(i,\tau) \tag{6.26}$$

where $i = (i_1, \cdots, i_n)$ identifies the register location(s) and τ is time.

A prior hypothesis for the channel can be obtained by assuming channel separability. The N_p coefficients can be obtained using a direct product:

$$\mathrm{x} = \begin{pmatrix} c_0(i=0,\tau) \\ c_1(i=0,\tau) \\ c_2(i=0,\tau) \\ c_3(i=0,\tau) \end{pmatrix} \times \cdots \times \begin{pmatrix} c_0(i=n-1,\tau) \\ c_1(i=n-1,\tau) \\ c_2(i=n-1,\tau) \\ c_3(i=n-1,\tau) \end{pmatrix}\;, \tag{6.27}$$

where $\times$ refers to the direct product.

Because the Pauli channel coefficients form a probability simplex, the natural way to model the probability distribution function $f_X(\mathrm{x})$ of the multi-dimensional Pauli channel distribution is the

Dirichlet distribution:

$$f_X(\mathrm{x}) \equiv \mathrm{Dirichlet}(\mathrm{x};\eta) \coloneqq \frac{\Gamma\left(\sum_{i=0}^{N_p-1}\eta_i\right)}{\prod_{i=0}^{N_p-1}\Gamma(\eta_i)}\left(\prod_{i=0}^{N_p-1}\mathrm{x}_i^{\eta_i-1}\right), \tag{6.28}$$

where $\eta_i \geq 0$ are the Dirichlet hyper-parameters, Γ is the Gamma function:

$$\Gamma(y) = \int_0^\infty t^{y-1}e^{-t}dt, \quad y>0 \tag{6.29}$$

and,

$$\int_{\mathrm{x}} \mathrm{Dirichlet}(\mathrm{x};\,\eta)\;\mathrm{dx} = 1. \tag{6.30}$$

Recall that we experimentally observe that the distribution of the decoherence times T_1 and T_2 fluctuates with time. Consequently, the distribution of the Pauli coefficients i.e. the Dirichlet distribution, varies with time. To be precise, the coefficients $\eta_i(\tau)$ vary with time τ. The distribution may be represented as $f_X(\mathrm{x};\tau)$.

At any circuit execution instance τ, the noise channel is obtained as a specific realization of the random variable X (the Pauli coefficients) which is just a sample from $f_X(\mathrm{x};\tau)$. The degree of channel non-stationarity can be quantified using the Hellinger distance H_x between the distributions at time τ and τ':

$$\begin{aligned}
&H_d = \sqrt{1-BC}\,,\\
&BC = \int_{\mathrm{x}} \sqrt{f(\mathrm{x};\tau)g(\mathrm{x};\tau')}d\mathrm{x}\\
&= \int_{\mathrm{x}} \sqrt{\frac{\Gamma\left(\eta_0(\tau)+\cdots+\eta_{N-1}(\tau)\right)}{\prod_{i=0}^{N-1}\Gamma\left(\eta_i(\tau)\right)}}\prod_{i=0}^{N-1}\mathrm{x}_i^{(\eta_i(\tau)-1)/2} \times \sqrt{\frac{\Gamma\left(\eta_0(\tau')+\cdots+\eta_{N_p-1}(\tau')\right)}{\prod_{i=0}^{N-1}\Gamma\left(\eta_i(\tau')\right)}}\prod_{i=0}^{N-1}\mathrm{x}_i^{(\eta_i(\tau')-1)/2}d\mathrm{x}\\
&= \frac{\sqrt{\Gamma\left(\sum_{i=0}^{N_p-1}\eta_i(\tau)\right)\Gamma\left(\sum_{i=0}^{N_p-1}\eta_i(\tau')\right)}}{\prod_{i=0}^{N_p-1}\sqrt{\Gamma(\eta_i(\tau))\Gamma(\eta_i(\tau'))}} \times \frac{\prod_{i=0}^{N_p-1}\Gamma\left(\frac{\eta_i(\tau)+\eta_i(\tau')}{2}\right)}{\Gamma\left(\frac{\sum_{i=0}^{N_p-1}\eta_i(\tau)+\eta_i(\tau')}{2}\right)}
\end{aligned} \tag{6.31}$$

where $N_p = 4^n$ and BC is the Bhattacharya coefficient.

The time-varying Pauli noise channel can be estimated using Bayes' rule:

$$f_X(\mathrm{x}|\text{observed data}) \propto f_X(\text{observed data}|\mathrm{x}) f_X^{\text{prior}}(\mathrm{x}) \,. \tag{6.32}$$

The prior is given by:

$$\text{Dirichlet}(\mathrm{x};\eta) = \frac{\Gamma\left(\sum_{j=0}^{N_p-1}\eta_j\right)}{\prod_{j=0}^{N_p-1}\Gamma(\eta_j)} \left(\prod_{j=0}^{N_p-1} \mathrm{x}_i^{\eta_j-1}\right) \tag{6.33}$$

where the η_i are estimated from the experimental data generated using Eqn. 6.27. This data can be old stale data and serves only to inform the prior.

To update the prior with new knowledge, we need to obtain a current dataset of circuit outcomes. This new new dataset will be used to obtain the Bayesian posterior using the likelihood.

This new dataset can be generated by executing the quantum circuit L times and recording the outcome v. The dataset is represented by $\{v_l\}$ where $l \in 0, \cdots, L-1$. The probability of observing v is given by:

$$p_v = \text{Tr}[\Pi_v \tilde{\mathcal{G}}_{\hat{\mathrm{x}}}(\rho_{\text{test}})] \,, \quad v \in \{0, \cdots, N-1\} \tag{6.34}$$

Here, Π_v is the orthogonal projection operator, $\tilde{\mathcal{G}}_{\mathrm{x}}$ is a noisy implementation of the ideal quantum operation $\mathcal{G}$, x is the noise parameterization and ρ_{test} is a known density matrix used for channel characterization.

The likelihood function is given by:

$$\begin{aligned}
\mathcal{L} = \Pr(\{v_\ell\}|\mathrm{x}) &= \Pr(V_0 = v_0, \cdots, V_{L-1} = v_{L-1}|\mathrm{x}) \\
&= \prod_{l=0}^{L-1} \Pr(V_l = v_l|\mathrm{x}) \\
&= \prod_{l=0}^{L-1} \text{Categorical}_{0,\cdots,N-1}(v_l; p_0, \cdots, p_{N-1}) \\
&= \prod_{l=0}^{L-1} p_0^{\delta_0(v_l)} \cdots p_{N-1}^{\delta_{N-1}(v_l)} \\
&= p_0^{C_0(\text{data})} \cdots p_{N-1}^{C_{N-1}(\text{data})} \\
&= \prod_{i=0}^{N-1} \left[\text{Tr}[\Pi_i \tilde{\mathcal{G}}_{\mathrm{x}}(\rho_{\text{test}})]\right]^{C_i(\{v_\ell\})} \,,
\end{aligned} \tag{6.35}$$

where, $C_i(\{v_\ell\}) = \sum_{\ell=0}^{L-1} \delta_i(v_\ell)$ (with $i \in \{0, 1, \cdots N-1\}$) is simply a counter function that counts how many times i appeared in the experimentally observed data post-measurement ($\delta_i(v)$ is Kronecker delta function which is 1 if $i = v$ and zero otherwise).

The log of the posterior is given by:

$$\begin{aligned} f_X^{\text{posterior}}(\mathrm{x}|\text{data}) = &\sum_{i=0}^{N-1} C_i(\text{data}) \log\left[\mathrm{Tr}[\Pi_i \tilde{\mathcal{G}}_{\mathrm{x}}(\rho_{\text{test}})]\right] + \log\Gamma\left(\sum_{j=0}^{N_p-1}\eta_j\right) - \sum_{j=0}^{N_p-1}\log\Gamma(\eta_j) \\ &+ \sum_{j=0}^{N_p-1}(\eta_j - 1)\log \mathrm{x}_j + \text{terms independent of x}\,. \end{aligned} \tag{6.36}$$

In the final step, the maximum-a-posterior (MAP) estimate is obtained as:

$$\hat{\mathrm{x}}(\tau) = \underset{\mathrm{x}}{\text{argmax}} \left[\sum_{i=0}^{N-1} C_i(\text{data}) \log\left[\mathrm{Tr}[\Pi_i \tilde{\mathcal{G}}_{\mathrm{x}}(\rho_{\text{test}})]\right] + \sum_{j=0}^{N_p-1}(\eta_j - 1)\log \mathrm{x}_j\right]\,. \tag{6.37}$$

$\hat{\mathrm{x}}(\tau)$ is our best guess for the time-varying noise at time τ. We will use this updated estimate for error mitigation using probabilistic error cancellation (PEC).

Probabilistic error cancellation (PEC) is a well-known error mitigation method [72, 153, 154]. The four broad steps of the PEC workflow are as follows. We use the convention that calligraphic symbols denote super-operators acting on density matrices:

$$\mathcal{G}(\rho) = G\rho G^\dagger. \tag{6.38}$$

First, expand an ideal unitary gate $\mathcal{G}$ as a (noise-model dependent) linear combination of implementable noisy gate set $\{\tilde{\mathcal{G}}_j\}$ (with its ideal counterpart $\{\mathcal{G}_j\}$), as follows:

$$\mathcal{G} = \sum_{j=0}^{N_p-1} \theta_j \tilde{\mathcal{G}}_j, \tag{6.39}$$

where θ_j are real coefficients, and $\mathcal{E}_{\mathrm{x}}$ is an error channel (such as Pauli noise channel), and $\tilde{\mathcal{G}}_j \equiv \mathcal{E}_{\mathrm{x}} \circ \mathcal{G}_{\mathrm{j}}$. The implementable noisy gate set $\{\tilde{\mathcal{G}}_j\}$ is also called noisy basis circuit set. For example, if $G = H \otimes H$ is a two-qubit Hadamard gate, then the noisy basis circuits are given by

$$\tilde{\mathcal{G}}_{\sigma\sigma'} = \mathcal{E}_{\mathrm{x}} \circ \mathcal{P}_{\sigma\sigma'} \circ \mathcal{G} \tag{6.40}$$

where

$$\mathcal{P}_{\sigma\sigma'}(\cdot) \equiv (\sigma \otimes \sigma')(\cdot)(\sigma \otimes \sigma') \quad , \tag{6.41}$$

and σ, σ' are picked from the set of Pauli matrices $\{I, X, Y, Z\}$.

The second step of PEC involves estimating the expectation value of the noise-mitigated observable as:

$$\sum_{j=0}^{N_p-1} \theta_j \left\langle \tilde{\mathcal{G}}_j \right\rangle . \tag{6.42}$$

The ideal gate can be approximated as:

$$\mathcal{G} = \sum_{w=0}^{N_p-1} p(w)[\gamma \mathrm{sgn}(\theta_w)\tilde{\mathcal{G}}_w], \tag{6.43}$$

where

$$\gamma = \sum_{w=0}^{N_p-1} |\theta_w| \tag{6.44}$$

and $p(w) = |\theta_w|/\gamma$ and w is a random variable such that $w \in 0, 1, \cdots, N_p - 1$. Said differently, the super-operator $\gamma \mathrm{sgn}(\theta_w)\tilde{\mathcal{G}}_w$ is an unbiased estimator for the ideal super-operator $\mathcal{G}$ since:

$$\mathcal{G} = \sum_{w=0}^{N_p-1} \mathrm{sgn}(\theta_w)|\theta_w|\tilde{\mathcal{G}}_w \tag{6.45}$$

$$= \sum_{w=0}^{N_p-1} \mathrm{sgn}(\theta_w)\frac{|\theta_w|}{\sum |\theta_w|}(\sum |\theta_w|)\tilde{\mathcal{G}}_w \tag{6.46}$$

$$= \sum_{w=0}^{N_p-1} p(w) \left[\gamma \mathrm{sgn}(\theta_w)\tilde{\mathcal{G}}_w\right] \tag{6.47}$$

$$= \mathbb{E}_w \left[\gamma \mathrm{sgn}(\theta_w)\tilde{\mathcal{G}}_w\right] . \tag{6.48}$$

The third step of PEC involves sampling from each of the noisy implementable circuits and computing the mean of the observable for these noisy basis circuits.

The final step of PEC involves inferring the (hopefully) noiseless observable using a weighted sum of the mean of the observable obtained from the noisy basis circuits in previous step (as per Eqn. 6.45).

In our 2-qubit application ($\mathbb{H} \otimes \mathbb{H}$), the observables are the projection operators $\Pi_0 = |00\rangle \langle 00|$, $\Pi_1 = |01\rangle \langle 01|$, $\Pi_2 = |10\rangle \langle 10|$, and $\Pi_3 = |11\rangle \langle 11|$. The complete noisy basis circuit set is shown in Fig. 6.6.

For our noise simulation, we assume that the mean of the stochastic T_1 coherence time for the first qubit decreases uniformly in a simple step-function-like manner over five time periods, deteriorating from 150 to 60 μs. Similarly, for the second qubit, we assume that the mean of the T_1 time deteriorates from 200 to 10 μs in five time periods. Additionally, we assume a simple step-function-like decrease in the mean of the T_2 coherence time for the first qubit, deteriorating from 70 to 50 μs in five time periods. Finally, for the second qubit, we assume that the mean of the T_2 coherence time deteriorates from 130 to 62.5 μs in five time periods. The coefficients of the Pauli channel are then computed using Eq. (6.21). This setup mimics intra-calibration deterioration of the noise in a quantum circuit. We assume a typical execution time of 100 μs for the Hadamard gate on the IBM transmon platform.

As described before, we model the distribution of the Pauli coefficients using a time-varying Dirichlet distribution. The Hellinger distance between the density at time τ=0 and a later time is a measure of the degree of non-stationarity, increasing from 0 to 57% as shown in Fig. 6.7. The true means of the time-varying noise Pauli channel coefficients are shown in Table 6.1. In period 0, the coefficient of the identity term in the Pauli noise channel is 38%, but by period 2, it degrades to 26%. This change is driven by the deterioration in the coherence times for qubit 0 and qubit 1, respectively, as described before.

At the outset, we assume we are in period 0, equipped with accurate knowledge of the Pauli noise channel. With this knowledge, we obtain the super-operator expression for the noisy basis circuits, which we linearly combine to estimate the ideal operation $(H \otimes H)$. The reconstructed operation is a weighted average using a quasi-probability distribution that uses the true noisy basis. Our results, shown in Fig. 6.9, indicate an accurate ideal gate implementation in the presence of noise in period 0, with a Hellinger distance between the expected and observed output of 0.34%. This small error stems from the shot noise due to finite sample size.

In subsequent time periods, the noise characteristics of the Pauli channel change as shown in Fig. 6.7 and Table 6.1. This change renders the previously implemented PEC approach invalid as the super-operators characterizing the noisy basis circuits are no longer accurate. Consequently, the PEC coefficients are also invalid.

The black bars in Fig. 6.9 indicate that the Hellinger distance between the output and ideal increases from 0.34% to 7%, and 15% in periods 1, and 2, respectively. To examine the raw data of the obtained histograms for the non-adaptive case, refer to the crimson colored bars in Fig. 6.8.

For the Bayesian update, we use the histogram of projective measurements obtained from applying the PEC circuits to the input density matrix shown in Eq. (6.49). The resulting observation stream of 2-bit strings belongs to one of four possibilities: $00, 01, 10, 11$ with probabilities $0.76, 0.08, 0.10,$

and 0.06, respectively. The input density matrix must have sufficient off-diagonal components to produce an observation rich histogram (as opposed to a bland uniform histogram which will make it impossible to differentiate between the Pauli coefficients).

$$\rho_{\text{test}} = \begin{pmatrix} 0.2 & 0.22 - 0.02j & 0.15 - 0.09j & 0.16 - 0.1j \\ 0.22 + 0.02j & 0.24 & 0.16 - 0.08j & 0.19 - 0.1j \\ 0.15 + 0.09j & 0.16 + 0.08j & 0.36 & 0.14 + 0.06j \\ 0.16 + 0.1j & 0.19 + 0.1j & 0.14 - 0.06j & 0.21 \end{pmatrix}. \tag{6.49}$$

Table 6.2 and Table 6.3 demonstrate the MAP estimation of Pauli coefficients and illustrates the quality of Bayesian estimation. With such updated estimates, we re-compute the super-operators for the noisy basis circuits and the linear combination coefficients. PEC is then implemented using the updated super-operators, resulting in improved output quality, as shown in Fig. 6.8. In this figure, the y-axis represents the probability of observing a particular computational basis state. In this two-qubit case, there are four possible states, and their probabilities sum to 1. The black bars indicate the ideal probabilities, while the red and orange bars represent the probabilities obtained with non-adaptive and adaptive PEC, respectively, for the fourth time-period. The graph demonstrates that adaptive PEC, which used adaptive estimation of the noise super-operators, improves accuracy compared to non-adaptive PEC. Specifically, the probability of observing 00 in period 2 increases from 57% for non-adaptive PEC to 72% for adaptive PEC. The likelihood and cost functions used in the Bayesian inference procedure are shown in Figures 6.10 and 6.11.

Fig. 6.9 compares the performance of adaptive and non-adaptive PEC implementations across four back-to-back time-periods in the presence of time-varying noise. The y-axis represents the Hellinger distance, which is the distance between two discrete probability distributions over the computational basis states. The black bars show the experimentally observed distribution when non-adaptive PEC is used, while the orange bars show the distribution when adaptive PEC is used. The x-axis represents the four time-periods. When using non-adaptive PEC, the Hellinger distance from the ideal distribution is 7%, and 15% for time-periods 1, and 2, respectively. When using adaptive PEC, the Hellinger distance significantly improves to 1.1%, and 3.1% for the same time-periods.

Note that we used a Pauli noise channel with all 4^n terms. However, in practical applications, it becomes necessary to reduce the number of terms. To achieve this, one can explore the use of a sparse Lindbladian noise model [27], which considers noise only in nearest-neighbor connections for Pauli terms with weight greater than 1. This reduction in terms leads to a linear scaling instead of exponential with the number of qubits, making the model more computationally efficient. Also note

that while it has been observed that single-qubit gate noise can be more than 10 times smaller than two-qubit gate noise [27], it cannot be disregarded in Probabilistic Error Cancellation (PEC) due to error propagation effects. Moreover, in the presence of time-varying quantum noise, it becomes even more critical to account for single-qubit noise to ensure accurate error cancellation.

Note that re-calibrating the noise model does not solve the challenge of dynamic estimation, as experimental evidence from various studies indicates significant fluctuations in decoherence times over time. These fluctuations, observed in studies like [19, 37], show that decoherence times can vary by approximately 50% within an hour due to the presence of oxides on superconductors' surfaces, represented as fluctuating two-level systems (TLS) [42, 36]. These fluctuations at the scale of minutes and hours are considered non-systematic noise, which cannot be addressed solely through re-calibrations. This emphasizes the necessity for Bayesian algorithms that can provide reliable error bars (as opposed to erroneous point estimates from MLE).

To summarize, this section delved into the behavior of non-stationary noise channels on real-world quantum computing platforms, with a focus on Pauli noise channels in superconducting qubits. The investigation specifically addressed the spatio-temporal non-stationarity of these noise channels, particularly in the context of ibm_belem transmon device. Spatial correlations within multi-qubit noise were explored by treating it as a collection of single-qubit channels while retaining the spatial correlations between individual qubits' T_1 and T_2 times. The Dirichlet distribution modeled the joint distribution of Pauli noise channel coefficients, while the Hellinger distance gauged the reliability of error channel characterization. The impact of time-varying Pauli noise on quantum information encoded in a n-qubit register was characterized, and coefficients of a separable 2-qubit Pauli noise channel were obtained. These coefficients directly relate to the decoherence times of individual elements, showing strong correlations among them. PEC is effective when noise is well-characterized, but non-stationary noise necessitates an adaptive approach. An adaptive Bayesian inference strategy is proposed to enhance PEC performance in the presence of time-varying noise. This approach dynamically estimated the time-varying Dirichlet distribution of Pauli coefficients using a Bayesian inference-based rolling update. An application of adaptive PEC for executing a Hadamard operation on two-qubits amidst time-varying noise was presented, revealing the need for adaptability due to changing noise characteristics. Without adaptive mitigation, outdated coefficients lead to inaccuracies and newly introduced noise affects circuit execution, severely impacting output accuracy.

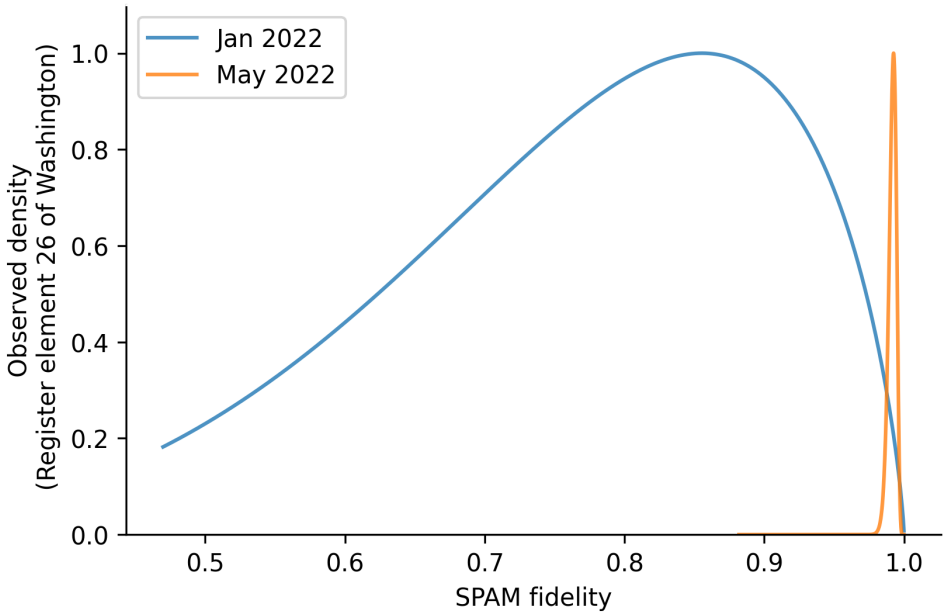


Figure 6.1: Time-varying density of SPAM fidelity. Data shown for register element #26.

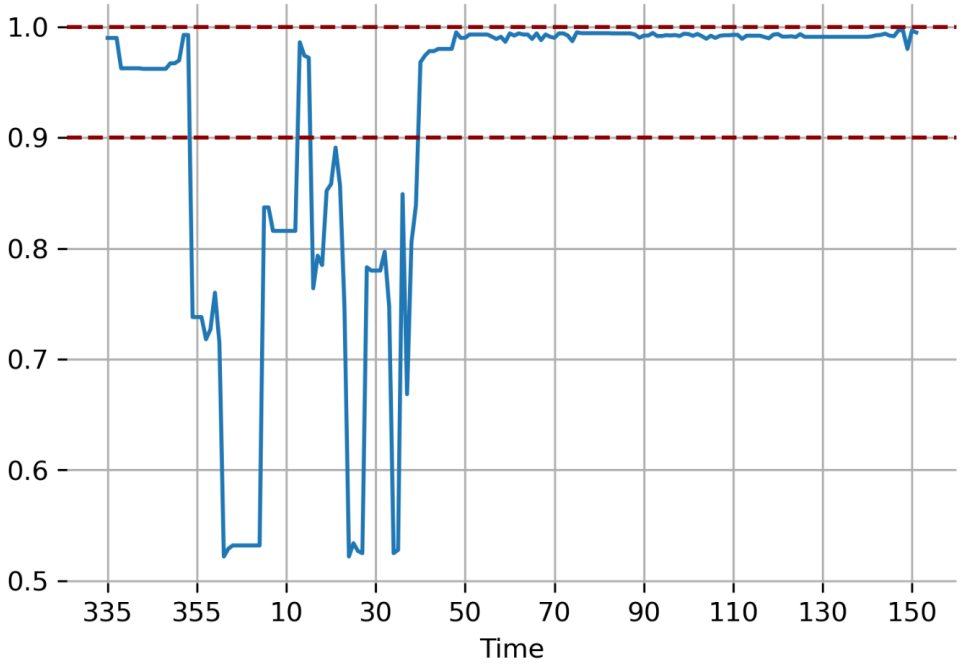


Figure 6.2: SPAM fidelity time-series for qubit #26 for Dec-May 2022.

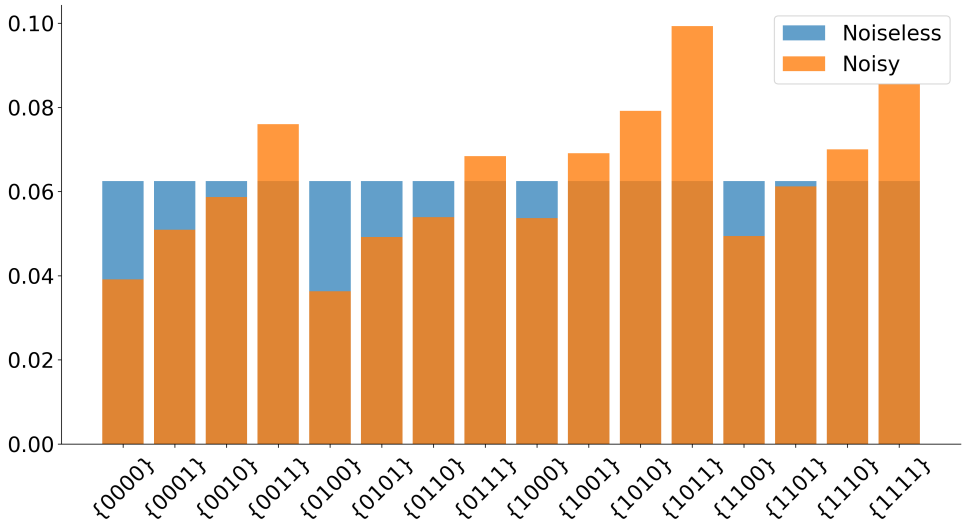


Figure 6.3: We simulated SPAM and gate error channels for a quantum circuit with 4 qubits that creates a uniform superposition across all the computational basis states using Hadamard gates. The blue bars represent the probability distribution across the measurement outcomes for the ideal, noiseless circuit while the orange bars represent the same for a realization of an execution on an unstable device.

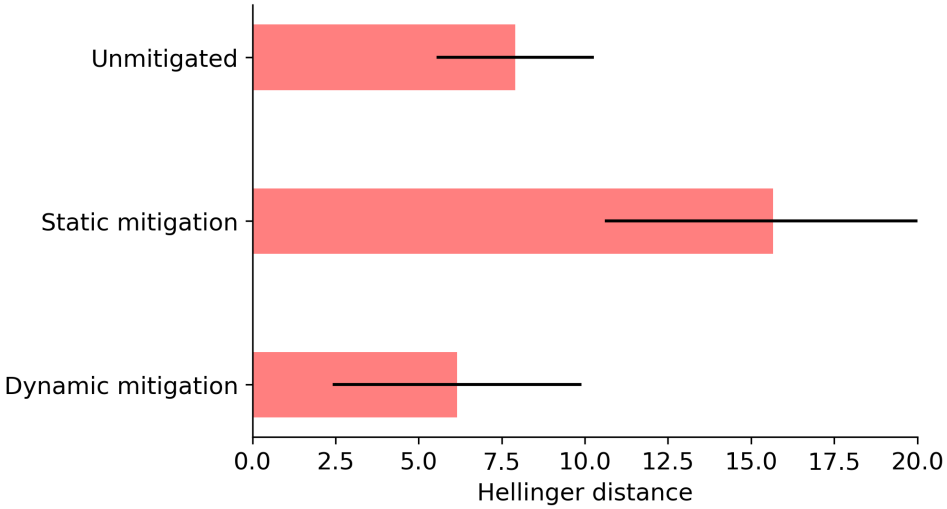


Figure 6.4: The plot shows improved accuracy achieved using Bayesian optimization in presence of time-varying noise. Mitigation using average parameters deteriorates accuracy. The number of samples used in the Bayesian estimation process is held constant at 10^4.

Table 6.1: True means of the time-varying noise Pauli channel coefficients

Pauli term [qubit0 $\otimes$ qubit1]	Period 0	Period 1	Period 2
II	0.379	0.326	0.26
IX	0.056	0.064	0.075
IY	0.056	0.064	0.075
IZ	0.076	0.078	0.079
XI	0.081	0.083	0.082
XX	0.012	0.016	0.024
XY	0.012	0.016	0.024
XZ	0.016	0.02	0.025
YI	0.081	0.083	0.082
YX	0.012	0.016	0.024
YY	0.012	0.016	0.024
YZ	0.016	0.02	0.025
ZI	0.127	0.12	0.108
ZX	0.019	0.024	0.031
ZY	0.019	0.024	0.031
ZZ	0.026	0.029	0.033

Table 6.2: Estimated Pauli coefficients for peiod 1

Pauli term [qubit0 $\otimes$ qubit1]	Estimated value	True value
II	0.311	0.326
IX	0.064	0.064
IY	0.064	0.064
IZ	0.097	0.078
XI	0.078	0.083
XX	0.016	0.016
XY	0.016	0.016
XZ	0.024	0.02
YI	0.078	0.083
YX	0.016	0.016
YY	0.016	0.016
YZ	0.024	0.02
ZI	0.114	0.12
ZX	0.023	0.024
ZY	0.023	0.024
ZZ	0.036	0.029

Table 6.3: Estimated coefficients for period 2

Pauli term [qubit0 $\otimes$ qubit1]	Estimated value	True value
II	0.243	0.26
IX	0.073	0.075
IY	0.073	0.075
IZ	0.103	0.079
XI	0.074	0.082
XX	0.022	0.024
XY	0.022	0.024
XZ	0.031	0.025
YI	0.074	0.082
YX	0.022	0.024
YY	0.022	0.024
YZ	0.031	0.025
ZI	0.103	0.108
ZX	0.031	0.031
ZY	0.031	0.031
ZZ	0.044	0.033

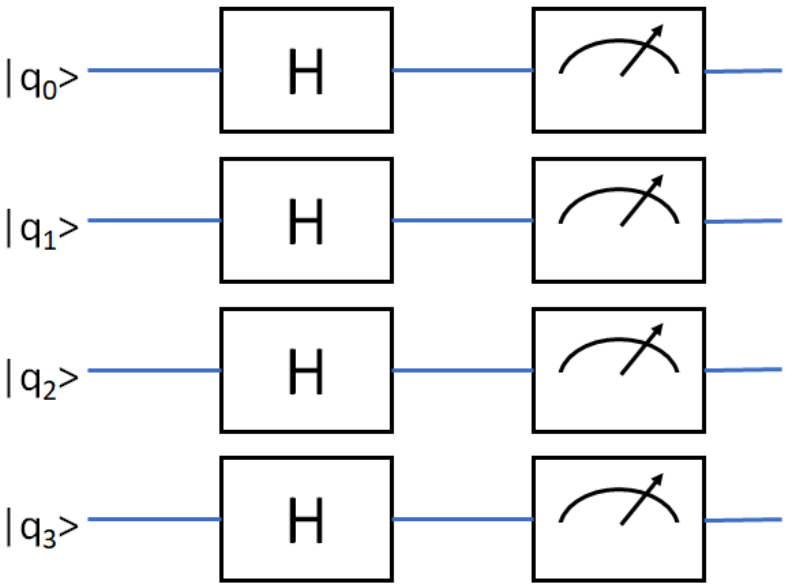


Figure 6.5: The 4-qubit quantum circuit used for the simulation experiment. Each qubit is assumed to have a different, independent SPAM error process. Each Hadamard gate is similarly assumed to have a different, independent gate error process.

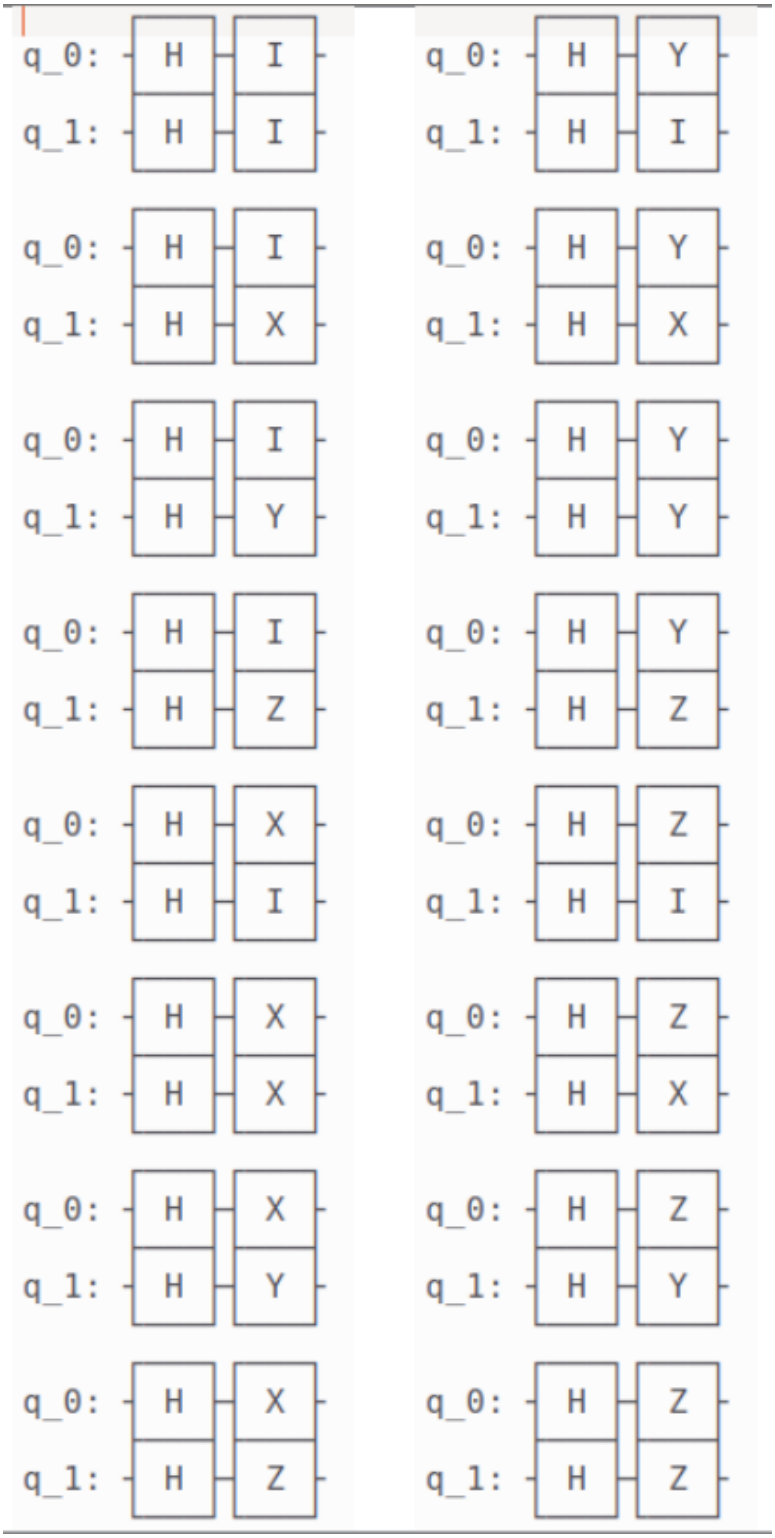


Figure 6.6: The figure displays the noisy basis circuits for the linear combination step (the first step) in PEC, where each figure represents the noisy operations $\mathcal{E}_{\mathrm{x}} \circ \mathcal{P}_{\sigma,\sigma'} \circ \mathcal{G}$ and $G = H \otimes H$ is the desired, ideal operation. Note that different basis set choices are possible, depending on the hardware.

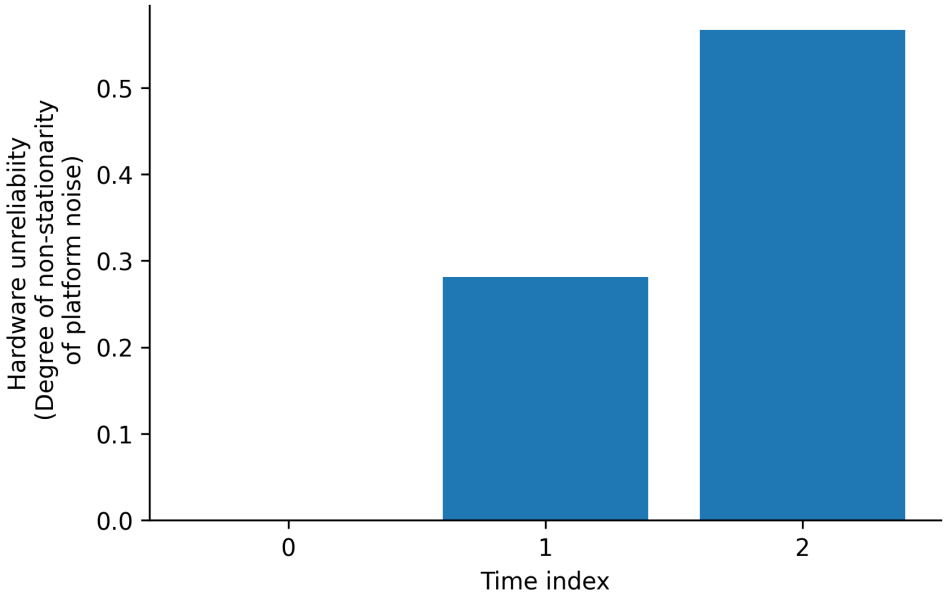


Figure 6.7: We model the degradation of a non-stationary Pauli noise channel by assuming the coherence time steadily decreases over time. Using a non-stationary Dirichlet distribution, we model the joint distribution of coefficients for a two-qubit circuit, which fluctuate as coherence times deteriorate. The y-axis represents the degree of non-stationarity, and the x-axis shows four time-periods. The Hellinger distance between the Dirichlet distributions at time τ=0 and a later time is a measure of non-stationarity, increasing from 0 to 57%. This model is based on transmon platforms and is used as an experimental setup for our simulation experiments.

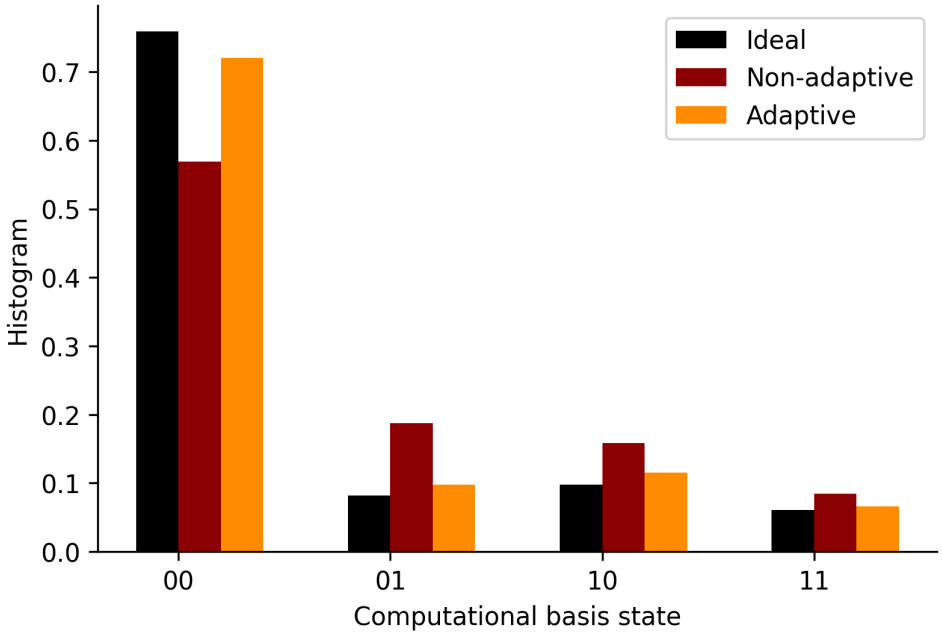


Figure 6.8: This graph compares adaptive and non-adaptive PEC for the $H \otimes H$ gate under time-varying noise. The y-axis shows the probability of observing a basis state. The black bars represent the ideal histogram for the test input, while red/orange bars are non-adaptive PEC results. Adaptive PEC improves accuracy. For example, the $|00\rangle$ probability increases from 57% to 72%.

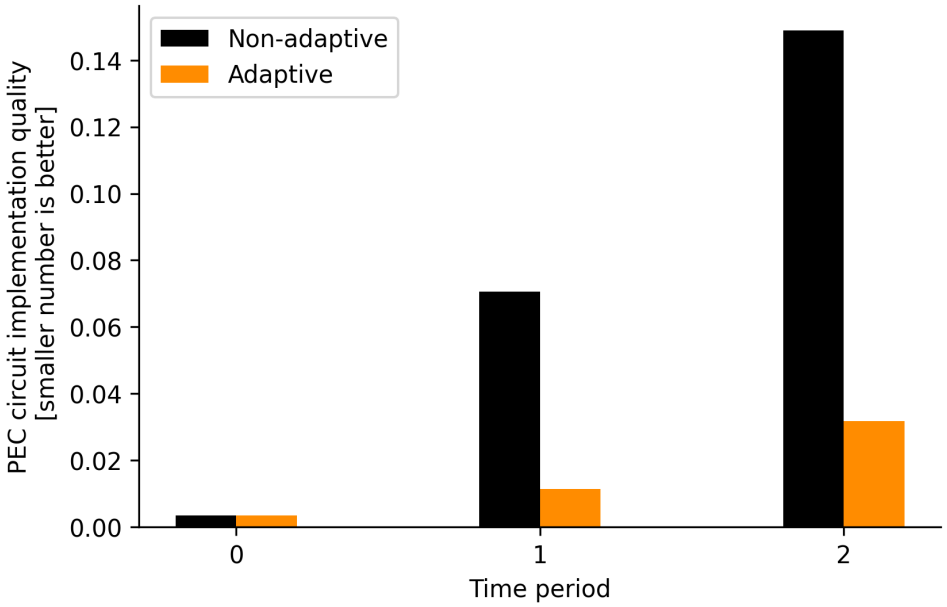


Figure 6.9: This graph compares adaptive and non-adaptive PEC implementations over four time-periods with time-varying noise, using Hellinger distance to measure the difference from the ideal distribution. The black and orange bars represent the observed distributions with non-adaptive and adaptive PEC, respectively. Adaptive PEC significantly outperforms non-adaptive PEC in reducing Hellinger distance to 1.1%, and 3.2% compared to 7%, and 15% for non-adaptive PEC across time-periods. Adaptive PEC uses adaptive estimation of noise super-operators to improve accuracy compared to non-adaptive PEC. Time-varying noise underscores the need for adaptive PEC.

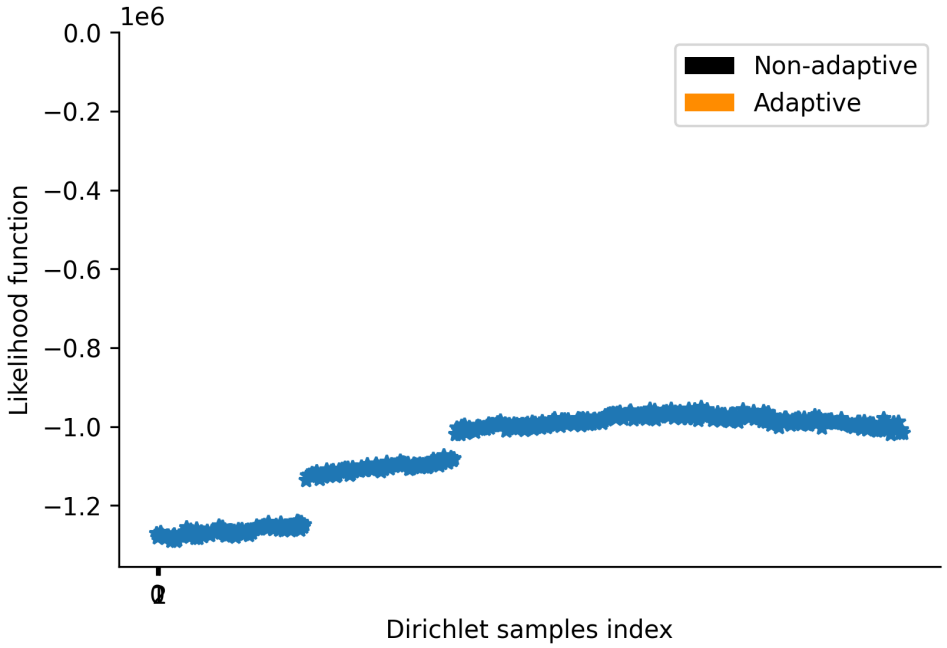


Figure 6.10: The likelihood function used in the Bayesian inference procedure for the second period.

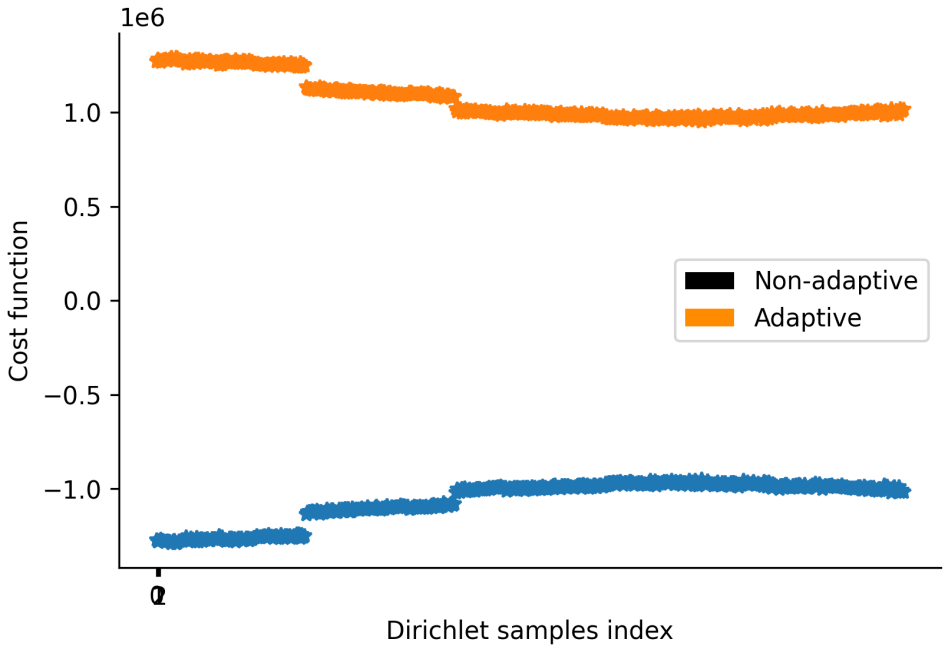


Figure 6.11: The cost function used in the Bayesian inference procedure for the second period.

Chapter 7

Adaptive probabilistic error cancellation

In this final chapter, we investigate the accuracy and stability of probabilistic error cancellation (PEC) outcomes in the presence of non-stationary noise, which is an obstacle to achieving accurate observable estimates. Leveraging Bayesian methods, we design a strategy to enhance PEC stability and accuracy.

The practical realization of quantum computing has witnessed rapid advancements [55], with quantum devices now operating as systems with hundreds of interacting qubits. However, these real quantum devices [30] are noisy [34], and practical efforts to realize a quantum computer introduce various noise processes like decay, de-coherence [155], environmental coupling, intra-register cross-talk [40, 41], and leakage from computational space [31]. Physical operations like quantum gates and measurements rely on electromagnetic fields susceptible to pulse distortion, attenuation, jitter, and drift, which further increase noise [54, 156]. Imperfections in thermodynamic controls (such as cryogenic cooling, magnetic shielding, vibration suppression, and imperfect vacuum chambers) can disturb the operating conditions of the quantum computer [70, 71].

These lead to computational errors that make it essential to address noise and implement error mitigation strategies [59] to improve the accuracy of quantum outcome [68].

Contemporary quantum computers are not only noisy but they also exhibit non-stationarity. Non-stationary noise processes [157, 44, 46, 42, 36] in superconducting qubits [33] are well-studied. For example, state preparation and measurement (SPAM) fidelities have been observed to fluctuate

significantly showing more than 25% deviation from their long-term average [110]. Similarly, the fidelity of CNOT gates have been noted to change by over 40% within similar time frames [112]. Moreover, the qubit relaxation times, known as T_1, have experienced fluctuations of up to 400% in just 30 minutes [121]. Likewise, de-phasing times, denoted as T_2, have been recorded to vary by over 50% within an hour [19, 37].

The non-stationarity observed in contemporary superconducting quantum computers stems from two primary sources linked to material defects: impurities within the material and ionization induced by cosmic rays. It is theorized that fluctuating two-level systems, possibly stemming from certain oxides on the superconductor's surface, contribute to non-stationarity [42, 36]. Additionally, cosmic rays [43, 44] contribute by ionizing the substrate upon impact, leading to the emission of high-energy phonons, which in turn triggers a burst of quasi-particles. These quasi-particles disrupt qubit coherence across the device. It has been shown that quantum computers can experience catastrophic errors in multi-qubit registers approximately every 10 seconds due to cosmic rays originating from outer space [44]. Studies that address non-stationary noise in superconducting quantum computers include investigations on output reproducibility [45], noise modeling [46], tracking the non-stationary profile of quantum noise [47], and quantum error mitigation using continuous control [48].

Quantum error mitigation is a set of techniques that employs statistical tools from estimation theory to reduce the impact of noise in quantum computations without directly correcting the quantum state [59]. Such techniques can become vulnerable to errors stemming from over or under estimation of noise due to the presence of non-stationarity [158].

A quantum noise channel [11] can be described as a stochastic process, allowing for the continuous update of its estimated characteristics in response to varying noise conditions. Such an approach treats the channel as a time-varying random variable. Consider Fig. 7.1 which shows the state preparation and measurement (SPAM) fidelity for the second register element on ibm_kolkata device on Jan 15, 2024. The probability density is clearly changing with time, even though the variance stays consistent. In light of such non-stationary data, a single-qubit SPAM noise channel can be described by the model: $\mathcal{E}(\rho) = f(t)\rho + [1 - f(t)]\mathbb{X}\rho\mathbb{X}$, where ρ represents the single-qubit density matrix, $\mathbb{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ denotes the Pauli-X matrix, and $f(t)$ denotes the SPAM fidelity drawn from a time-dependent distribution. From Fig. 7.1, non-stationarity of noise in this superconducting device is apparent within 24 hours. The noise parameters estimated during device re-calibration quickly becomes outdated, compromising the accuracy of noise channel information essential for mitigation. Thus re-calibration alone is insufficient and there is a need for adaptive error mitigation techniques that can function in between calibration intervals in the face of changing noise conditions.

This study focuses on probabilistic error cancellation (PEC) [72] in the presence of non-stationary noise. PEC is a quantum error mitigation approach that aims to construct unbiased estimates of the means of quantum observables from noisy observations. Effective implementation requires an accurate noise characterization [159]. For example, learning correlated noise channels in large quantum circuits on a superconducting quantum processor has proven to be difficult [27]. Yet, leveraging sparse noise models, PEC has been successful in estimating the mean of observables in circuits comprising 2,880 CNOT gates, executed on a 127-qubit noisy superconducting processor - a task that conventional brute-force computing could not match [35].

We demonstrate that adaptive probabilistic error cancellation, which views quantum noise channels as evolving random variables, outperforms its non-adaptive counterpart in devices subject to non-stationary noise. To achieve this aim, we will make use of a 5-qubit implementation of the Bernstein-Vazirani algorithm [109].

The manuscript is organized as follows. In Sec. 7.1, we provide background for probabilistic error cancellation (PEC). Sec. 7.2 develops a Bayesian [73, 74] approach for adapting the method of probabilistic error cancellation to non-stationary noise. It also sets up a performance evaluation framework for the accuracy and stability of PEC results. Sec. 7.3 presents a numerical validation of adaptive probabilistic error cancellation using a 5-qubit implementation of the Bernstein-Vazirani algorithm, where we treat the noise parameters (qubit-specific state preparation and measurement (SPAM) fidelities and depolarizing parameters characterizing noise in CNOT gate) as non-stationary random variables. In Sec. 7.4, we present the results of experiments conducted on a real, noisy quantum device to test the adaptive PEC algorithm. Concluding remarks are provided in Sec. 8.

7.1 Background

In this section, we provide background for the quantum error mitigation method called probabilistic error cancellation (PEC) [72, 27, 35] which aims to mitigate errors by approximating the noiseless mean of the observable as a weighted sum of noisy observables. We use calligraphic symbols to denote super-operators acting on density matrices (ρ):

$$\mathcal{G}\rho = G\rho G^{\dagger}. \tag{7.1}$$

where $\mathcal{G}$ is the super-operator and G is a unitary quantum operator. For example, if G denotes the CNOT operator, then $\mathcal{G}$ is the super-operator for the CNOT operation.

Typically we do not have access to a noiseless implementation of $\mathcal{G}$. Let $\tilde{\mathcal{G}}$ denote the super-operator corresponding to the available, noisy implementation of $\mathcal{G}$. We have access to other synthesized implementations of $\tilde{\mathcal{G}}$ by subjecting $\tilde{\mathcal{G}}$ to basis operations available on a noisy device. Originally, the noisy basis set was specified as the set of native gates that a quantum computer could implement [72]. However, with advancements in superconducting quantum computers, the noise associated with single-qubit Pauli operators has become negligible [27, 28]. Consequently, Pauli operators, which might be a composition of multiple native gates, can be employed as the basis set [160]. We denote by $\{\tilde{\mathcal{G}_k}\}$ the set of all noisy super-operators, with $k = 0$ denoting the noisy implementation of $\mathcal{G}$ without additional operations from the basis set.

As an example, consider a 2-qubit quantum gate G. The set of noisy super-operators composed under a Pauli channel assumption are given by: $\{\mathcal{P} \circ \tilde{\mathcal{G}}\}$ where $\mathcal{P} \circ \tilde{\mathcal{G}}(\rho) = \mathcal{P}(\tilde{\mathcal{G}}(\rho))$ and $\mathcal{P}(\cdot) \equiv [\mathbb{P}_0 \otimes \mathbb{P}_1](\cdot)[\mathbb{P}_0 \otimes \mathbb{P}_1]$. Here 0 and 1 refer to qubit 0 and qubit 1 respectively and $\mathbb{P}_0, \mathbb{P}_1$ are picked from the set of Pauli operators $\{\mathbb{I}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}\}$. Thus, by varying P_0 and P_1, we obtain the set $\{\mathcal{P} \circ \tilde{\mathcal{G}}\}$ which forms a basis $\{\tilde{\mathcal{G}_k}\}_{k=0}^{15}$ that spans the super-operator space. In this 2-qubit example, the map for k is derived from the cartesian product of $\{\mathbb{I}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}\} \times \{\mathbb{I}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}\}$, with the sequence of this ordered set determining the value of k.

In general, the super-operator $\mathcal{G}$ can be expressed as a linear combination of the basis super-operators:

$$\mathcal{G} = \sum_{k=0}^{N_p - 1} \eta_k \tilde{\mathcal{G}}_k, \tag{7.2}$$

where N_p is the dimension of the super-operator space. The PEC coefficients η_k in the linear combination are determined either analytically, under a noise model assumption for single and two-qubit gates, which can be extended to larger circuits, or numerically, by minimizing the one-norm between high-dimensional matrices [160, 72]. We employ the analytical approach.

The circuits corresponding to the super-operators $\tilde{\mathcal{G}}_k$ in Eqn. 7.2 are constructed by subjecting each of the gates in the quantum circuit for $\tilde{\mathcal{G}}$ to operations from the noisy basis set. If we execute these noisy circuits and collect the mean of the observable, then, from Eqn. 7.2, we recover the ideal noiseless mean of an observable $\langle \mathcal{O} \rangle$ as:

$$\langle \mathcal{O} \rangle = \mathrm{Tr}\left[\mathcal{O}\mathcal{G}\rho\right] = \mathrm{Tr}\left[\mathcal{O} \sum_k \eta_k \tilde{\mathcal{G}}_k \rho\right] \tag{7.3}$$

where ρ is the input density matrix to the circuit. Thus, an observable $\langle \mathcal{O} \rangle$ may be estimated as a weighted sum of the mean of the observables from the noisy circuits.

Eqn. 7.3 can be re-written as:

$$\langle \mathcal{O} \rangle = \gamma \sum_{k} \text{sgn}(\eta_{\text{k}}) \mathbb{Q}_{\text{k}} \text{Tr}\left[\mathcal{O}\tilde{\mathcal{G}}_{\text{k}}\rho\right] \tag{7.4}$$

with $\gamma = \sum |\eta_{\text{k}}|$ and $\mathbb{Q}_{\text{k}} = |\eta_{\text{k}}|/\gamma$. The sign function, sgn$(\cdot)$, returns +1 for positive inputs, -1 for negative inputs, and 0 for an input of 0.

Note that the set $\{\mathbb{Q}_k\}$ forms a valid probability distribution because all its elements are positive and sum to 1. However, $\{\mathbb{Q}_k\}$ is termed a quasi-probability distribution (QPD) [72, 160]. To see this, consider a random integer $K \in \{0, \cdots N_p - 1\}$ which follows the probability distribution function denoted by $\mathbb{Q}_k$. This random integer K can be mapped one-to-one to the random variable $\text{sgn}(\eta_{\text{K}})\text{Tr}\left[\mathcal{O}\tilde{\mathcal{G}_{\text{K}}}\rho\right]$, which inherits the probability distribution function $\mathbb{Q}_k$. By repeatedly sampling the random variable K, we realize a set of random integers represented as $\{\mathfrak{m}\}$. For each specific $\mathfrak{m}$ that is realized, we execute the corresponding noisy quantum circuit $\tilde{\mathcal{G}_{\mathfrak{m}}}$ multiple times and obtain a set of noisy means of observables denoted by the set $\left\{\text{Tr}\left[\mathcal{O}\tilde{\mathcal{G}_{\mathfrak{m}}}\rho\right]\right\}$.

When computing the average over the set of noisy means of observables $\left\{\text{Tr}\left[\mathcal{O}\tilde{\mathcal{G}_{\mathfrak{m}}}\rho\right]\right\}$, we need to adjust the sign of each element of the set by the sign of η_{m} and scale it by γ. This average then converges to the mean of the observable from a noiseless gate $\mathcal{G}$, in the asymptotic limit of a large number of repeated samplings of the random variable K. The need for adjustment by the sign function leads us to denote $\mathbb{Q}_k$ as quasi-probabilities.

Achieving an accuracy of $O(\epsilon)$ using the empirical mean of the random variable $\gamma\text{sgn}(\eta_{\text{K}})\text{Tr}\left[\mathcal{O}\tilde{\mathcal{G}_{\text{K}}}\rho\right]$ (which is an unbiased estimator of $\langle\mathcal{O}\rangle$) requires $O(\gamma/\epsilon)^2$ PEC circuit samples and the result has variance of order $O(\gamma^2)$ [72].

7.2 Adaptive PEC

Estimating channel noise parameters is crucial for determining the PEC coefficients $\{\eta_{\text{k}}\}$. However, the non-stationary nature of noise, along with drift and latency in characterization, complicates this task. This in turn makes it difficult to accurately assess the PEC coefficients. In this section, we demonstrate how adaptive parameter estimation can be applied to PEC.

The parameters characterizing the noise during idle time (such as qubit decoherence time [18]) and quantum operations (such as CNOT fidelity [110]) exhibit random non-stationary behavior in some hardware. Estimating non-stationary stochastic processes is challenging and their predictive value is limited because the patterns identified from historical data may not reliably indicate future behavior,

making it difficult to discern underlying trends. A model that is effective at one time point can become inaccurate at another.

However, we can utilize intermittent incremental measurements from quantum circuits to devise a Bayesian [73, 74, 75, 76] update for the current state of the device noise:

$$\Pr(\mathbf{x}|\text{data}) \propto \Pr(\text{data}|\mathbf{x}) \Pr(\mathbf{x}) \tag{7.5}$$

where data refers to measurements obtained from circuit execution, $\mathbf{x}$ denotes the multi-dimensional vector of parameters characterizing the device noise (such as connection-specific CNOT fidelity and qubit-specific SPAM fidelities), $\Pr(\mathbf{x}|\text{data})$ is the posterior noise distribution, $\Pr(\text{data}|\mathbf{x})$ is the likelihood, and $\Pr(\mathbf{x})$ is the prior noise distribution. We use this to update $\{\eta_{\text{k}}\}$ via updates to model-specific parameters.

Obtaining the posterior distribution for x is a two-step process: first, we estimate the posterior for uncorrelated parameters, and second, for correlated parameters in the underlying noise model.

7.2.1 Uncorrelated parameters

Uncorrelated parameters refer to those model parameters for which there exist datasets (such as the partial trace of the observable on a single qubit) where the observed data is modeled by error due to only one noise parameter. Here, the analysis is simpler as univariate priors can be used. For example, if we disregard noise from single-qubit rotations, then the measurements obtained from any qubit that was not subjected to entangling operations, can be described by a one-parameter model under symmetric SPAM noise model.

Let $\mathcal{D}_q = \{b_q(0), \cdots, b_q(L-1)\}$ represent a dataset of L samples obtained for qubit q after measurement in the computational basis. Each $b_q(l)$ is a single-bit measured after the l-th execution of $\tilde{\mathcal{G}}_0$. We can adopt a beta distribution as the prior for the SPAM fidelity f_q because the beta distribution is well-suited for values restricted to the [0,1] interval and can effectively accommodate the experimentally observed unimodal and skewed density as seen in Fig. 7.1 for qubit 2. The beta distribution's flexibility allows for an accurate fit to these characteristics. The Beta distribution with parameters α_q, β_q is given by $\text{Beta}(f_q; \alpha_q, \beta_q) = f_q^{\alpha_q - 1}(1-f_q)^{\beta_q - 1}/B(\alpha_q, \beta_q)$ where the normalizing denominator $B(\alpha_q, \beta_q)$ is the Beta function defined as $B(m,n) = \Gamma(m)\Gamma(n)/\Gamma(m+n)$, and $\Gamma(\cdot)$ is the gamma function given by: $\Gamma(y) = \int_0^\infty t^{y-1}e^{-t}dt$, defined for any positive y.

The prior for the mean μ_q^{prior} and variance v_q^{prior} of the SPAM fidelity f_q can be derived from historical characterization data (e.g. using data post calibration). If such data is not accessible, we can obtain the starting point from a small perturbation to the ideal value. The prior parameters are then obtained as: $\alpha_q^{\text{prior}} = \mu_q^{\text{prior}} \left[\mu_q^{\text{prior}}(1-\mu_q^{\text{prior}})/v_q^{\text{prior}} - 1\right]$ and $\beta_q^{\text{prior}} = (1-\mu_q^{\text{prior}}) \left[\mu_q^{\text{prior}}(1-\mu_q^{\text{prior}})/v_q^{\text{prior}} - 1\right]$.

The likelihood is obtained as:

$$\mathcal{L}(f_q) = f_q^{C_0[\mathcal{D}_q]}(1-f_q)^{L-C_0[\mathcal{D}_q]}/B(\alpha_q,\beta_q)^L \tag{7.6}$$

where $C_0[\mathcal{D}_q]$ counts the number of 0's in the dataset $\mathcal{D}_q$. The updated posterior density (indicated by the prime on the updated parameters) for the SPAM fidelity (f_q) [161] :

$$f_q|\mathcal{D}_q \sim \text{Beta}(\alpha_q', \beta_q') \tag{7.7}$$

where $\alpha_q' = \alpha_q + L - C_0[\mathcal{D}_q]$ and $\beta_q' = \beta_q + C_0[\mathcal{D}_q]$. This shows the influence of incremental measurements on posterior noise density. The qubit-wise updated mean [162] of the SPAM fidelity, obtained as $\mu_q' = \alpha_q'/(\alpha_q' + \beta_q')$, are then used in estimating the PEC coefficients.

7.2.2 Correlated parameters

The second task involves the adaptive estimation of correlated noise parameters, which is more complex due to the measurements being influenced by multiple noise processes simultaneously. These are the noise parameters for which there exist datasets (typically the partial trace of an observable across a subset of qubits) which reflect errors from various noise parameters jointly. The analysis uses the probability of observing classical bit-strings on this qubit subset as the random variables.

For a subset of m qubits from a total of n qubits, there are $M = 2^m$ possible observed bit-strings, each with a probability denoted by $p_0, p_1, \ldots, p_{M-1}$, where p_0 represents the probability of observing all zeros and p_{M-1} that of all ones. These probabilities are treated as random variables which form a probability simplex, as they are all positive and add up to 1. Hence the natural way to model the joint density is using a Dirichlet prior given by:

$$\Pr(p_0, \cdots, p_{M-1}) = \frac{\Gamma(\sum a_i)}{\prod_0^{M-1} \Gamma(a_i)} \prod_0^{M-1} p_i^{a_i - 1} \tag{7.8}$$

where the gamma function $\Gamma(\cdot)$ was already defined previously and $\{a_i : a_i > 0, i \in \{0, \cdots M-1\}\}$ are the parameters characterizing the Dirichlet distribution that need to be estimated and updated – a multi-variate task analogous to the uni-variate task for the beta distribution in the previous section.

We denote the dataset derived from L measurements of the m qubits in the computational basis as $\mathcal{D} = \{w(0), \cdots, w(L-1)\}$, where each $w(i)$ is a binary string of length m. The likelihood function is given by:

$$\mathcal{L}(p_0, \cdots, p_{M-1}) = \prod_{i=0}^{M-1} p_i^{C_{v_i}[\mathcal{D}]} \tag{7.9}$$

where $v_i \in \{0,1\}^m$, $C_{v_i}[\mathcal{D}]$ is the count of occurrences of the string v_i in the dataset $\mathcal{D}$. Note that $w(i)$ is a specific realization post-measurement that takes one of the values in the set $\{v_i\}$.

The posterior joint density is also a Dirichlet distribution [163]:

$$\Pr(p_0, \cdots, p_{M-1} \mid \mathcal{D}) = \frac{\Gamma(\sum a_i')}{\prod_{i=0}^{M-1} \Gamma(a_i')} \prod_{i=0}^{M-1} p_i^{a_i'-1} \tag{7.10}$$

with parameters $a_i' = a_i + C_{v_i}(\mathcal{D})$. Upon marginalizing this joint density, the marginals follow a Beta distribution. For example, the random variable p_i is Beta-distributed with parameters $\alpha_i = a_i'$ and $\beta_i = -a_i' + \sum_{j=0}^{M-1} a_j'$ [164].

The updated marginals from the Dirichlet distribution give us the evolving densities for each p_i, allowing us to calculate their time-varying means and variances. Specifically, the mean for p_i updates to $a_i' / \sum a_i'$, and its variance updates to $a_i' \left(\sum_j a_j' - a_i'\right) / \left(\sum_j a_j'\right)^2 / \left(1 + \sum_j a_j'\right)$.

At the final stage, the method employs the relationship between the probabilities $p_0, \cdots, p_{M-1}$ and the noise model parameters to derive the time-dependent parameter means and variances of the noise parameters. The process concludes by updating the quasi-probability distribution using the updated noise parameters per Eqn. 7.5.

7.2.3 Accuracy and stability

We next evaluate the performance of PEC in presence of non-stationary noise using the lens of accuracy and stability. A mean of a quantum observable $\mathcal{O}$ is represented as $\langle \mathcal{O} \rangle_{\mathrm{x}}$ (where x labels the noise instance), while the observed mean after mitigation is denoted as $\langle \mathcal{O} \rangle_{\mathrm{x}}^{\mathrm{mit}}$, which equals $\langle \mathcal{O} \rangle$ only in the asymptotic limit of infinite samples and zero noise. When we want to say the mitigated noisy mean at a specific time t, we will use $\langle \mathcal{O} \rangle_{\mathrm{x}}^{\mathrm{mit}}(t)$.

We say that a PEC mitigated observable is $\epsilon-$accurate if the absolute difference between the mitigated and noiseless observable is upper bounded by ϵ:

$$| \langle O \rangle_{\mathrm{x}}^{\mathrm{mit}} - \langle O \rangle | \leq \epsilon \tag{7.11}$$

Now, suppose the underlying noise, being non-stationary, is characterized by a time-dependent density $f(\mathrm{x}; t)$. We say that a PEC mitigated observable is $\epsilon-$stable between times t_1 and t_2 if:

$$| \langle O \rangle_{\mathrm{x}}^{\mathrm{mit}} (t_1) - \langle O \rangle_{\mathrm{x}}^{\mathrm{mit}} (t_2)| \leq \epsilon \tag{7.12}$$

where

$$\begin{aligned} \langle O \rangle_{\mathrm{x}}^{\mathrm{mit}} (t) &= \int_{\mathrm{x}} \langle O \rangle_{\mathrm{x}}^{\mathrm{mit}} f(\mathrm{x}; t) d\mathrm{x} \\ &= \int_{\mathrm{x}} \sum_{k} \eta_{\mathrm{k}}(\mathrm{x}, t) \mathrm{Tr} \left[\mathcal{O} \tilde{\mathcal{G}}_k \rho \right] f(\mathrm{x}; t) d\mathrm{x} \end{aligned} \tag{7.13}$$

Note the dependence of η_{k} on the random variable x and time t.

When measuring $\mathcal{O}_q$ for qubit q, the qubit-wise accuracy metric is:

$$\epsilon_q = \left| \langle \mathcal{O}_q \rangle_{\mathrm{x}} - \langle \mathcal{O}_q \rangle \right| \tag{7.14}$$

without mitigation, and

$$\epsilon_q = \left| \langle \mathcal{O}_q \rangle_{\mathrm{x}}^{\mathrm{mit}} - \langle \mathcal{O}_q \rangle \right| \tag{7.15}$$

with error mitigation. The register average, for a n-qubit register, is:

$$\epsilon_R = \sum_{q=0}^{n-1} \epsilon_q / n \tag{7.16}$$

Similarly, the qubit-wise stability metric is:

$$s_q(t) = \left| \langle \mathcal{O}_q \rangle_{\mathrm{x}} (t) - \langle \mathcal{O}_q \rangle_{\mathrm{x}} (0) \right| \tag{7.17}$$

without mitigation, and

$$s_q(t) = \left| \langle \mathcal{O}_q \rangle_{\mathrm{x}}^{\mathrm{mit}} (t) - \langle \mathcal{O}_q \rangle_{\mathrm{x}}^{\mathrm{mit}} (0) \right| \tag{7.18}$$

with error mitigation. The register average, for a n-qubit register, is:

$$s_R(t) = \sum_{q=0}^{n-1} s_q(t)/n. \tag{7.19}$$

We expect the accuracy (ϵ_q, ϵ_R) and stability (s_q, s_R) metrics to be smaller when using adaptive PEC because of more accurate estimates for the time-varying PEC coefficients $\{\eta_k\}$.

7.3 Numerical Validation

For studying the stability and accuracy of PEC in presence of non-stationary noise, we use an implementation of the Bernstein-Vazirani algorithm [109], a standard benchmarking circuit that requires only a modest number of gates. The purpose of the algorithm is to recover a n-bit secret string r, encoded in a black-box oracle function. The algorithm identifies the secret with a single query, while the classical method needs 2^n queries (worst-case).

A quantum circuit for the Bernstein-Vazirani algorithm is shown in Fig. 7.2. We conceptualize the circuit as having four principal layers ($\mathbb{L}_1, \mathbb{L}_2, \mathbb{L}_3$, and $\mathbb{L}_4$) separated by dashed vertical lines. Note that the total number of qubits needed by an $n-1$-bit secret string is n, comprising $n-1$ data qubits and one ancilla qubit. So, a 5-qubit implementation of the Bernstein-Vazirani algorithm has $n = 5$ but the secret string length is $n - 1 = 4$.

The circuit is initialized with all the qubits in the register in the $|0\rangle$ state. Thus, a pure state description yields an initial state $|\psi_0\rangle = |0\rangle^{\otimes n}$. In the first layer, all the qubits are subjected to Hadamard gates. The ancilla qubit (which is the last qubit with index n) is additionally subjected to a $\mathbb{Z}$ gate. The output after the first layer $\mathbb{L}_1 = \mathbb{H}^{\otimes n-1} \otimes (\mathbb{Z}_n \mathbb{H}_n)$ is $|\psi_1\rangle = |+\rangle^{\otimes n-1} |-\rangle$.

The second layer implements the oracle function for which we can use CNOT gates. Each CNOT's control qubit corresponds to one of the bits in the secret string r, while the target qubit remains fixed at qubit n. Specifically, if the bit r_i of the secret string r is 1, we add a CNOT between qubit i (control) and qubit n (target). The input to the second layer is $\mathbb{H}^{\otimes n-1} |0\rangle |-\rangle$ while the output is $|\psi_2\rangle = \left(\mathbb{H}^{\otimes n-1} |r\rangle\right) |-\rangle$ where r is the secret string.

To retrieve the secret string r, the third layer requires another layer of Hadamard gates. A $\mathbb{Z}$ gate is applied to the ancilla qubit (qubit n) before the application of the Hadamard layer to make the computing reversible. The output quantum state after the third layer $\mathbb{L}_3 = \mathbb{H}^{\otimes (n-1)} \otimes (\mathbb{H}_n \mathbb{Z}_n)$ is $|\psi_3\rangle = |r\rangle |0\rangle$.

7.3.1 Observable

The fourth layer is for measurement in the $\mathbb{Z}$-basis and is not a unitary layer. We measure the state of each of the n-qubits after projection onto the computational basis states. Post-measurement, the observation obtained is a classical bit (either 0 or 1) for each of the n-qubits measured. Thus, the final observed output is a bit-string of length n, including the ancilla.

We make a distinction between the qubit-wise observables $\mathcal{O}_q$, where $q \in \{0, \cdots n\}$ and the measurement operator for the register $M_Z = \mathbb{Z}_0 \otimes \cdots \mathbb{Z}_n$ in computational basis. The qubit-wise observable $\mathcal{O}_q$ has identity across the n-qubit tensor product except in the q-th position. For example, $\mathcal{O}_0 = \mathbb{Z}_0 \otimes \mathbb{I}^{\otimes n-1}$ and $\mathcal{O}_n = \mathbb{I}^{\otimes n-1} \otimes \mathbb{Z}_n$. The eigenvalues of $\mathcal{O}_q$ are $+1$ (corresponding to classical bit 0) and -1 (corresponding to classical bit 1). The theoretical mean of the qubit-wise observable $\mathcal{O}_q$ is denoted by $\langle \mathcal{O}_q \rangle = \mathrm{Tr}[\mathcal{O}_q \rho]$ where $q \in \{0, \cdots, n\}$, which reflects the measurement process being fundamentally probabilistic.

The experimentally observed mean of the observable, denoted by $\langle \mathcal{O}_q \rangle_{\mathrm{x}}$ is calculated as the sum of eigenvalues weighted by their empirically observed frequencies. In a noiseless circuit, the observed mean asymptotically converges to the theoretical mean as the sample size tends to infinity. However, in presence of shot noise and non-zero variance of x, $\langle \mathcal{O}_q \rangle_{\mathrm{x}}$ may not equal $\langle \mathcal{O}_q \rangle$.

The experimentally observed measurements of the M_Z operator in computational basis belong to one of the 2^n eigenstates, each of which can be represented by a n-bit binary string. These observations contain the information necessary for computing $\mathcal{O}_q$ for all q. In the context of the Bernstein-Vazirani algorithm, we discard the ancilla bit and declare the search a success when the first n-bits of the observed value of the M_Z operator matches the secret string r.

7.3.2 Modeling circuit noise

Our experimental focus is on superconducting hardware in next section. The potential sources of noise in a superconducting implementation of the circuit depicted in Fig. 7.2 are: (i) state preparation noise, (ii) noise in the implementation of the Hadamard gate, (iii) noise in the implementation of the $\mathbb{Z}$ gate, (iv) noise in the implementation of the CNOT gate, and (v) measurement noise (also known as readout noise). The error resulting from the first and last noise sources, namely state preparation noise and measurement noise, are often measured collectively due to experimental limitations. This combined noise is commonly referred to as SPAM (state preparation and measurement) noise. An effective model assumes that the state preparation is noiseless, and the noise impacts the measurement (or readout) process. The relative magnitudes of the different types of noise are different. State preparation and measurement (SPAM) noise typically has the largest contribution to errors [28]. The

next most significant contribution arises from imperfect implementations of entangling gates, such as the CNOT. After that the strength of the noise, for single-qubit rotations, decreases by two orders of magnitude or more [27, 28]. The $\mathbb{Z}$ gate is a software-based operation [165] and error-free. After disregarding single-qubit rotation errors, only two predominant types of noise emerge: SPAM noise and CNOT noise.

Next, we create a quantum channel based description of the circuit noise. The circuit noise can be modeled layer-wise. The output density matrix prior to measurement can be represented as: $\tilde{\rho} = \mathcal{E}_4(\mathcal{E}_3(\mathbb{L}_3\mathcal{E}_2(\mathbb{L}_2\mathcal{E}_1(\mathbb{L}_1\rho\mathbb{L}_1^\dagger)\mathbb{L}_2^\dagger)\mathbb{L}_3^\dagger)$ where $\mathcal{E}_k(\cdot)$ denotes the noise channel for the k-th layer of the circuit in Fig. 7.2. We approximate $\mathcal{E}_1$ as identity channel because we ignore single-qubit errors.

We model noise in the CNOT gate using a 2-qubit depolarizing model. Let $\mathrm{x}_C(t)$ and $\mathrm{x}_T(t)$ represent the stochastic depolarizing parameters for the control and target qubits at time t. The depolarizing noise model for the CNOT gate is represented by:

$$\begin{aligned}
&\mathcal{E}_{\mathrm{CNOT}}(\cdot) = [1-\mathrm{x}_T(t)][1-\mathrm{x}_C(t)](\cdot)+ \\
&\frac{1-\mathrm{x}_C(t)}{3}\mathrm{x}_T\sum_{\mathbb{P}'_T}(\mathbb{I}_C\otimes\mathbb{P}'_T)(\cdot)(\mathbb{I}_C\otimes\mathbb{P}'_T) \\
&+\frac{1-\mathrm{x}_T(t)}{3}\mathrm{x}_C\sum_{\mathbb{P}'_C}(\mathbb{P}'_C\otimes I_T)(\cdot)(\mathbb{P}'_C\otimes I_T) \\
&+\frac{\mathrm{x}_C(t)\mathrm{x}_T}{9}\sum_{\mathbb{P}'_C,\mathbb{P}'_T}(\mathbb{P}'_C\otimes\mathbb{P}'_T)(\cdot)(\mathbb{P}'_C\otimes\mathbb{P}'_T)
\end{aligned} \tag{7.20}$$

In this expression, we use $\mathbb{P}'_C, \mathbb{P}'_T \in \{\mathbb{X}, \mathbb{Y}, \mathbb{Z}\}$ as the single-qubit Pauli operators excluding identity $\mathbb{I}$, acting on the control (C) and target (T) qubits respectively. The sum is over all the single-qubit Pauli operators, excluding identity. The effect of the quantum noise channel $\mathcal{E}_2(\cdot)$ for layer 2, on a 5-qubit state, combines the identity channel $\mathcal{I}$ (which acts on qubits without CNOT connection) and $\mathcal{E}_{\mathrm{CNOT}}$ (which acts on qubits linked by CNOT connections).

Similar to layer 1, layer 3 also comprises single-qubit rotations only. Hence, we treat $\mathcal{E}_3(\cdot)$ as identity channel. Lastly, the fourth layer is subjected to SPAM noise which we adapt to a noise channel description [105]. It effectively handles measurement noise by corrupting the density matrix post execution but pre measurement, and then conducting noise-free projective measurements on the corrupted output. The SPAM noise channel for qubit q has two Kraus operators M_0 and M_1:

$$\begin{aligned}
M_0 =& \sqrt{f_q}\,|0\rangle\langle 0| + \sqrt{1-f_q}\,|1\rangle\langle 1| \\
M_1 =& \sqrt{1-f_q}\,|0\rangle\langle 0| + \sqrt{f_q}\,|1\rangle\langle 1|
\end{aligned} \tag{7.21}$$

where f_q represents the SPAM fidelity of qubit q. The probability of observing 0 is given by $\mathrm{Tr}[M_0^\dagger M_0 \rho]$ and the probability of observing 1 is given by $\mathrm{Tr}[M_1^\dagger M_1 \rho]$. Note that we have assumed a symmetric model for SPAM noise with f_q denoting the average SPAM fidelities for the initial states prepared as $|0\rangle$ and $|1\rangle$.

Neglecting inter-qubit cross-talk, we then have the noise channel representation for the last layer $\mathcal{E}_4$ as a separable SPAM noise channel:

$$\mathcal{E}_4(\cdot) = \left[\bigotimes_{q=0}^{n} \mathcal{E}_q^{\mathrm{SPAM}}\right](\cdot) \tag{7.22}$$

7.3.3 PEC coefficients under Pauli noise

We first discuss CNOT noise mitigation using PEC. Following that, we discuss single-qubit SPAM noise mitigation using PEC. Then, we integrate the two discussions for noise mitigation in the quantum circuit implementation of the Bernstein-Vazirani algorithm using PEC.

7.3.3.1 CNOT noise

Let $\tilde{\mathcal{G}_0}$ denote the super-operator for the noisy CNOT operation. Under the Pauli channel assumption, there are 16 basis super-operators, indexed by k:

$$\tilde{\mathcal{G}_k}\rho = (\mathbb{P}_C \otimes \mathbb{P}_T)\left(\tilde{\mathcal{G}_0}\rho\right)(\mathbb{P}_C \otimes \mathbb{P}_T) \tag{7.23}$$

where ρ is a 2-qubit density matrix, $\mathbb{P}_C, \mathbb{P}_T \in \{\mathbb{I}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}\}$ are the single-qubit Paulis acting on the control (C) and target (T) qubits respectively, and $k \in \{0, \cdots 15\}$. The index k is determined by the specific combination of $\mathbb{P}_C$ and $\mathbb{P}_T$, starting from $\mathbb{I} \otimes \mathbb{I}$ for $k = 0$ and ending with $\mathbb{Z} \otimes \mathbb{Z}$ for $k = 15$, incrementing k for each subsequent combination in the sequence $\mathbb{I}, \mathbb{X}, \mathbb{Y}, \mathbb{Z}$ applied to $\mathbb{P}_C$ and $\mathbb{P}_T$.

When $\mathbb{P}_C = \mathbb{P}_T = \mathbb{I}$, the circuit corresponds to the noisy super-operator $\tilde{\mathcal{G}_0}$ denoting the noisy CNOT gate available. An example of one of the remaining 15 PEC circuits for CNOT noise mitigation is shown in Fig. 7.3.

Using the requirement that the linear combination $\sum \eta_k \tilde{\mathcal{G}_k}$ should equal the noiseless CNOT operation, we derive the PEC coefficients as follows.

Let $c_0 = 3/[\mathrm{x}_T(1-\mathrm{x}_C)]^2 + [\mathrm{x}_C(1-\mathrm{x}_T)]^2 - 3c_1^2]$ and, $c_1 = \mathrm{x}_C + \mathrm{x}_T - \mathrm{x}_C\mathrm{x}_T - 1$.

When $\mathbb{P}_C = \mathbb{P}_T = \mathbb{I}$,

$$\eta_0 = c_0 c_1.$$

When $\mathbb{P}_C = \mathbb{I}$ and $\mathbb{P}_T \in \{\mathbb{X}, \mathbb{Y}, \mathbb{Z}\}$,

$$\eta_{\mathrm{k}} = c_0 \frac{\mathrm{x}_T(1 - \mathrm{x}_C)}{3}, \quad k \in \{1, 2, 3\}.$$

When $\mathbb{P}_T = \mathbb{I}$ and $\mathbb{P}_C \in \{\mathbb{X}, \mathbb{Y}, \mathbb{Z}\}$,

$$\eta_{\mathrm{k}} = c_0 \frac{\mathrm{x}_C(1 - \mathrm{x}_T)}{3}, \quad k \in \{4, 8, 12\}.$$

For all the remaining terms,

$$\eta_{\mathrm{k}} = c_0 \frac{\mathrm{x}_C \mathrm{x}_T}{9}, \quad k \in \{9, 10, 11, 13, 14, 15\}.$$

This yields the quasi-probability distribution as: $\{\mathbb{Q}_k\} = \{|\eta_0|/\gamma_{\mathrm{CNOT}}, \cdots, |\eta_{15}|/\gamma_{\mathrm{CNOT}}\}$ where $\gamma_{\mathrm{CNOT}} = |\eta_0| + \cdots + |\eta_{15}|$.

7.3.3.2 SPAM noise for qubit q

Consider a single-qubit SPAM noise channel for qubit q. The four noisy super-operators that can be implemented in this case are: (i) the SPAM noise channel: $\tilde{\mathcal{G}}_{\mathbb{I}}(\rho) = f_q\rho + (1 - f_q)\mathbb{X}\rho\mathbb{X}$, (ii) the SPAM noise channel followed by an X error: $\tilde{\mathcal{G}}_{\mathbb{X}}(\rho) = f_q\mathbb{X}\rho\mathbb{X} + (1 - f_q)\rho$, (iii) the SPAM noise channel followed by a Y error: $\tilde{\mathcal{G}}_{\mathbb{Y}}(\rho) = f_q\mathbb{Y}\rho\mathbb{Y} + (1 - f_q)\mathbb{Z}\rho\mathbb{Z}$ and, (iv) the SPAM noise channel followed by a Z error: $\tilde{\mathcal{G}}_{\mathbb{Z}}(\rho) = f_q\mathbb{Z}\rho\mathbb{Z} + (1 - f_q)\mathbb{Y}\rho\mathbb{Y}$.

Solving the linear equation:

$$\mathcal{I} = \eta_0\tilde{\mathcal{G}}_I + \eta_1\tilde{\mathcal{G}}_X + \eta_2\tilde{\mathcal{G}}_Y + \eta_3\tilde{\mathcal{G}}_Z \tag{7.24}$$

we get the quasi-probability distribution as: $\{|\eta_0|/\gamma_{\mathrm{SPAM}}, |\eta_1|/\gamma_{\mathrm{SPAM}}, 0, 0\}$ with

$$\begin{aligned} &\eta_0 = \frac{f_q}{2f_q - 1}, \quad \eta_1 = -\frac{1 - f_q}{2f_q - 1}, \quad \eta_2 = \eta_3 = 0 \\ &\mathrm{sgn}(\eta_0) = +1, \quad \mathrm{sgn}(\eta_1) = -1 \\ &\gamma_{\mathrm{SPAM}}(q) = |\eta_0| + |\eta_1| \end{aligned} \tag{7.25}$$

Estimating f_q therefore provides a complete specification of the PEC coefficients.

7.3.3.3 Composite noise in 5-qubit implementation of Bernstein-Vazirani algorithm

The two noisy basis super-operators for each of the 5 distinct SPAM noise channels and the 16 noisy basis super-operators for the CNOT noise channel leads to 512 (16×2^5) noisy basis circuits $\{\mathcal{G}_k\}$ for the 5-qubit Bernstein-Vazirani circuit, where k runs from 0 to 511.

Under the SPAM noise separability assumption, γ is obtained as:

$$\gamma = \gamma_{CNOT} \prod_{q=0}^{4} \gamma_{SPAM}(q) \tag{7.26}$$

where $\gamma_{SPAM}(q)$ refers to the γ_{SPAM} for the q-th qubit. The PEC coefficients for the Bernstein-Vazirani circuit are the elements of:

$$\left\{\eta_0^{\text{CNOT}}, \cdots, \eta_{15}^{\text{CNOT}}\right\} \times \prod_{q=0}^{4} \left\{\eta_0^{\text{SPAM}}(q), \eta_1^{\text{SPAM}}(q)\right\} \tag{7.27}$$

The quasi-probability distribution then follows as: $\{\mathbb{Q}_k\} = \left\{|\eta_{\text{k}}^{\text{BV}}|/\gamma\right\}$.

7.3.4 Adaptive noise model

In the context of our 5-qubit Bernstein-Vazirani setup, the noise parameters f_0, f_1, and f_2 are estimated using the method for uncorrelated parameters as they do not have CNOT correlations. The adaptive estimation of parameters f_3, f_4, x_C, and x_T employs the method for handling correlated parameters, detailed in Sec. 7.2. The process is initiated within the Bayesian inference framework, which considers the probabilities of observing outcomes 00, 01, 10, and 11 on qubits 3 and 4, as depicted in Fig. 7.2, to be random variables. The estimation of the time-varying means and variances of these correlated noise parameters ($f_3, f_4, \text{x}_C, \text{x}_T$) is achieved by associating the mean values of the estimated densities directly with the parameters of the correlated noise model using:

$$\Pr(i) = \text{Tr}\left[\Pi_i \mathcal{E}_4(\mathbb{L}_3 \mathcal{E}_2(\mathbb{L}_2 \mathbb{L}_1 \rho \mathbb{L}_1^\dagger \mathbb{L}_2^\dagger) \mathbb{L}_3^\dagger)\right] \tag{7.28}$$

where $\Pi_i = |i\rangle\langle i|$ are the projection operators and $i \in \{00, 01, 10, 11\}$. The probabilities are given by:

$$\begin{aligned} \Pr(00) =& f_3 f_4(-1 - \text{x}_C\text{x}_T + \text{x}_C + \text{x}_T) \\ &+ f_3(\text{x}_C\text{x}_T/2 - \text{x}_T/2) \\ &+ f_4(1 - \text{x}_T - \text{x}_C/2 + \text{x}_C\text{x}_T/2) \\ &- \text{x}_C\text{x}_T/4 + \text{x}_T/2 \end{aligned} \tag{7.29}$$

$$\begin{aligned}\Pr(01) =& f_3 f_4(1 - \mathrm{x}_C - \mathrm{x}_T + \mathrm{x}_C\mathrm{x}_T) \\ &+ f_3(-1 - \mathrm{x}_C\mathrm{x}_T/2 + \mathrm{x}_C - \mathrm{x}_T/2) \\ &+ f_4(-1 - \mathrm{x}_C\mathrm{x}_T/2 + \mathrm{x}_C + \mathrm{x}_T) \\ &+ 1 + \mathrm{x}_C\mathrm{x}_T/4 - \mathrm{x}_C/2 - \mathrm{x}_T/2\end{aligned} \tag{7.30}$$

$$\begin{aligned}\Pr(10) =& f_3 f_4(1 - \mathrm{x}_T - \mathrm{x}_C + \mathrm{x}_C\mathrm{x}_T) \\ &+ f_3(\mathrm{x}_T/2 - \mathrm{x}_C\mathrm{x}_T/2) \\ &+ f_4(\mathrm{x}_C/2 - \mathrm{x}_C\mathrm{x}_T/2) \\ &+ \mathrm{x}_C\mathrm{x}_T/4\end{aligned} \tag{7.31}$$

$$\begin{aligned}\Pr(11) =& f_3 f_4(-1 + \mathrm{x}_C + \mathrm{x}_T - \mathrm{x}_C\mathrm{x}_T) \\ &+ f_3(1 - \mathrm{x}_T/2 - \mathrm{x}_C + \mathrm{x}_C\mathrm{x}_T/2) \\ &+ f_4(\mathrm{x}_C\mathrm{x}_T/2 - \mathrm{x}_C/2) \\ &+ \mathrm{x}_C/2 - \mathrm{x}_C\mathrm{x}_T/4\end{aligned} \tag{7.32}$$

In the last step, the updated PEC coefficients are obtained using Eqn. 7.27.

7.3.5 Numerical simulation

For our 5-qubit circuit implementing the Bernstein-Vazirani algorithm, we used secret string $r =$ "1000". Thus, the qubit-wise mean of the $\mathbb{Z}$ observable for the noiseless case is given by: $\langle\mathcal{O}_0\rangle = +1, \langle\mathcal{O}_1\rangle = +1, \langle\mathcal{O}_2\rangle = +1, \langle\mathcal{O}_3\rangle = -1, \langle\mathcal{O}_4\rangle = +1$, qubit 4 being the ancilla.

To validate our method, we used a numerical experiment that conducts a density matrix simulation of a 5-qubit noisy quantum circuit implementing the Bernstein-Vazirani algorithm using the Qiskit [23] software. The simulation begins with the mean of the beta distributions characterizing the SPAM fidelities for qubits 0-4 set at $0.96, 0.95, 0.94, 0.93, 0.92$, respectively, and the mean of the depolarizing channel parameters for the control and target qubits in the CNOT gate both fixed at 0.017. Over the course of ten simulated time periods, the average SPAM fidelity for each qubit decreased by 0.01 per period, resulting in final mean SPAM fidelities of $0.86, 0.85, 0.84, 0.83, 0.82$ for qubits 0-4, respectively. Similarly, the average depolarizing parameter for the CNOT gate also declined by 0.01 per time

period, leading to a final mean value of 0.117 for both control and target qubits by the simulation's end. The noise parameters were adaptively estimated, per the methodology described in previous sections, using data generated by executing each of the 512 PEC circuits using $10,000$ shots.

Fig. 7.5 demonstrates the simulated efficacy of the adaptive PEC algorithm. The plot compares the register accuracy and stability (averaged over the 5 qubits) achieved with the adaptive approach against a non-adaptive approach. It shows an improvement in accuracy of 59.5% in the final time period, when device noise is at its peak, with an average accuracy improvement of 53.4% across all ten periods. Similarly, stability improved by 58.0% in the last period, and by an average of 51.5% over the entire span of ten periods. The improvement in accuracy and stability of the outcomes from adaptive PEC occur due to more accurate noise characterizations using Bayesian inference.

7.4 Experimental Testing

We tested the adaptive PEC method on the 27-qubit superconducting device called ibm_kolkata. Qubits 0,1,2,3,4 in Fig. 7.2 map to physical qubits 0,1,2,3,5 on the device shown in Fig. 7.6. The CNOT gate is between the physical qubits 3 (control) and 5 (target) on ibm_kolkata. Our dataset spans 24 hours and comprises 13 complete PEC datasets. It was collected on January 15, 2024, and have the following time-stamps: 00:01 hrs, 02:03 hrs, 03:50 hrs, 05:43 hrs, 07:23 hrs, 09:11 hrs, 10:53 hrs, 12:39 hrs, 14:22 hrs, 16:09 hrs, 17:50 hrs, 19:33 hrs, and 21:17 hrs. Each dataset is derived from measurements made in the computational basis, with each observation being a 5-bit string. The observations were obtained from the 512 noisy basis circuits, as described in Sec. 7.3.3, with each circuit repeated using $L = 10,000$ shots, resulting in approximately 67 million observations in total.

7.4.1 Non-stationary noise estimates

Fig. 7.7 illustrates how noise in the quantum computer changed over time. The blue line in plot (a) shows the depolarizing parameter for the target qubit of the CNOT gate, while the black line represents the depolarizing parameter for the control qubit. Plot (b) shows five lines, each representing the SPAM fidelity for the register elements. The x-axis denotes intra-calibration timestamps.

The graph shows periods where the noise levels in the depolarizing parameter for qubit 3 (the control qubit) are steady, notably between 2:00 am and 7:30 am, contrasting with times of significant fluctuation, as observed between 9:00 am and 2:30 pm. The depolarizing parameter fluctuates between 1% and 4% for qubit 4 (the target qubit) and between 1% and 3% for qubit 3, both peaking sharply at 10:53 am.

Qubit 4 experiences the most significant impact from SPAM noise, with its values fluctuating between 0.99 and 0.94, a notable range given the sensitivity of PEC to accurate noise estimations. In contrast, Qubit 3 maintains a consistent SPAM fidelity throughout the same period, indicated by small error margins and a steady average value of 0.99. Meanwhile, Qubit 2 demonstrates a gradual drift in its values, starting from below 0.94 and rising to 0.96. This progressive change suggests a systematic, non-random trend that might be rectifiable with bias shift corrections. However, such patterns are not uniform across the entire register, implying the necessity to consider non-stationary statistics for modeling the system. Conducting experiments in times of significant non-stationary activity, like from 7:30 am to 2:30 pm, lead to more unstable outcomes when using non-adaptive PEC.

7.4.2 Non-stationary quasi-probability distribution

Fig. 7.8 underscores the importance of considering the non-stationary nature of the quasi-probability distribution when implementing PEC, especially given the lengthy data collection process required for a single PEC mitigation (approximately 2 hours in our example). The abrupt change observed at 12:39 p.m. in the quasi-probability distribution directly correlates with the sharp change in the noise parameters characterizing the quantum circuit at the same time, as illustrated in Fig. 7.7.

To maintain clarity in Fig. 7.8, we have not plotted all 512 bins of the histogram in one plot. Fig. 7.8 (a) displays the time-varying weight $\mathbb{Q}_0$ for the Bernstein-Vazirani circuit as-is without any additional Pauli-gate added (we also refer to this as the raw Bernstein-Vazirani circuit). If the circuit were noiseless, then the weight for the raw circuit will be a constant 1. We observe a decrease in weight for the raw circuit, dipping below 78% around 10:53 am from a peak of almost 83%, coinciding with a peak in circuit noise as seen in Fig. 7.7. This decrease in weight is expected as the circuit noise peaks.

The values of the quasi-probability bins for the next 10 basis circuits are shown in plot (b), with values approximately 10 times lower than the first circuit. Ignoring seemingly small coefficients in the quasi-probability distribution without considering the precision of final reported results can be risky. Subsequent basis circuits, not shown here, have significantly smaller quasi-probability weights (around 10^{-4}). Yet their collective impact in a sum of 500 can be substantial, contributing up to 0.05. Given our reported accuracy and stability are around 10^{-2}, these coefficients, though small, can significantly influence the results. Therefore, in our analysis, we included all basis circuits without approximation, focusing on the effects of non-stationary noise on PEC, rather than on resource optimization.

7.4.3 Non-stationary PEC outcomes

The impact of the non-stationary noise can be seen in Fig. 7.9. The first plot, labeled No mitigation, presents the accuracy and stability metrics defined in Eqns. 7.16 and 7.19 respectively, for the raw Bernstein Vazirani circuit without any form of quantum error mitigation. The second plot, labeled ROEM, which stands for readout error mitigation, displays the metrics after performing SPAM noise mitigation. In this case, the SPAM noise parameters are held constant after initial device characterization. It deploys the standard matrix inversion [28] technique for mitigation. The third plot, labeled non-adaptive PEC, exhibits the accuracy and stability metrics for the Bernstein Vazirani circuit with non-adaptive PEC. This method incorporates SPAM noise mitigation within the PEC framework, as detailed in Sec. 7.3.3. The fourth plot, labeled adaptive PEC, presents the metrics for the adaptive PEC method, as discussed in Sec. 7.2.

Fig. 7.9 (a) shows the effectiveness of the adaptive PEC in enhancing result accuracy. It reveals approximately a 42% improvement in accuracy on average compared to the non-adaptive method. The observed accuracy benefit ranges from a minimum of 25% to a maximum of 78%. Fig. 7.9 (b) shows the impact on result stabilization. It shows an approximately 60% enhancement in stability on average compared to the non-adaptive method. The observed stability benefit ranges from a minimum of 8% to a maximum of 200%.

Observing the plots, it is evident that adaptive PEC significantly outperforms standard PEC. Additionally, all four methods (no mitigation, ROEM, PEC, and adaptive PEC) exhibit a time-series trend that deteriorates notably at 10:53 am. The observation correlates with the abrupt change in the underlying quasi-probability distribution at 10:53 am as seen in Fig. 7.8.

Both the accuracy and stability metrics at 12:39 p.m. are slightly worse for adaptive PEC compared to non-adaptive PEC. This discrepancy stands out as the only instance where adaptive PEC performs sub-optimally. It seems, surprisingly, that the stale data serves as a better noise estimate for this specific time-point. However, the noise at 12:39 p.m. is not necessarily closer to the noise at the starting time-stamp of 00:01 a.m., as demonstrated in Fig. 7.7. While adaptive PEC succeeds in most cases, the learning process is not instantaneous. There exists a slight lag in learning due to the influence of prior information on the final estimate. This phenomenon reflects a fundamental aspect of learning methods: the presence of memory, which can aid learning but also slows down adaptation to the fast, spiky changes. The sharp change in noise at 10:53 a.m. leads to an overestimation of the noise estimate compared to when using the initial value at 00:01 a.m. as a reference point. Residual errors in the noise estimates likely arises from inaccurate models and non-stationary processes changing at a faster rate than sampled here.

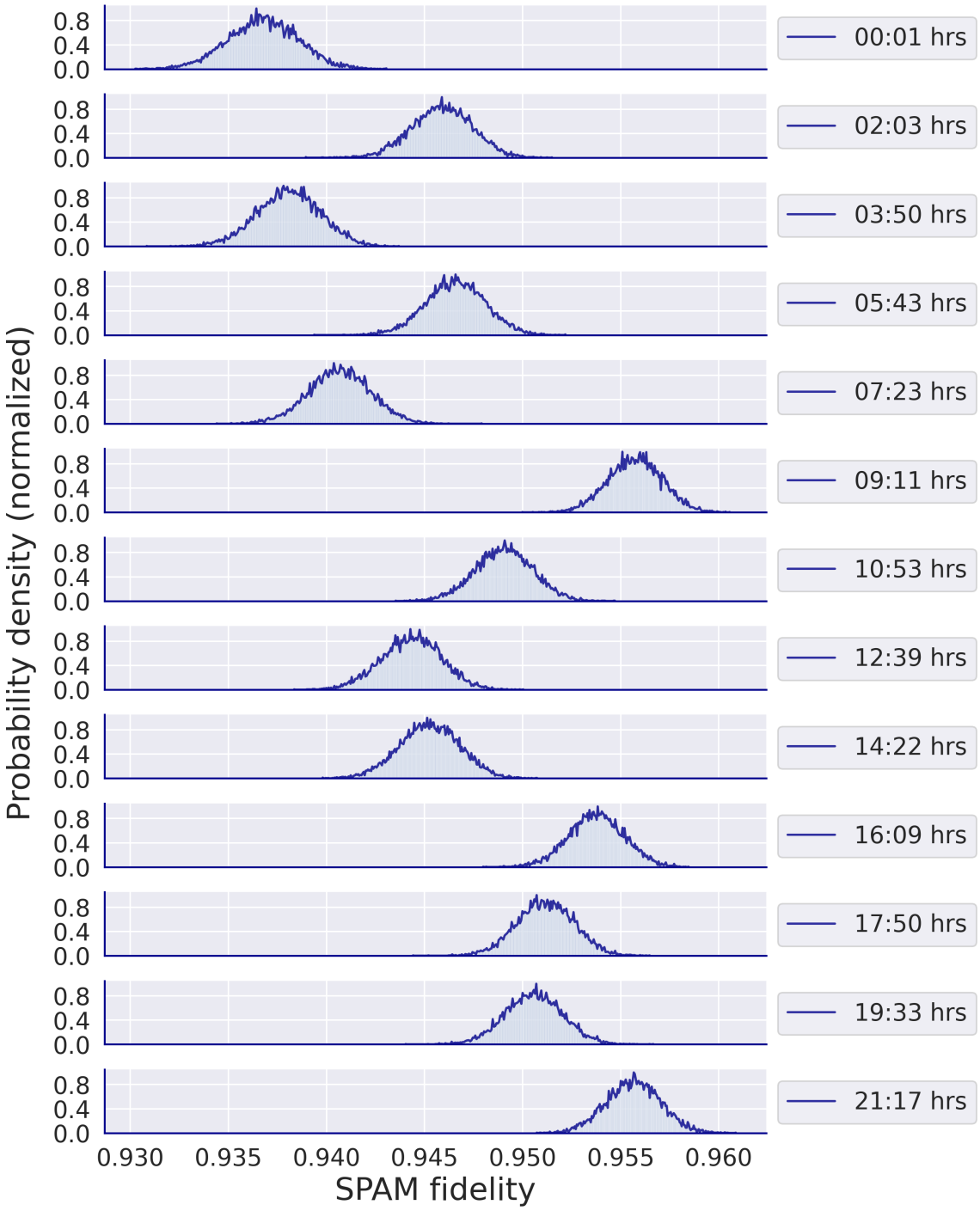


Figure 7.1: Non-stationary distribution functions of the state preparation and measurement (SPAM) fidelity for qubit 2 on ibm_kolkata superconducting device collected on Jan 15, 2024.

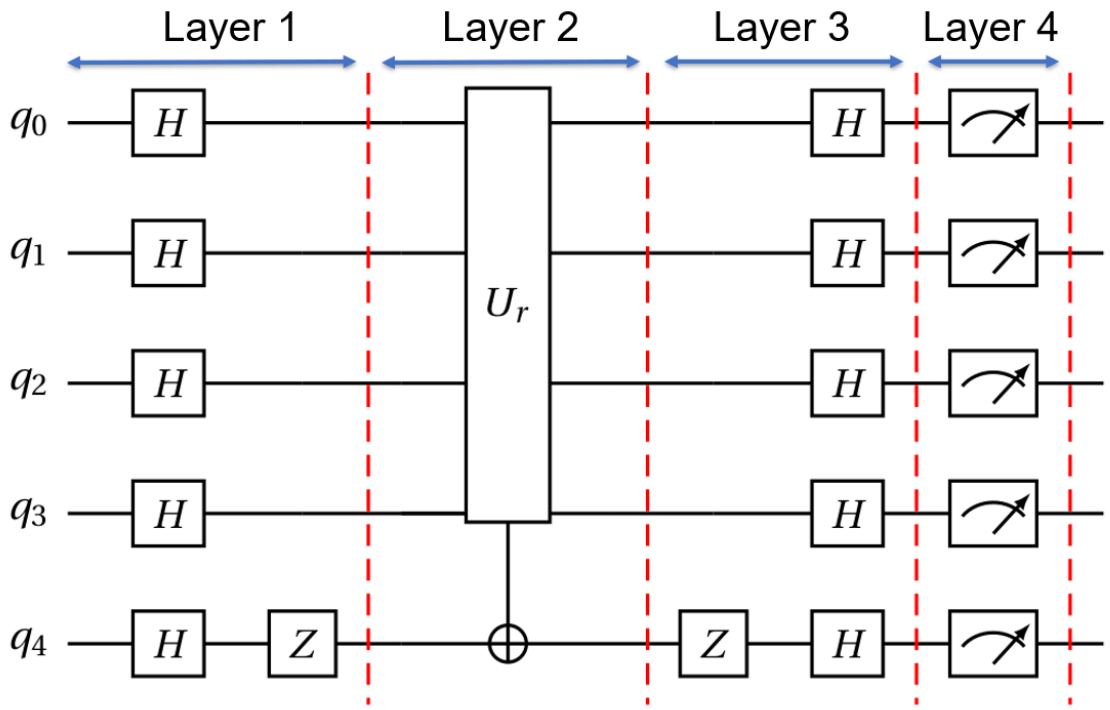


Figure 7.2: A 5-qubit implementation of the Bernstein-Vazirani algorithm with secret bit-string r.

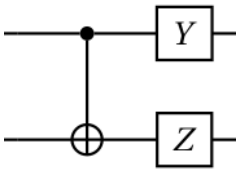


Figure 7.3: Circuit diagram for implementing one of the 16 noisy super-operators for CNOT noise mitigation using PEC. The initial noisy CNOT gate is followed up by $\mathbb{Y}$ and $\mathbb{Z}$ gates on the control and target qubits respectively.

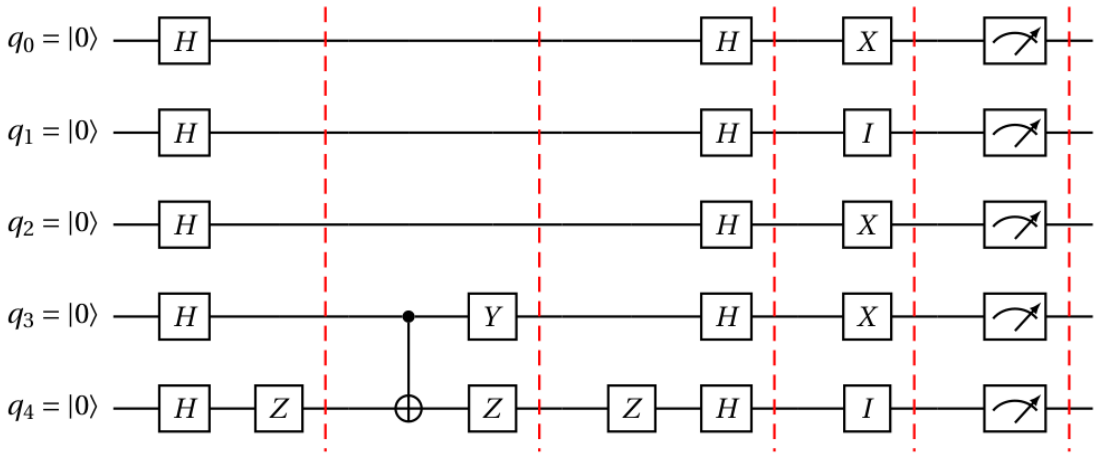
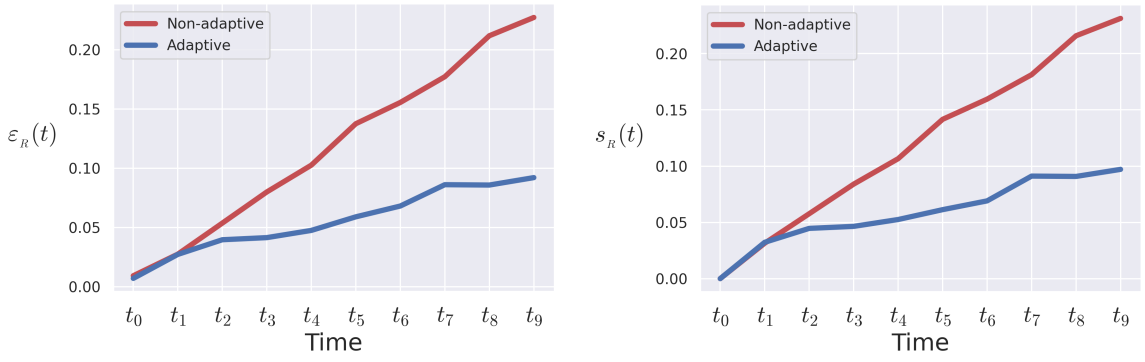


Figure 7.4: One of the 512 noisy basis circuits for mitigation using PEC in the 5-qubit implementation of the Bernstein-Vazirani circuit. The initial noisy CNOT gate is followed up with $\mathbb{Y}$ and $\mathbb{Z}$ gates on the control and target qubits respectively. The readout lines for qubit $0, 1, 2, 3, 4$, and 5 are subjected to the Pauli gates $\mathbb{X}, \mathbb{I}, \mathbb{X}, \mathbb{X}, \mathbb{I}$ respetively, prior to measurement, in this specific noisy basis circuit for PEC.



(a) (b)

Figure 7.5: Comparison of the adaptive PEC algorithm with a non-adaptive approach for a density matrix simulation that implements a noisy 5-qubit circuit for solving the Bernstein Vazirani problem. (a) Average accuracy across the 5-qubits and (b) Average stability across the 5-qubits.

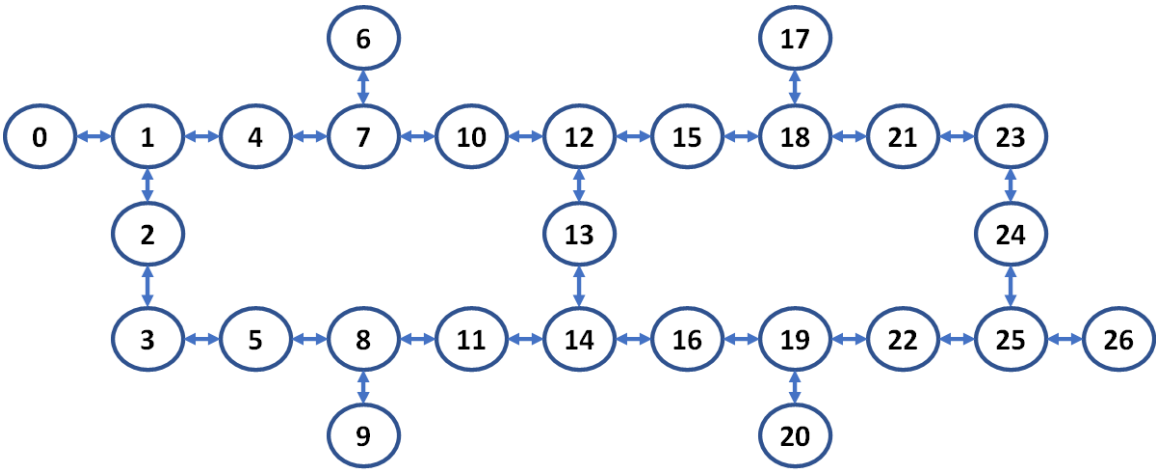
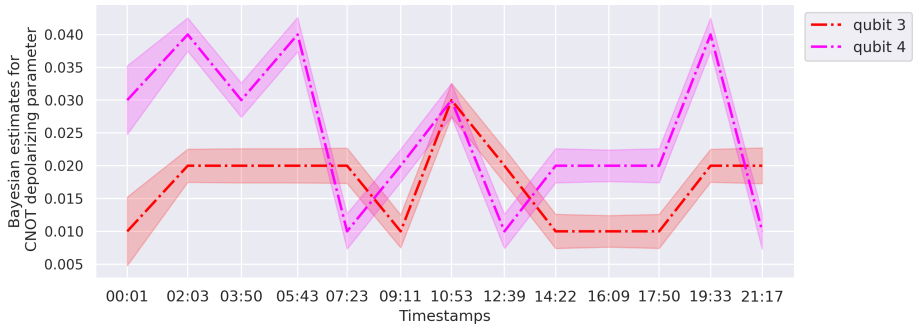
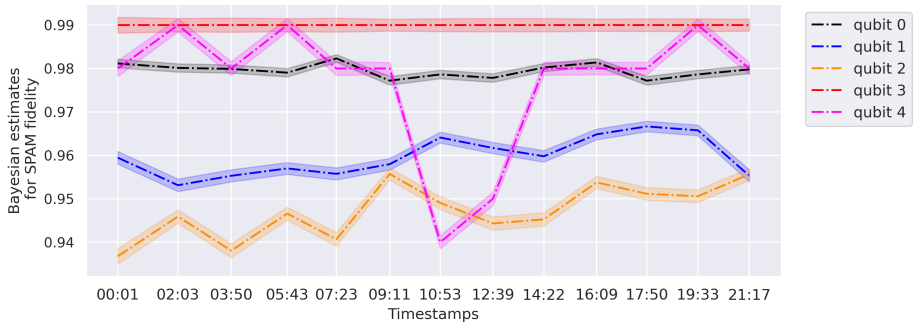


Figure 7.6: Qubit layout of the 27-qubit superconducting device ibm_kolkata. Circles represent superconducting transmons and lines indicate possible gate operations between sites.

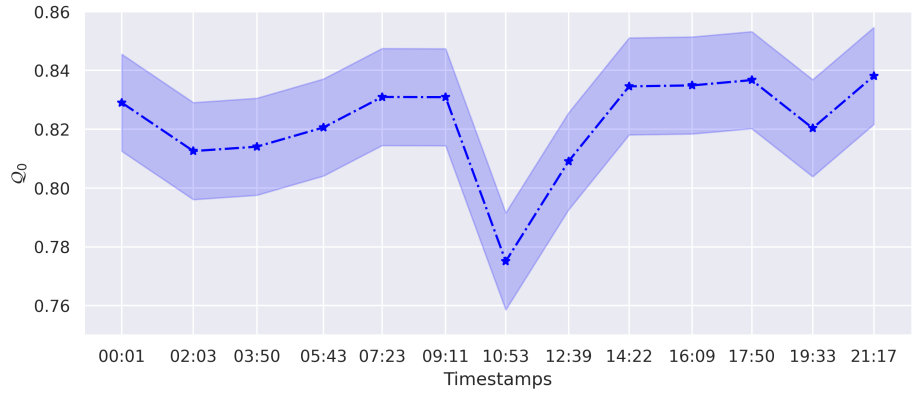


(a)



(b)

Figure 7.7: This figure depicts the non-stationary noise on the experimental device ibm_kolkata. In plot (a), the blue line represents the depolarizing parameter for the target qubit, while the black line denotes the depolarizing parameter for the control qubit in the CNOT gate. Plot (b) illustrates five lines, each indicating the SPAM fidelity for the register elements. The x-axis corresponds to intra-calibration timestamps for January 15. The shaded regions denote the time-varying standard deviations.

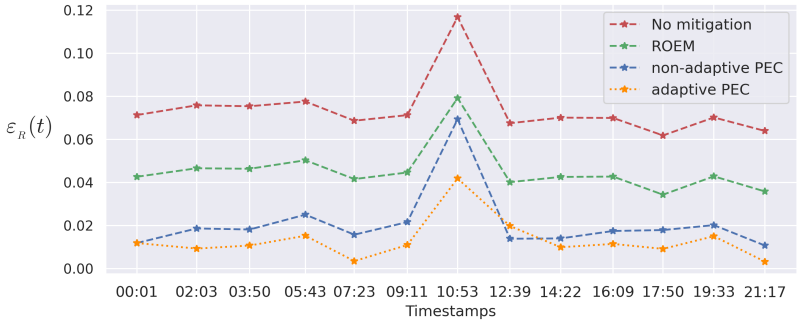


(a)

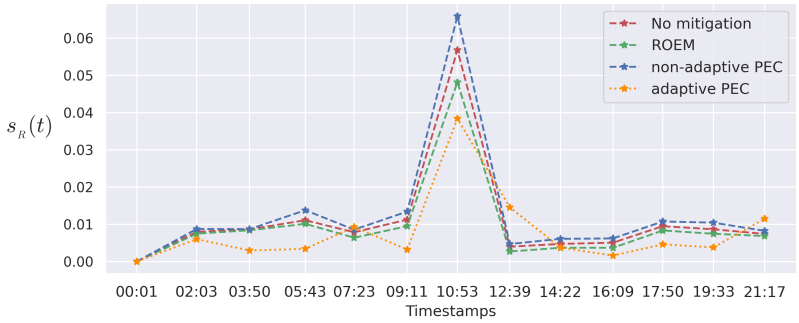


(b)

Figure 7.8: These plots depict the non-stationary nature of the quasi-probability distribution. To maintain clarity, we opted not to plot all 512 bins of the histogram in a single plot. In (a), we observe the time-varying weight $\mathbb{Q}_0$ for the raw Bernstein-Vazirani circuit which carries the most substantial weight, accounting for almost 80% of the distribution. In (b), we show the values of the quasi-probability bins for the subsequent 10 basis circuits, which are more than 10 times lower in magnitude compared to the first circuit. The non-stationary nature of the quasi-probability distribution becomes crucial given the lengthy data collection process required for PEC mitigation because the noise estimation becomes inaccurate in these time-frames. Our experiment took approximately 2 hours for each dataset comprising 512 circuits.



(a)



(b)

Figure 7.9: The figure comprises two plots, each depicting four graphs: (1) "No mitigation" presents raw Bernstein-Vazirani metrics without error mitigation, (2) "ROEM" shows metrics after readout error mitigation, with constant SPAM noise parameters, (3) "Non-adaptive PEC" displays metrics for PEC, and (4) "Adaptive PEC" exhibits metrics for PEC with adaptive noise mitigation. The x-axis denotes intra-calibration time-stamps (UTC) for Jan 15. Plot (a) illustrates the time-varying accuracy metric from Eqn. 7.16. It demonstrates the adaptive PEC's 42% accuracy improvement over non-adaptive PEC. Plot (b) shows the time-varying stability metric from Eqn. 7.19. It illustrates the adaptive PEC's 60% stability enhancement compared to non-adaptive PEC. These plots underscore the significant impact of non-stationary noise on PEC resilience. Due to PEC's lengthy completion time (a couple of hours), adaptive methods are able to handle non-stationary noise conditions better.

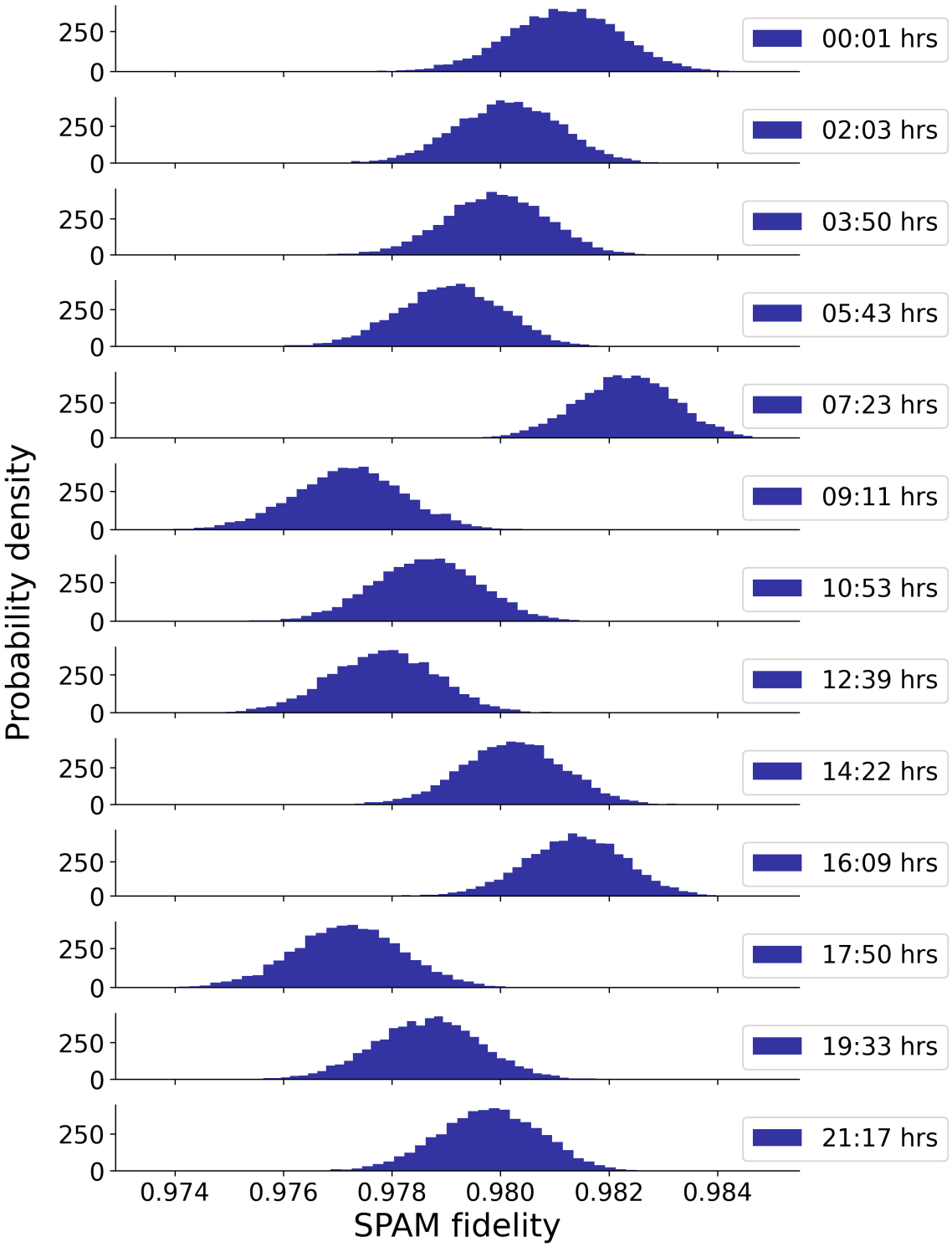


Figure 7.10: Non-stationary distribution function of SPAM fidelity for qubit 0 of ibm_kolkata as observed on Jan 15, 2024.

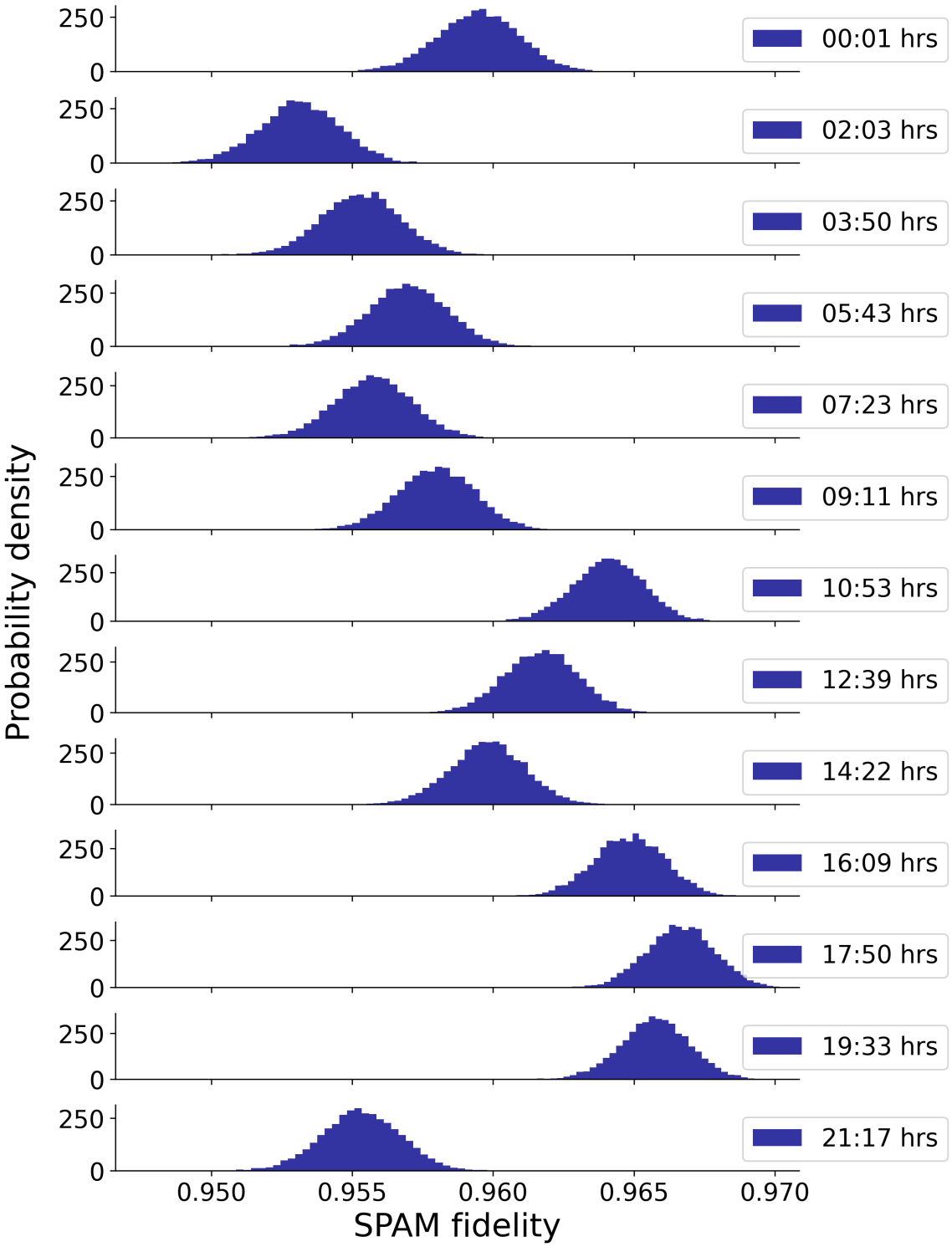


Figure 7.11: Non-stationary distribution function of SPAM fidelity for qubit 1 of ibm_kolkata as observed on Jan 15, 2024.

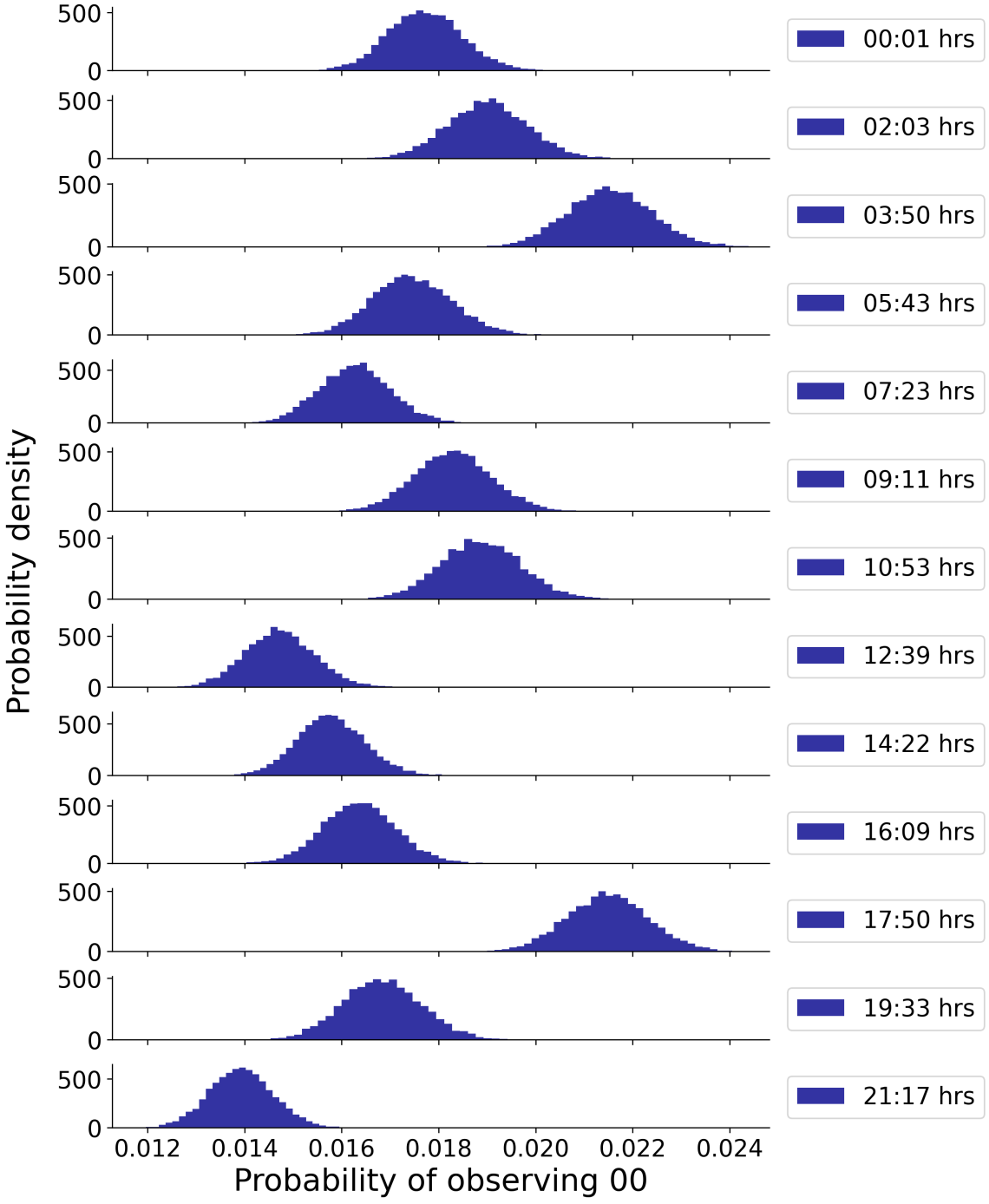


Figure 7.12: Time-varying density for probability of observing '00' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024.

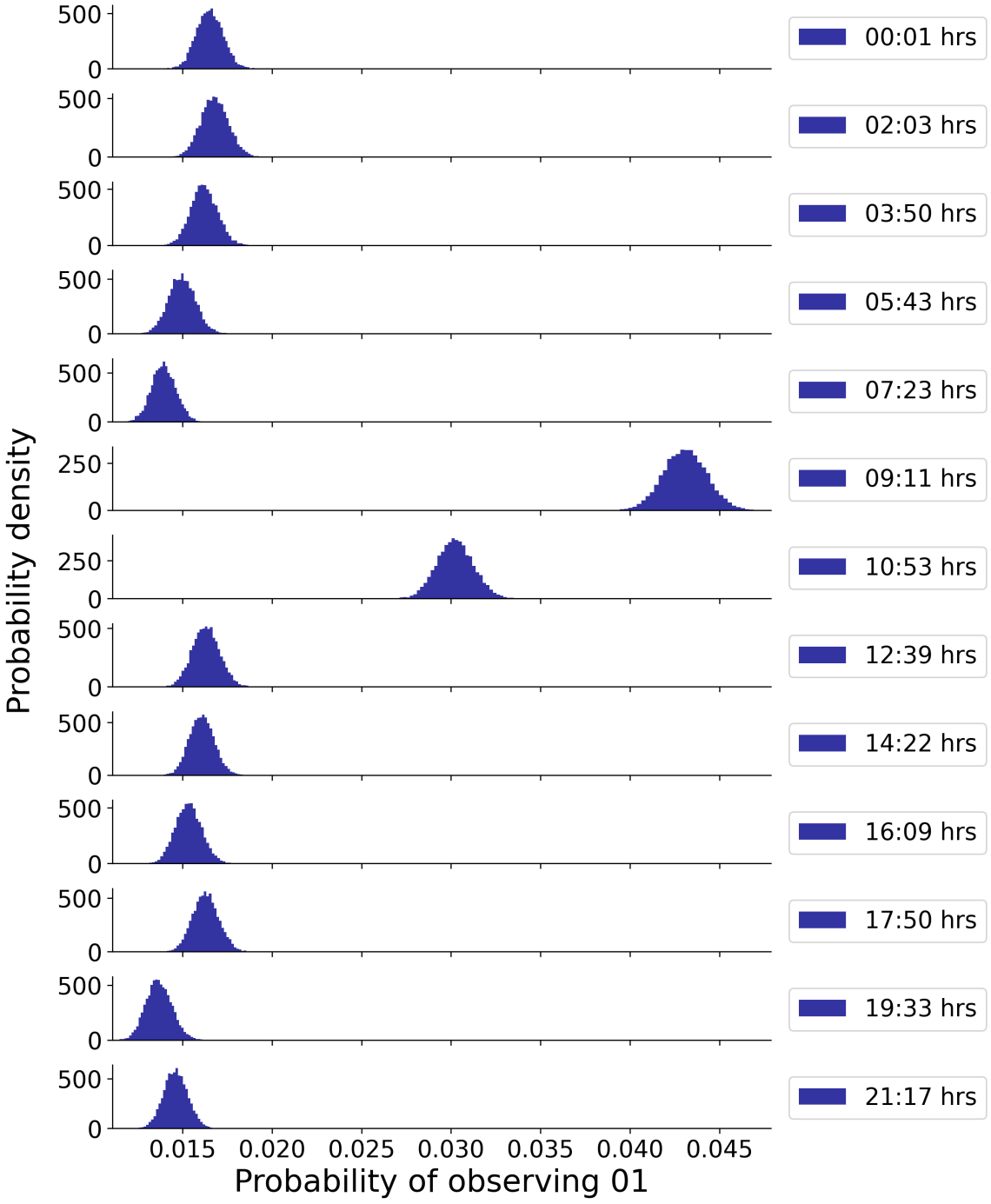


Figure 7.13: Time-varying density for probability of observing '01' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024.

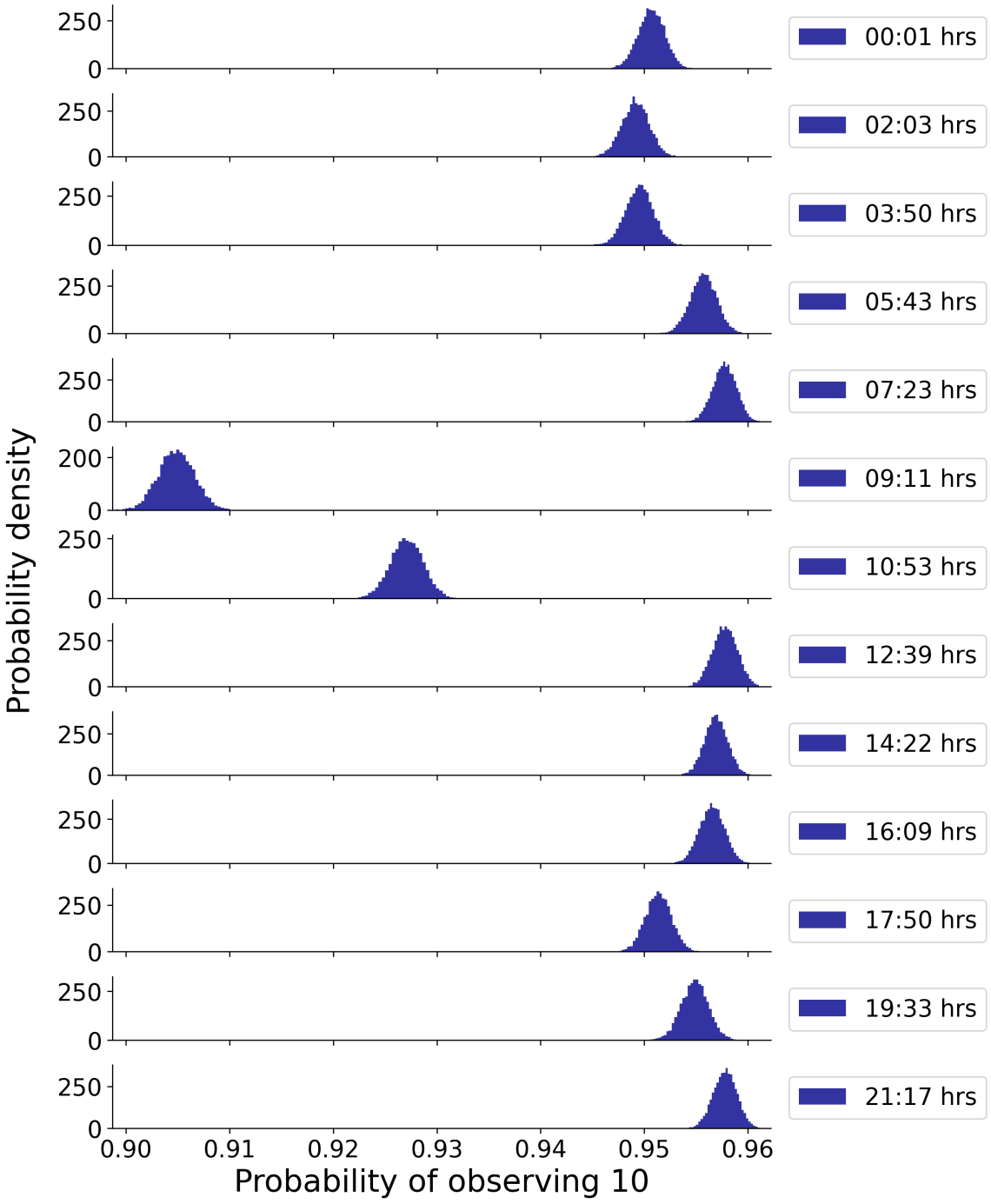


Figure 7.14: Time-varying density for probability of observing '10' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024.

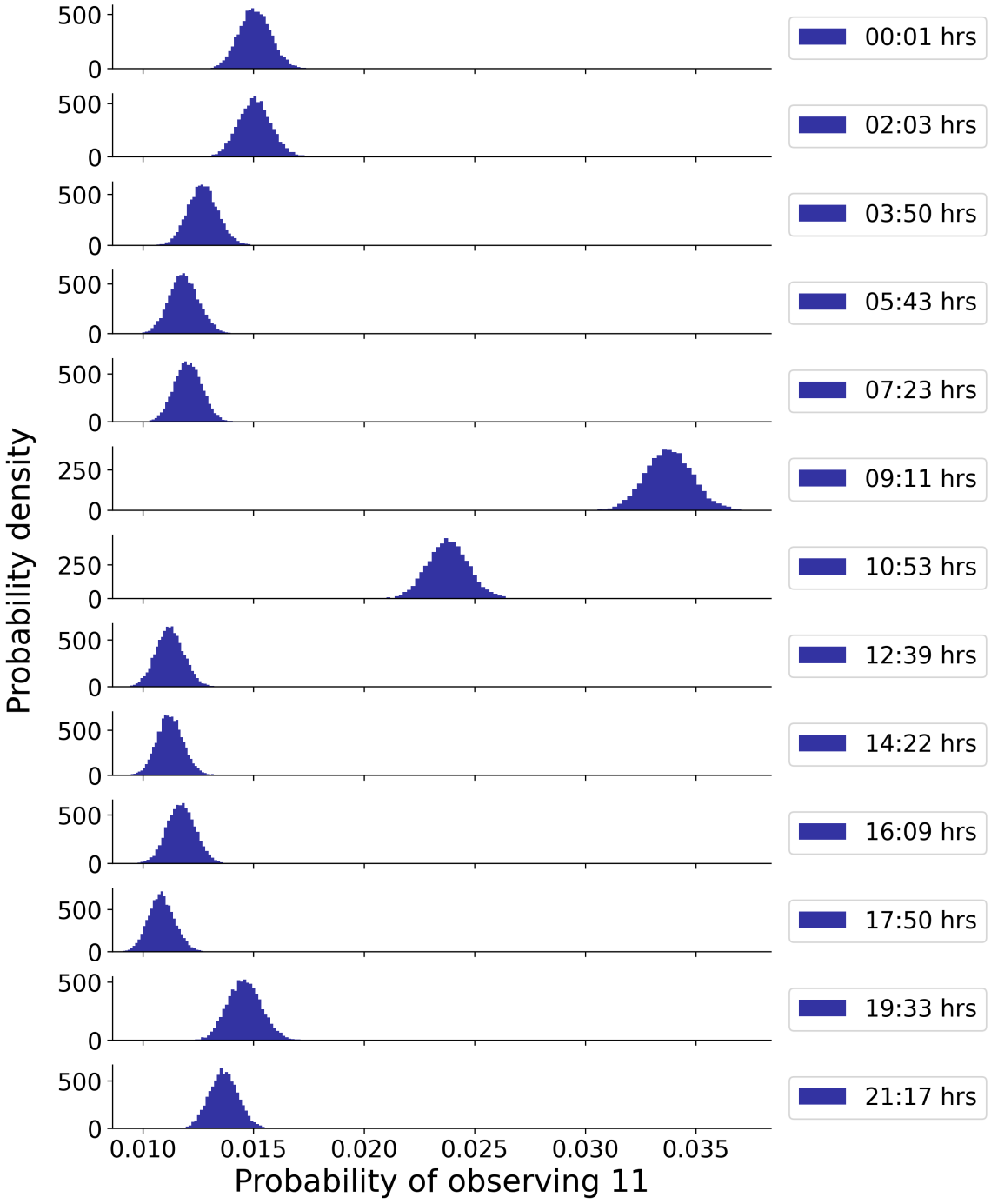


Figure 7.15: Time-varying density for probability of observing '11' on qubits 3 and 4 of ibm_kolkata, for the Bernstein-Vazirani circuit, as observed on Jan 15, 2024.

Chapter 8

Conclusion

Quantum computing's tremendous potential [1] is curtailed by noise. Characterizing noise in contemporary quantum computers remains challenging due to its non-stationary statistics. The non-stationarity manifests across time, across different parts of the chip, and across devices. It impacts the verification and validation of quantum computing demonstrations and adversely impacts error mitigation strategies [59]. This hinders the production of trustworthy results.

This dissertation focused on the non-stationarity [46, 42, 36] of noise in superconducting processors [30]. It established a framework comprising computational accuracy, device reliability, outcome stability, and result reproducibility for assessing noisy outcomes. It determined upper and lower bounds for the performance metrics, in terms of the Hellinger distance between time-varying noise densities. It demonstrated that if the noise stays within the theoretical bounds, outcome stability can be ensured with high confidence. It developed a Bayesian strategy to improve the stability and accuracy of results obtained from probabilistic error cancellation. Refinement in noise model selection, optimization algorithms, and data collection frequencies can yield further improvements and is a rich space to explore further.

In particular, chapter 2 discussed noise in quantum computing and experimentally analysed the decoherence of superconducting qubits over a 24-hour period on September 12, 2023. The time-varying decoherence characterization was connected to the coefficients of a Pauli channel model.

While the majority of quantum computing research has focused on achieving accuracy [54, 71, 68], limited attention has been given to reproducibility, reliability, and stability. Chapter 3 precisely defined these terms (computational accuracy, result reproducibility, device reliability and stability of error mitigated outcomes) for assessing the performance of noisy, quantum computers. Chapter 4

evaluated hardware reliability by analyzing experimental data across various time scales (monthly, daily, hourly, and seconds). It quantified the degree of non-stationarity in SPAM fidelity, CNOT fidelity, duty cycle, and addressability. It developed and validated a method to examine holistic reliability using the method of copulas [124].

Chapter 5 developed bounds on a proxy parameter to encapsulate the array of parameters characterizing a noisy device. Such a proxy parameter is valuable for streamlining high-dimensional noise analysis. Experimental validation of the theoretical bound was performed using a 27-qubit superconducting device. The chapter further illustrated how to determine the minimum sample size to achieve reproducibility with $1 - \delta$ confidence. It established bounds on device reliability necessary for achieving an ϵ-stable outcome.

Chapter 6 applied adaptive techniques [73, 74, 75, 76] to enhance accuracy of histograms, demonstrating a reduction in the Hellinger distance from 15% to 3.1%, while chapter 7 examined the impact of non-stationary noise on probabilistic error cancellation. It introduced a Bayesian approach to improve stability and accuracy of PEC outcomes. The algorithm was tested on the ibm_kolkata device on January 15, 2024. The dataset covered a 24-hour period, consisting of 13 complete PEC datasets for the Bernstein-Vazirani test circuit, with approximately 67 million observations. Results indicated a 42% increase in accuracy and a 60% enhancement in stability. Consistent improvement trends across time-stamps and qubits demonstrated the effectiveness of the algorithm.

The improvements in accuracy and stability of adaptive PEC vs non-adaptive PEC, across time and qubits, presented in this dissertation, indicate that the noise was well-characterized. In general, the choice of noise model influences the noise estimation process. If the chosen noise model is incorrect or insufficiently granular, the estimated noise parameters will reflect the specific dataset's patterns rather than accurately representing device noise. Mitigation using such mis-estimated noise models will be unsuccessful because the mitigation algorithm will introduce additional errors to the already noisy data. However, scalability becomes an issue with the Bayesian approach in the presence of correlations. While the method can accurately estimate numerous noise parameters in the absence of correlations, scalability diminishes when dealing with highly correlated parameters due to the need to estimate joint distributions and perform Monte Carlo simulations for maximum-a-posterior optimization.

Before concluding, I offer a perspective on the significance of this dissertation. Firstly, this work delves into the intricate nature of non-stationary noise in contemporary quantum computers. It sheds light on strategies for improving the reproducibility and stability of noisy quantum computations, by incorporating adaptive processes to manage non-stationary noise.

Quantum error correction (QEC) methods represent a formal way to mitigate and manage errors but in practice need to be tailored to the noise. This is similar to 5G networks where low density parity check (LDPC) [166, 167] codes are employed to meet the high data transfer demands. These codes use a quasi-cyclic structure that adapt their encoding and decoding graph size depending on prevailing error rates. The wireless channel noise is often analyzed using methods that take into account the variability in space, time, and frequency when using correlation functions and power spectrum densities. In parallel, when dealing with quantum noise channels, this work adopted a strategy that incorporates time-varying correlated distributions to reflect the non-stationary nature of quantum noise and utilized quasi-probability distributions (in probabilistic error cancellation) whose weights are influenced by the strength of the prevailing noise.

This dissertation can also help draw parallels to classical cellular communications in handling non-stationary noise. Just as non-stationary channel conditions in cellular networks once caused frustrating service quality issues due to inadequate noise characterization, quantum computing faces similar issues with error mitigation fluctuating due to poor noise characterization. Moreover, the concept of temporal decoherence, caused by interactions with the environment, can be likened to the phenomenon of signal fading in classical cellular systems due to atmospheric interactions. Both provide case studies for how non-stationary processes can impact the reliability of information transmission. Lastly, just like early cellular communication systems suffered high error-rates from poor multiplexing capabilities due to limited spectrum resources, similarly, contemporary quantum computing systems are limited in their error mitigation and correction capabilities due to constraints in qubit resources.

The unpredictable and rapidly changing nature of non-stationary noise presents a challenge for fault-tolerant quantum computing. This type of noise can lead to error patterns that change more quickly than contemporary error correction codes can adjust, making it difficult to maintain fault tolerance. For instance, cosmic rays can cause sudden and sporadic error bursts, temporarily pushing error rates beyond the limits that quantum error correction codes are designed to handle. This can result in uncorrected errors and, potentially, the failure of logical qubits. Additionally, superconducting processors exhibit varying error rates across the chip, which challenges the assumption of a uniform threshold error rate, as often applied in techniques like the surface code. Such issues of non-stationary temporal and spatial error rates affecting error correction have previously been encountered in the field of cellular communications, indicating that the challenges faced by quantum computing are not entirely unprecedented. Note that non-stationary noise does not prevent the application of fault-tolerant quantum error correction but needs a tailored approach.

The results from this study can also help in noise modeling in quantum communication scenarios involving entanglement distribution over long distances as the quality of entanglement can be degraded by non-stationary atmospheric conditions when using free-space optical channels. These conditions can change rapidly due to weather phenomena, causing fluctuations in entanglement fidelity and compromising the integrity of information transfer.

Lastly, this dissertation underscores the need for interdisciplinary collaboration to advance the field of quantum computing. For example, it stresses the role of reliability engineering in improving manufacturing processes for quantum chips to ensure high standards of quality. It urges software developers to develop adaptive algorithms to tackle real-time challenges posed by non-stationary noise in quantum systems. It calls for the utilization of advanced statistical models and database management techniques by data scientists to aid quantum physicists using models should embrace the inherent quantum noise model uncertainties as a feature rather than a flaw. And, it challenges information theorists to extend their work beyond simple noise models. While foundational insights are invaluable, there is a need to extend, validate and apply these theories using real-world data to make them useful. By fostering collaboration across such diverse disciplines, we can accelerate the development of quantum computing technologies that are reliable, reproducible, and stable.

Bibliography

[1] Travis S Humble, Himanshu Thapliyal, Edgard Munoz-Coreas, Fahd A Mohiyaddin, and Ryan S Bennink. Quantum computing circuits and devices. *IEEE Design & Test*, 36(3):69–94, 2019. 1, 18, 182

[2] Antoine Browaeys and Thierry Lahaye. Many-body physics with individually controlled rydberg atoms. *Nature Physics*, 16(2):132–142, 2020. 1

[3] Ashley Montanaro. Quantum algorithms: an overview. *npj Quantum Information*, 2(1):1–8, 2016. 1

[4] Jiangwei Shang, Yi-Lin Seah, Hui Khoon Ng, David John Nott, and Berthold-Georg Englert. Monte carlo sampling from the quantum state space. i. *New Journal of Physics*, 17(4):043017, 2015. 1

[5] Giovanny Espitia. On the role of quantum computing in science and cybersecurity. *arXiv preprint arXiv:2105.09942*, 2021. 1

[6] Meng-Leong How and Sin-Mei Cheah. Business renaissance: Opportunities and challenges at the dawn of the quantum computing era. *Businesses*, 3(4):585–605, 2023. 1

[7] Richard P Feynman, Robert B Leighton, Matthew Sands, and Everett M Hafner. The feynman lectures on physics. *American Journal of Physics*, 1965. 2

[8] C Cohen-Tannoudji, B Diu, and F Laloë. Quantum mechanics, volume i, ii, paris-new york: Èditions hermann and john wiley & sons, 1977. 2

[9] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics*. Cambridge University Press, 2020. 2, 3

[10] N David Mermin. *Quantum computer science: an introduction*. Cambridge University Press, 2007. 2, 4, 7

[11] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002. 2, 3, 4, 29, 57, 98, 132, 152

[12] R Bhatia. Matrix analysis, 1997. 2

[13] Karl Kraus, Arno Böhm, John D Dollard, and WH Wootters. *States, Effects, and Operations: Lectures in Mathematical Physics*. Springer, 1983. 2

[14] Amine Zeguendry, Zahi Jarir, and Mohamed Quafafou. Quantum machine learning: A review and case studies. *Entropy*, 25(2):287, 2023. 3

[15] Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738, 1995. 3

[16] Charles H. Bennett and Peter W. Shor. Quantum information theory. *IEEE Transactions on Information Theory*, 44(6):2724–2742, 1998. 3, 130

[17] JH Béjanin, CT Earnest, AS Sharafeldin, and M Mariantoni. Interacting defects generate stochastic fluctuations in superconducting qubits. *Physical Review B*, 104(9):094106, 2021. 3, 10, 18, 20

[18] Jonathan J Burnett, Andreas Bengtsson, Marco Scigliuzzo, David Niepce, Marina Kudra, Per Delsing, and Jonas Bylander. Decoherence benchmarking of superconducting qubits. *npj Quantum Information*, 5(1):1–8, 2019. 3, 10, 18, 19, 20, 29, 73, 155

[19] Malcolm Carroll, Sami Rosenblatt, Petar Jurcevic, Isaac Lauer, and Abhinav Kandala. Dynamics of superconducting qubit relaxation times. *arXiv preprint arXiv:2105.15201*, 2021. 3, 10, 18, 19, 20, 141, 152

[20] Dephasing and relaxation. available online: *https://wright.chem.wisc.edu/content/dephasing-and-relaxation-0/* (accessed on 2 dec 2022). 3, 21

[21] Thomas H Cormen, CE Leiserson, Ronald L Rivest, and Cliford Stein. Introduction to algorithms cambridge, massachusetts, 1990. 4, 7

[22] Stephen A Cook. The complexity of theorem-proving procedures. In *Logic, Automata, and Computational Complexity: The Works of Stephen A. Cook*, pages 143–152. 2023. 4

[23] Thomas Alexander, Naoki Kanazawa, Daniel J Egger, Lauren Capelluto, Christopher J Wood, Ali Javadi-Abhari, and David C McKay. Qiskit pulse: programming quantum computers through the cloud with pulses. *Quantum Science and Technology*, 5(4):044006, 2020. 7, 112, 130, 132, 166

[24] Harrison Ball, Michael J Biercuk, and Michael R Hush. Quantum firmware and the quantum computing stack. *Physics Today*, 74(3):28–34, 2021. 8

[25] Keith A. Britt, Fahd A. Mohiyaddin, and Travis S. Humble. Quantum accelerators for high-performance computing systems. In *2017 IEEE International Conference on Rebooting Computing (ICRC)*, pages 1–7, 2017. doi: 10.1109/ICRC.2017.8123664. 8

[26] Alexander McCaskey, Eugene Dumitrescu, Dmitry Liakh, and Travis Humble. Hybrid programming for near-term quantum computing systems. In *2018 IEEE International Conference on Rebooting Computing (ICRC)*, pages 1–12, 2018. doi: 10.1109/ICRC.2018.8638598. 8

[27] Ewout Van Den Berg, Zlatko K Minev, Abhinav Kandala, and Kristan Temme. Probabilistic error cancellation with sparse pauli-lindblad models on noisy quantum processors. *Nature Physics*, pages 1–6, 2023. 8, 140, 141, 153, 154, 162

[28] Sergey Bravyi, Sarah Sheldon, Abhinav Kandala, David C Mckay, and Jay M Gambetta. Mitigating measurement errors in multiqubit experiments. *Physical Review A*, 103(4):042605, 2021. 8, 99, 154, 161, 162, 169

[29] Philip Krantz, Morten Kjaergaard, Fei Yan, Terry P Orlando, Simon Gustavsson, and William D Oliver. A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews*, 6(2): 021318, 2019. 8, 18

[30] Thomas E Roth, Ruichao Ma, and Weng C Chew. An introduction to the transmon qubit for electromagnetic engineers. *arXiv preprint arXiv:2106.11352*, 2021. 8, 18, 72, 128, 151, 182

[31] Suhas Vittal, Poulami Das, and Moinuddin Qureshi. Eraser: Towards adaptive leakage suppression for fault-tolerant quantum computing. In *Proceedings of the 56th Annual IEEE/ACM International Symposium on Microarchitecture*, pages 509–525, 2023. 8, 151

[32] Avraham Ben-Aroya and Amnon Ta-Shma. Approximate quantum error correction for correlated noise. *IEEE Transactions on Information Theory*, 57(6):3982–3988, 2011. 8, 108

[33] John M Martinis. Qubit metrology for building a fault-tolerant quantum computer. *npj Quantum Information*, 1(1):1–3, 2015. 8, 18, 151

[34] John Preskill. Quantum computing in the nisq era and beyond. *Bulletin of the American Physical Society*, 64, 2019. 9, 19, 61, 72, 151

[35] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout Van Den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, et al. Evidence for

the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, 2023. 9, 153

[36] PV Klimov, Julian Kelly, Z Chen, Matthew Neeley, Anthony Megrant, Brian Burkett, Rami Barends, Kunal Arya, Ben Chiaro, Yu Chen, et al. Fluctuations of energy-relaxation times in superconducting qubits. *Physical Review Letters*, 121(9):090502, 2018. 10, 19, 29, 105, 141, 151, 152, 182

[37] Corey Rae H McRae, Gregory M Stiehl, Haozhi Wang, Sheng-Xiang Lin, Shane A Caldwell, David P Pappas, Josh Mutus, and Joshua Combes. Reproducible coherence characterization of superconducting quantum devices. *Applied Physics Letters*, 119(10):100501, 2021. 10, 19, 141, 152

[38] Rochus Klesse and Sandra Frank. Quantum error correction in spatially correlated quantum noise. *Physical Review Letters*, 95(23):230503, 2005. 10, 19

[39] Riddhi Swaroop Gupta, Luke CG Govia, and Michael J Biercuk. Integration of spectator qubits into quantum computer architectures for hardware tune-up and calibration. *Physical Review A*, 102(4):042611, 2020. 10, 20

[40] Pedro Parrado-Rodríguez, Ciarán Ryan-Anderson, Alejandro Bermudez, and Markus Müller. Crosstalk suppression for fault-tolerant quantum error correction with trapped ions. *Quantum*, 5:487, 2021. 10, 20, 151

[41] Chao Fang, Ye Wang, Shilin Huang, Kenneth R Brown, and Jungsang Kim. Crosstalk suppression in individually addressed two-qubit gates in a trapped-ion quantum computer. *Physical Review Letters*, 129(24):240504, 2022. 10, 20, 151

[42] Clemens Müller, Jürgen Lisenfeld, Alexander Shnirman, and Stefano Poletto. Interacting two-level defects as sources of fluctuating high-frequency noise in superconducting circuits. *Physical Review B*, 92(3):035442, 2015. 10, 105, 141, 151, 152, 182

[43] Qian Xu, Alireza Seif, Haoxiong Yan, Nam Mannucci, Bernard Ousmane Sane, Rodney Van Meter, Andrew N Cleland, and Liang Jiang. Distributed quantum error correction for chip-level catastrophic errors. *Physical Review Letters*, 129(24):240502, 2022. 10, 152

[44] Matt McEwen, Lara Faoro, Kunal Arya, Andrew Dunsworth, Trent Huang, Seon Kim, Brian Burkett, Austin Fowler, Frank Arute, Joseph C Bardin, et al. Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits. *Nature Physics*, 18(1): 107–111, 2022. 10, 151, 152

[45] Timothy Proctor, Melissa Revelle, Erik Nielsen, Kenneth Rudinger, Daniel Lobser, Peter Maunz, Robin Blume-Kohout, and Kevin Young. Detecting and tracking drift in quantum information processors. *Nature Communications*, 11(1):1–9, 2020. 10, 152

[46] Josu Etxezarreta Martinez, Patricio Fuentes, Pedro Crespo, and Javier Garcia-Frias. Time-varying quantum channel models for superconducting qubits. *npj Quantum Information*, 7(1): 1–10, 2021. 10, 20, 32, 59, 76, 105, 128, 133, 134, 151, 152, 182

[47] Arshag Danageozian, Ashe Miller, Pratik J Barge, Narayan Bhusal, and Jonathan P Dowling. Noisy coherent population trapping: applications to noise estimation and qubit state preparation. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 55(15):155503, 2022. 10, 152

[48] Swarnadeep Majumder, Leonardo Andreta de Castro, and Kenneth R Brown. Real-time calibration with spectator qubits. *npj Quantum Information*, 6(1):19, 2020. 10, 152

[49] Josu Etxezarreta Martinez, Patricio Fuentes, Pedro M Crespo, and Javier Garcia-Frias. Approximating decoherence processes for the design and simulation of quantum error correction codes on classical computers. *IEEE Access*, 8:172623–172643, 2020. 10, 31

[50] Antonio deMarti iOlius, Patricio Fuentes, Román Orús, Pedro M Crespo, and Josu Etxezarreta Martinez. Decoding algorithms for surface codes. *arXiv e-prints*, pages arXiv–2307, 2023. 10

[51] Antonio deMarti iOlius, Josu Etxezarreta Martinez, Patricio Fuentes, Pedro M Crespo, and Javier Garcia-Frias. Performance of surface codes in realistic quantum hardware. *Physical Review A*, 106(6):062428, 2022. 10

[52] Vladimir Vapnik. *The nature of statistical learning theory*. Springer science & business media, 1999. 10, 50, 51, 112

[53] Jose Carrasco, Andreas Elben, Christian Kokail, Barbara Kraus, and Peter Zoller. Theoretical and experimental perspectives of quantum verification. *PRX Quantum*, 2(1):010102, 2021. 10, 50

[54] Martin Kliesch and Ingo Roth. Theory of quantum system certification. *PRX Quantum*, 2(1): 010201, 2021. 11, 50, 51, 151, 182

[55] Quantum roadmap. Available online: *https://www.ibm.com/roadmaps/quantum/* (accessed on 12 feb 2024). 11, 50, 72, 113, 151

[56] Antonio Acín, Immanuel Bloch, Harry Buhrman, Tommaso Calarco, Christopher Eichler, Jens Eisert, Daniel Esteve, Nicolas Gisin, Steffen J Glaser, Fedor Jelezko, et al. The quantum

technologies roadmap: a european community view. *New Journal of Physics*, 20(8):080201, 2018. 11, 50, 113

[57] Richard Hughes and Todd Heinrichs. A quantum information science and technology roadmap. *Rep. LA-UR-04-1778, ARDA*, 2004. 11, 20, 50, 113

[58] Sumeet Khatri, Kunal Sharma, and Mark M Wilde. Information-theoretic aspects of the generalized amplitude-damping channel. *Physical Review A*, 102(1):012401, 2020. 11, 32, 50, 114, 133

[59] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, 94(1):015004, 2022. 11, 50, 102, 151, 152, 182

[60] Robin Harper, Steven T Flammia, and Joel J Wallman. Efficient learning of quantum noise. *Nature Physics*, 16(12):1184–1188, 2020. 11, 50

[61] Alexandru Gheorghiu, Theodoros Kapourniotis, and Elham Kashefi. Verification of quantum computation: An overview of existing approaches. *Theory of computing systems*, 63(4):715–808, 2019. 11, 50

[62] Ye-Chao Liu, Jiangwei Shang, Xiao-Dong Yu, and Xiangdong Zhang. Efficient verification of quantum processes. *Phys. Rev. A*, 101:042315, Apr 2020. doi: 10.1103/PhysRevA.101.042315. URL `https://link.aps.org/doi/10.1103/PhysRevA.101.042315`. 11, 50

[63] A. McCaskey, E. Dumitrescu, M. Chen, D. Lyakh, and T. S. Humble. Validating quantum-classical programming models with tensor network simulations. *PLoS ONE*, 13:e0206704, 2018. doi: 10.1371/journal.pone.0206704. 11, 50

[64] T. Nguyen, D. Lyakh, P. Lotshaw, A. McCaskey, R. Bennink, and W Jong. Scalable programming workflows for validation of quantum computers. *2021 IEEE/ACM Second International Workshop on Quantum Computing Software (QCS)*, 2021. doi: 10.1109/qcs54837.2021.00013. 11, 50

[65] Alexander J McCaskey, Zachary P Parks, Jacek Jakowski, Shirley V Moore, Titus D Morris, Travis S Humble, and Raphael C Pooser. Quantum chemistry as a benchmark for near-term quantum computers. *npj Quantum Information*, 5(1):1–8, 2019. 11, 50

[66] Robin Blume-Kohout and Kevin C. Young. A volumetric framework for quantum computer benchmarks. *Quantum*, 4:362, November 2020. ISSN 2521-327X. doi: 10.22331/q-2020-11-15-362. URL `https://doi.org/10.22331/q-2020-11-15-362`. 11, 50

[67] J. Helsen and S. Wehner. A benchmarking procedure for quantum networks. *npj Quantum Inf*, 9, 2023. doi: 10.1038/s41534-022-00628-x. 11, 50

[68] Samuele Ferracin, Seth T Merkel, David McKay, and Animesh Datta. Experimental accreditation of outputs of noisy quantum computers. *arXiv preprint arXiv:2103.06603*, 2021. 11, 51, 151, 182

[69] David P DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000. 11, 61, 73, 76

[70] Robin J Blume-Kohout. Modeling and characterizing noise in quantum processors. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2020. 11, 151

[71] Robin Blume-Kohout. Optimal, reliable estimation of quantum states. *New Journal of Physics*, 12(4):043034, 2010. 11, 151, 182

[72] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. Error mitigation for short-depth quantum circuits. *Physical Review Letters*, 119(18):180509, 2017. 12, 133, 137, 153, 154, 155

[73] Joseph M Lukens, Kody JH Law, Ajay Jasra, and Pavel Lougovski. A practical and efficient approach for bayesian quantum state estimation. *New Journal of Physics*, 22(6):063038, 2020. 12, 128, 153, 156, 183

[74] Muqing Zheng, Ang Li, Tamás Terlaky, and Xiu Yang. A bayesian approach for characterizing and mitigating gate and measurement errors. *arXiv preprint arXiv:2010.09188*, 2020. 12, 72, 128, 153, 156, 183

[75] Neil J Gordon, David J Salmond, and Adrian FM Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F-radar and Signal Processing*, volume 140, pages 107–113. IET, 1993. 12, 128, 156, 183

[76] Jayesh H Kotecha and Petar M Djuric. Gaussian sum particle filtering. *IEEE Transactions on Signal Processing*, 51(10):2602–2612, 2003. 12, 128, 156, 183

[77] Daniel Gottesman. The heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998. 12

[78] D Spehner, F Illuminati, M Orszag, and W Roga. Geometric measures of quantum correlations with bures and hellinger distances. *Lectures on general quantum correlations and their applications*, pages 105–157, 2017. 12, 128

[79] Roger B Nelsen. *An introduction to copulas*. Springer science & business media, 2007. 13, 110, 114

[80] Yong Wan, Daniel Kienzler, Stephen D Erickson, Karl H Mayer, Ting Rei Tan, Jenny J Wu, Hilma M Vasconcelos, Scott Glancy, Emanuel Knill, David J Wineland, et al. Quantum gate teleportation between separated qubits in a trapped-ion processor. *Science*, 364(6443):875–878, 2019. 18

[81] K. W. Chan, W. Huang, C. H. Yang, J. C. C. Hwang, B. Hensen, T. Tanttu, F. E. Hudson, K. M. Itoh, A. Laucht, A. Morello, and A. S. Dzurak. Assessment of a silicon quantum dot spin qubit environment via noise spectroscopy. *Phys. Rev. Applied*, 10:044017, Oct 2018. doi: 10.1103/PhysRevApplied.10.044017. URL `https://link.aps.org/doi/10.1103/PhysRevApplied.10.044017`. 18

[82] XY Jin, A Kamal, AP Sears, T Gudmundsen, D Hover, J Miloshi, R Slattery, F Yan, J Yoder, TP Orlando, et al. Thermal and residual excited-state population in a 3d transmon qubit. *Physical Review Letters*, 114(24):240501, 2015. 19

[83] Joseph Emerson, Marcus Silva, Osama Moussa, Colm Ryan, Martin Laforest, Jonathan Baugh, David G Cory, and Raymond Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–1896, 2007. 31

[84] Marcus Silva, Easwar Magesan, David W Kribs, and Joseph Emerson. Scalable protocol for identification of correctable codes. *Physical Review A*, 78(1):012347, 2008. 31

[85] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, 60(3):1888, 1999. 31, 32, 133

[86] Tilo Eggeling and Reinhard F Werner. Separability properties of tripartite states with $u \otimes u \otimes u$ symmetry. *Physical Review A*, 63(4):042111, 2001. 31

[87] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A*, 80(1):012304, 2009. 31

[88] Easwar Magesan. Gaining information about a quantum channel via twirling. Master's thesis, University of Waterloo, 2008. 31

[89] Zhenyu Cai, Xiaosi Xu, and Simon C Benjamin. Mitigating coherent noise using pauli conjugation. *npj Quantum Information*, 6(1):17, 2020. 32, 133

[90] Samuele Ferracin, Theodoros Kapourniotis, and Animesh Datta. Accrediting outputs of noisy intermediate-scale quantum computing devices. *New Journal of Physics*, 21(11):113038, 2019. 51

[91] Rudolf Beran. Minimum hellinger distance estimates for parametric models. *The annals of Statistics*, pages 445–463, 1977. 51, 100, 103

[92] Hamed Rahimian, Güzin Bayraksan, and Tito Homem-de Mello. Identifying effective scenarios in distributionally robust stochastic programs with total variation distance. *Mathematical Programming*, 173(1-2):393–430, 2019. 51

[93] Hubert W Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, 62(318):399–402, 1967. 51

[94] Shiming Xiang, Feiping Nie, and Changshui Zhang. Learning a mahalanobis distance metric for data clustering and classification. *Pattern recognition*, 41(12):3600–3612, 2008. 51

[95] Juan Antonio Cuesta and Carlos Matrán. Notes on the wasserstein metric in hilbert spaces. *The Annals of Probability*, pages 1264–1276, 1989. 51

[96] Lloyd Fisher and John W Van Ness. Distinguishability of probability measures. *The Annals of Mathematical Statistics*, pages 381–392, 1969. 51

[97] Tim Van Erven and Peter Harremos. Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014. 51

[98] Frank Nielsen. On a generalization of the jensen–shannon divergence and the jensen–shannon centroid. *Entropy*, 22(2):221, 2020. 51

[99] Rui F Vigelis, Luiza HF De Andrade, and Charles C Cavalcante. Properties of a generalized divergence related to tsallis generalized divergence. *IEEE Transactions on Information Theory*, 66(5):2891–2897, 2019. 51

[100] Luczka J. Hänggi P. Dajka, J. Distance between quantum states in the presence of initial qubit-environment correlations: A comparative study. *Phys. Rev. A*, 84, 2011. doi: 10.1103/physreva.84.032120. 51

[101] Jasra A. Law K. J. H. Lougovski P. Bennink, R. S. Estimation and uncertainty quantification for the output from quantum simulators. *Foundations of Data Science*, 1:157–176, 2019. doi: 10.3934/fods.2019007. 51

[102] Géza Tóth and Iagoba Apellaniz. Quantum metrology from a quantum information science perspective. *Journal of Physics A: Mathematical and Theoretical*, 47(42):424006, 2014. 51

[103] Costin Bădescu, Ryan O'Donnell, and John Wright. Quantum state certification. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 503–514, 2019. 56

[104] Kenneth Rudinger, Timothy Proctor, Dylan Langharst, Mohan Sarovar, Kevin Young, and Robin Blume-Kohout. Probing context-dependent errors in quantum processors. *Physical Review X*, 9(2):021045, 2019. 56

[105] Alistair WR Smith, Kiran E Khosla, Chris N Self, and MS Kim. Qubit readout error mitigation with bit-flip averaging. *Science Advances*, 7(47):eabi8009, 2021. 58, 99, 100, 162

[106] Filip B Maciejewski, Flavio Baccari, Zoltán Zimborás, and Michał Oszmaniec. Modeling and mitigation of cross-talk effects in readout noise with applications to the quantum approximate optimization algorithm. *Quantum*, 5:464, 2021. 62

[107] Carlos A Pérez-Delgado and Pieter Kok. Quantum computers: Definition and implementations. *Physical Review A*, 83(1):012303, 2011. 62

[108] Samudra Dasgupta and Travis S Humble. Assessing the stability of noisy quantum computation. In *Quantum Communications and Quantum Imaging XX*, volume 12238, pages 44–49. SPIE, 2022. 63

[109] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. In *Proceedings of the twenty-fifth annual ACM symposium on theory of computing*, pages 11–20, 1993. 65, 76, 108, 153, 160

[110] Samudra Dasgupta and Travis S Humble. Reliability of noisy quantum computing devices. *arXiv preprint arXiv:2307.06833*, 2023. 72, 75, 128, 152, 155

[111] Quantum computing software and programming tools. Available online: *https://www.ibm.com/quantum-computing/experience/* (accessed on 21 august 2021). 72, 101, 132

[112] Samudra Dasgupta and Travis S Humble. Stability of noisy quantum computing devices. *arXiv preprint arXiv:2105.09472*, 2021. 72, 106, 152

[113] Processor types. Available online: *https://docs.quantum.ibm.com/run/processor-types/* (accessed on 12 feb 2024). 72

[114] Gadi Aleksandrowicz, Thomas Alexander, Panagiotis Barkoutsos, Luciano Bello, Yael Ben-Haim, D Bucher, FJ Cabrera-Hernández, J Carballo-Franquis, A Chen, CF Chen, et al. Qiskit: An open-source framework for quantum computing. *Accessed on: Mar*, 16, 2019. 73

[115] Oak Ridge Leadership Computing Facility. Available online: *https://www.olcf.ornl.gov/* (accessed on 12 feb 2024). 73

[116] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roee Ozeri, Signe Seidelin, and David J Wineland. Randomized benchmarking of quantum gates. *Physical Review A*, 77(1):012307, 2008. 73

[117] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005. 73

[118] Travis Hurant and Daniel D Stancil. Asymmetry of cnot gate operation in superconducting transmon quantum processors using cross-resonance entangling. *arXiv preprint arXiv:2009.01333*, 2020. 73

[119] Ivar Martin, L Bulaevskii, A Shnirman, and YM Galperin. Stationary and non-stationary noise in superconducting quantum devices. In *Noise and Fluctuations in Circuits, Devices, and Materials*, volume 6600, pages 35–46. SPIE, 2007. 74

[120] Tianyu Xie, Zhiyuan Zhao, Shaoyi Xu, Xi Kong, Zhiping Yang, Mengqi Wang, Ya Wang, Fazhan Shi, and Jiangfeng Du. 99.92%-fidelity cnot gates in solids by noise filtering. *Physical Review Letters*, 130(3):030601, 2023. 74

[121] Samudra Dasgupta and Travis S Humble. Characterizing the stability of nisq devices. In *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 419–429. IEEE, 2020. 75, 152

[122] E. Novais and H. Baranger. Decoherence by correlated noise and quantum error correction. *Phys. Rev. Lett.*, 97, 2006. doi: 10.1103/physrevlett.97.040501. 76

[123] M. Biercuk, T. Stace, S. Flammia, and M. Biercuk. Effect of noise correlations on randomized benchmarking. *Phys. Rev. A*, 93, 2016. doi: 10.1103/physreva.93.022303. 76

[124] X. Tang, D. Li, C. Zhou, and L. Zhang. Bivariate distribution models using copulas for reliability analysis. *Proceedings of the Institution of Mechanical Engineers, Part O:*, 227:499–512, 2013. doi: 10.1177/1748006x13481928. 77, 110, 183

[125] Michel Verleysen and Damien François. The curse of dimensionality in data mining and time series prediction. In *International work-conference on artificial neural networks*, pages 758–770. Springer, 2005. 77

[126] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S347, 2005. 98

[127] Michał Oszmaniec, Filip B Maciejewski, and Zbigniew Puchała. Simulating all quantum measurements using only projective measurements and postselection. *Physical Review A*, 100 (1):012351, 2019. 99

[128] Michael R Geller. Rigorous measurement error correction. *Quantum Science and Technology*, 5 (3):03LT01, 2020. 99

[129] A Sklar. Fonctions de répartition ¡ n dimensions et leur marges. *Publ. Inst. Stat. Paris*, 8: 131–229, 1959. 110, 112

[130] C. Klüppelberg and G. Kuhn. Copula structure analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 71:737–753, 2009. doi: 10.1111/j.1467-9868.2009.00707.x. 110

[131] A. Dutfoy and R. Lebrun. Practical approach to dependence modelling using copulas. *Proceedings of the Institution of Mechanical Engineers, Part O:*, 223:347–361, 2009. doi: 10.1243/1748006xjrr226. 110

[132] C. Chesneau. Extensions of two bivariate strict archimedean copulas. *Computational Journal of Mathematical and Statistical Sciences*, 0:159–180, 2023. doi: 10.21608/cjmss.2023.205330.1007. 110

[133] Elton Yechao Zhu, Sonika Johri, Dave Bacon, Mert Esencan, Jungsang Kim, Mark Muir, Nikhil Murgai, Jason Nguyen, Neal Pisenti, Adam Schouela, et al. Generative quantum learning of joint probability distribution functions. *Physical Review Research*, 4(4):043092, 2022. 110

[134] Giovanni De Luca and Giorgia Rivieccio. Multivariate tail dependence coefficients for archimedean copulae. In *Advanced statistical methods for the analysis of large data-sets*, pages 287–296. Springer, 2012. 110

[135] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition.* Princeton university press, 2015. 110

[136] Sascha Wilkens and Joe Moorhouse. Quantum computing for financial risk measurement. *Quantum Information Processing*, 22(1):51, 2023. 110

[137] Christian Genest, Johanna Nešlehová, and Johanna Ziegel. Inference in multivariate archimedean copula models. *Test*, 20:223–256, 2011. 110

[138] Ben Calderhead. A general construction for parallelizing metropolis- hastings algorithms. *Proceedings of the National Academy of Sciences*, 111(49):17408–17413, 2014. 112, 131

[139] F. Chapeau-Blondeau. Modeling and simulation of a quantum thermal noise on the qubit. *Fluct. Noise Lett.*, 21, 2022. doi: 10.1142/s0219477522500602. 114

[140] Thomas G Wong and Raqueline AM Santos. Exceptional quantum walk search on the cycle. *Quantum Information Processing*, 16:1–17, 2017. 128

[141] Simon J Devitt. Performing quantum computing experiments in the cloud. *Physical Review A*, 94(3):032329, 2016. 128

[142] Stephen M Stigler. Thomas bayes's bayesian inference. *Journal of the Royal Statistical Society: Series A (General)*, 145(2):250–258, 1982. 129

[143] Amine Mezghani, Michel T Ivrlac, and Josef A Nossek. Achieving near-capacity on large discrete memoryless channels with uniform distributed selected input. In *2008 International Symposium on Information Theory and Its Applications*, pages 1–6. IEEE, 2008. 129

[144] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American statistical Association*, 112(518):859–877, 2017. 129, 130

[145] Robert Bassett and Julio Deride. Maximum a posteriori estimators as a limit of bayes estimators. *Mathematical Programming*, 174:129–144, 2019. 130

[146] Eric R Ziegel. The elements of statistical learning, 2003. 130

[147] Roman Kápl and Pavel Parízek. Endicheck: Dynamic analysis for detecting endianness bugs. In *Tools and Algorithms for the Construction and Analysis of Systems: 26th International Conference, TACAS 2020, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2020, Dublin, Ireland, April 25–30, 2020, Proceedings, Part II 26*, pages 254–270. Springer, 2020. 130

[148] Arjun K Gupta and Saralees Nadarajah. *Handbook of beta distribution and its applications*. CRC press, 2004. 131

[149] Scott R Eliason. *Maximum likelihood estimation: Logic and practice*. Number 96. Sage, 1993. 131

[150] Bo Yang, Rudy Raymond, and Shumpei Uno. Efficient quantum readout-error mitigation for sparse measurement outcomes of near-term quantum devices. *Physical Review A*, 106(1): 012423, 2022. 132

[151] Kriste Krstovski, David A Smith, Hanna M Wallach, and Andrew McGregor. Efficient nearest-neighbor search in the probability simplex. In *Proceedings of the 2013 Conference on the Theory of Information Retrieval*, pages 101–108, 2013. 133

[152] Tzu-Tsung Wong. Generalized dirichlet distribution in bayesian analysis. *Applied Mathematics and Computation*, 97(2-3):165–181, 1998. 133

[153] Suguru Endo, Simon C Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications. *Physical Review X*, 8(3):031027, 2018. 133, 137

[154] Shuaining Zhang, Yao Lu, Kuan Zhang, Wentao Chen, Ying Li, Jing-Ning Zhang, and Kihwan Kim. Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system. *Nature Communications*, 11(1):587, 2020. 137

[155] Abhinav Kandala, Kristan Temme, Antonio D Córcoles, Antonio Mezzacapo, Jerry M Chow, and Jay M Gambetta. Error mitigation extends the computational reach of a noisy quantum processor. *Nature*, 567(7749):491–495, 2019. 151

[156] G.A.L. White, F.A. Pollock, L.C.L. Hollenberg, K. Modi, and C.D. Hill. Non-markovian quantum process tomography. *PRX Quantum*, 3:020344, May 2022. doi: 10.1103/PRXQuantum.3.020344. URL `https://link.aps.org/doi/10.1103/PRXQuantum.3.020344`. 151

[157] Ted Thorbeck, Andrew Eddins, Isaac Lauer, Douglas T McClure, and Malcolm Carroll. Two-level-system dynamics in a superconducting qubit due to background ionizing radiation. *PRX Quantum*, 4(2):020356, 2023. 151

[158] Ivan Henao, Jader P Santos, and Raam Uzdin. Adaptive quantum error mitigation using pulse-based inverse evolutions. *npj Quantum Information*, 9(1):120, 2023. 152

[159] Giacomo Torlai, Christopher J Wood, Atithi Acharya, Giuseppe Carleo, Juan Carrasquilla, and Leandro Aolita. Quantum process tomography with unsupervised learning and tensor networks. *arXiv preprint arXiv:2006.02424*, 2020. 153

[160] Quasi-probability representation of an ideal gate. Available online: *https://mitiq.readthedocs.io/en/v0.11.1/examples/pec-tutorial.html#task-2-sampling-from-the-quasi-probability-representation-of-an-ideal-gate/* (last accessed on 25 mar 2024). 154, 155

[161] Andrew Gelman, John B Carlin, Hal S Stern, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 1995. 157

[162] Devinderjit Sivia and John Skilling. *Data analysis: a Bayesian tutorial*. OUP Oxford, 2006. 157

[163] Christian P Robert, George Casella, and George Casella. *Introducing monte carlo methods with r*, volume 18. Springer, 2010. 158

[164] Peter M Lee. *Bayesian statistics*. Oxford University Press London:, 1989. 158

[165] David C McKay, Christopher J Wood, Sarah Sheldon, Jerry M Chow, and Jay M Gambetta. Efficient z gates for quantum computing. *Physical Review A*, 96(2):022330, 2017. 162

[166] William E Ryan et al. An introduction to ldpc codes. *CRC Handbook for Coding and Signal Processing for Recording Systems*, 5(2):1–23, 2004. 184

[167] Sergey Bravyi, Andrew W Cross, Jay M Gambetta, Dmitri Maslov, Patrick Rall, and Theodore J Yoder. High-threshold and low-overhead fault-tolerant quantum memory. *Nature*, 627(8005): 778–782, 2024. 184