Tutorial: Quantum Noise Characterization and Mitigation

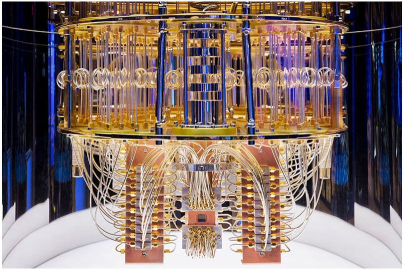
Daniel Volya & Prabhat Mishra

Computer & Information Science & Engineering
University of Florida

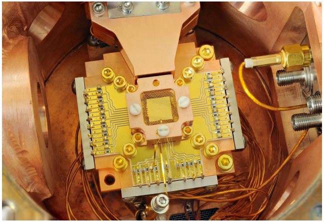
September 19, 2022

https://volya.xyz/blog/qce2022

Quantum Computing



(a) IBM Q System



(b) NIST Ion trap 2011

Why?

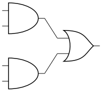
1. Quantum computers of today are “noisy” and computationally weak.
2. But we want answers today (e.g. groundstate energies!)
3. What can we do?

- Beginner: 20%
- Intermediate: 60%
- Advanced: 20%

Table of contents

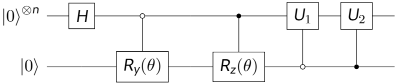
Introduction - Classical vs. Quantum

Classical



0	0	0	0	0
0	0	0	1	0
⋮	⋮	⋮	⋮	⋮
1	1	1	1	1

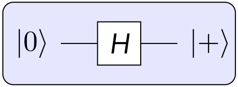
Quantum



$$U\left|\psi\right\rangle = \begin{pmatrix} 1 & i & \dots & e^{i\theta} \\ -i & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ e^{-i\theta} & 1 & \dots & 0 \end{pmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Introduction - Example

Hadamard Gate, H



$$H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) = |+\rangle$$

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{bmatrix}1\\0\end{bmatrix}$$

Introduction - Example (Density matrix)

Hadamard Gate, H

$$H|0\rangle\langle 0|H^\dagger = |+\rangle\langle +|$$

$$\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

Introduction - Behind the Scenes

Discrete quantum operations (gates) are *implemented* as time evolutions of a quantum system.

$$i\frac{\partial}{\partial t}|\psi\rangle = \hat{H}(t)|\psi\rangle, \quad i\frac{\partial}{\partial t}\rho = [\hat{H}(t), \rho]$$

$$|\psi(t_f)\rangle = U(t_f, t_i)|\psi(t_i)\rangle$$

$$U(t_f, t_i) = \mathcal{T}\exp\left(-i\int_{t_i}^{t_f}\hat{H}(t)dt\right)$$

Introduction - Behind the Scenes

$$H(t) = H_0 + \sum_j a_j(t) H_j$$

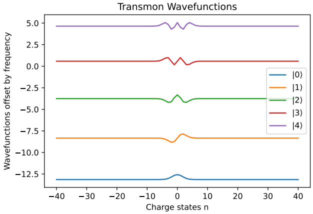
E.g. The Cooper Pair Box (CPB) Hamiltonian:

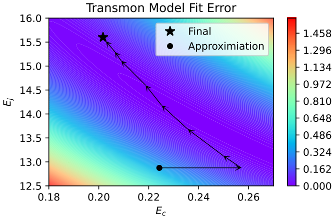
$$H_T = H_{el} + H_j$$

$$H_{el} = 4E_C(\hat{n} - n_g)^2 \approx 4E_C(n - n_g)^2 |n\rangle\langle n|$$

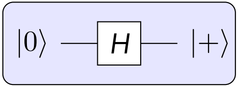
$$H_j = -\frac{E_j}{2}\cos\hat{\phi} \approx -\frac{E_j}{2}\sum_n |n\rangle\langle n{+}1| + |n{+}1\rangle\langle n|$$

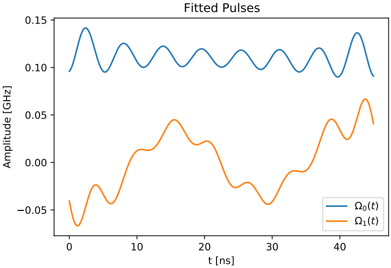
$$UH_dU^\dagger = \epsilon Re(B)(\tilde{c} + \tilde{c}^\dagger) - i\epsilon Im(B)(\tilde{c} - \tilde{c}^\dagger)$$



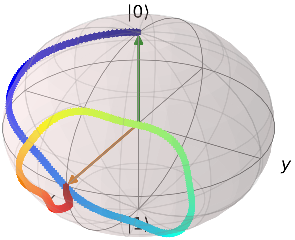


Introduction - Zooming In





(a) Optimal quantum control pulses



(b) Simulated "path" of evolution

Introduction - Universal Computation

- Solovay–Kitaev theorem
- Quantum circuit of m constant-qubit gates can be approximated to ϵ error by a quantum circuit of $O(m \log^c(m/\epsilon))$ gates from a universal gate set
- "Fast" to do so (logarithmic)

Example universal gate set:

$$\{H, CNOT, T\}$$

Introduction - Summary

- Classical computers are restricted to one particular state at a time. Quantum computers can be in superposition
- Quantum circuits represent a sequence of discrete (entangling) gates
- Quantum gates are implemented as evolutions of some underlying quantum system
- Can reach[1] any arbitrary quantum state with a crafted sequence of gates

[1] ϵ-close given a *universal* gate set

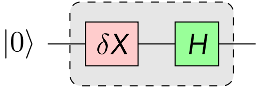
Introduction - The Catch

Quantum computers must satisfy conflicting requirements:

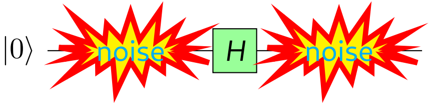
- Externally control, entangle, and measure qubits
- Isolate the qubits from their environment

noise ☹

Introduction - Coherent Noise



Introduction - Deoherent Noise



- Quantum Channels
- Church of the larger Hilbert space?

Introduction - Quantum threshold theorem

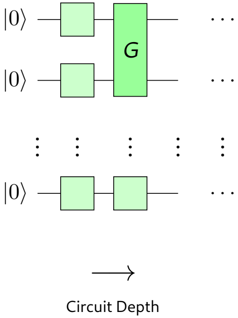
Quantum computer with a physical error rate below a certain threshold can, through application of quantum error correction schemes, suppress the logical error rate to arbitrarily low levels.

Table: Some ”static” error correction codes

Code	# Qubits	Error Threshold
Bit-flip/Phase-flip	3	$p < 0.5$
Shor	9	$p \lessapprox 0.0323$
Steane	7	$p \lessapprox 0.0579$

Estimates: 0.1% probability of depolarizing error, surface code would need 1000-10000 physical qubits per one logical qubit (E. Campbell et al. arXiv:1612.07330)

Introduction - Consequence of Noise



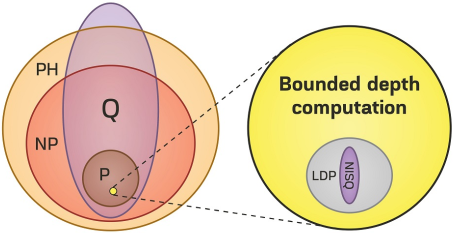
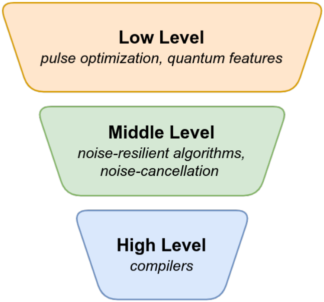


Figure: Image credits: Gil Kalai. NISQ circuits are computationally very weak.

1. The depth of a circuit is constrained
2. Noisy Intermediate Scale Quantum (NISQ) computers (John Preskill)


Low Level
pulse optimization, quantum features
Middle Level
noise-resilient algorithms,
noise-cancellation
High Level
compilers

https://volya.xyz/blog/qce2022

Quantum State (Process) Tomography

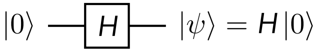
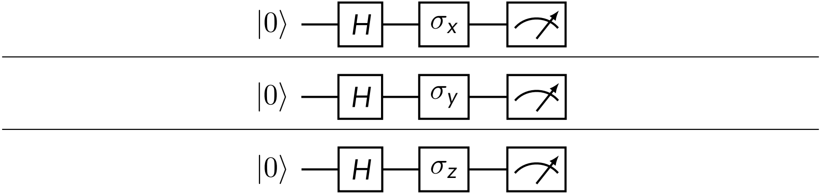


Figure: Hadamard gate

What is state $|\psi\rangle$?
(What is the matrix H?)

Quantum State (Process) Tomography



Each measurement has 2 possible outcomes, hence 6 different possibilities. Need to fit these 6 possibilities into a 2-by-2 matrix ρ.

Quantum State (Process) Tomography

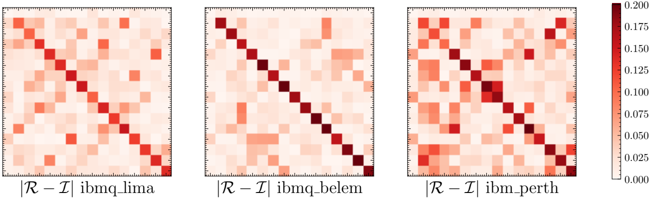


Figure: 2-Qubit Process Tomography

Other Methods

Summary of Noise Characterization Techniques

Technique	Purpose	Advantage	Disadvantage
1. Quantum State (Process) Tomography	$\rho\ (\mathcal{E}(\rho))$	Full Classification	$O(2^{n^2})$
• Pseudoinverse	ρ	Simple Least-Squares	Matrix inversion
• MLE	ρ	No Matrix Inversion	Prediction not justifiable
• BME	ρ	Honest Solution	Optimization routine
• Matrix Completion	ρ	$O(r2^n \log^2 2^n)$	ρ must be close to pure
• Permutation Invariant Tomography	ρ_{PI}	$O(n^2)$	Not complete ρ
2. Measurement Error Characterization	A	Full Classification	$O(2^n)$
• Tensor Product Model	$A = (A_i \otimes A_{i-1} \ldots)$	O(2)	No cross-talk
• CTMP	$A = e^G$	$O(e^{5\gamma} poly(n))$	Assumes small noise
3. Randomized Benchmarking	α	Avg. Error Rate	$O(n^{2^2}(\ln n^2)^3)$
• 2 (t)-design	α	$O(m \log(1/\epsilon))$	Only clifford gates
4. Comparing Output Distribution		Fewer Measurements	Classical simulation
• Kullback-Leibler	D_{KL}		Asymmetric
• Jenses-Shannon	D_{JS}	Symmetric	
• Average Cross Entropy	α_f	SPAM Resilient	Considers average
5. Quantum Volume	QV	Performance Metric	Considers average

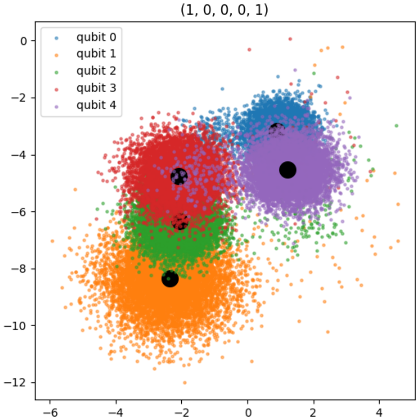


Figure: In-phase and quadrature components (IQ) for state $|10001\rangle$ (*ibmq_rome*)

Noise Mitigation

NISQ computers are... weak. Can we help them with classical computers?

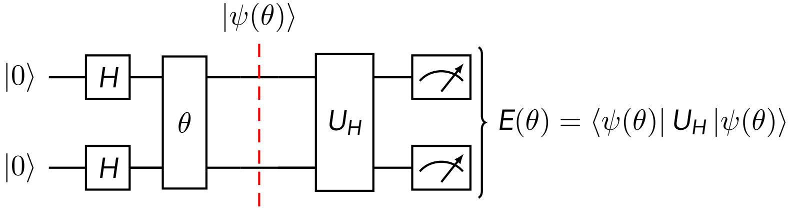
Noise Mitigation

Summary of Noise Mitigation and Correction Techniques

Technique	Purpose	Advantage	Disadvantage
1. Quantum Error Correction	$\|\psi\rangle = \sum_i \alpha_i \|\phi_i\rangle$	Reliable Qubits	Costly; Noise-model
2. Classical-Postprocessing	$\langle A \rangle$	Minimizes quantum resources	Needs regression models
• RNE	$\langle A \rangle = f(0; \vec{a})$	No noise-model	f-bias
• CDR	$\langle A \rangle = f(\langle A \rangle^{\text{noisy}}, \vec{a})$	Self-tuning	Clifford gate approx.
• vnCDR	$\langle A \rangle = f(0; \langle A \rangle^{\text{noisy}}, \vec{a})$	Enhanced self-tuning	Clifford gate approx.
• PEC	$\langle A \rangle = \gamma \mathbb{E}\{\dots\}$	Unbiased estimate	Gate tomography
3. Passive Noise Mitigation	$\|\psi\rangle$ (ρ)	Inherent robustness	System Model
• Optimal Control	$a_{i,n}$	Direct control	Runtime
• Open Quantum Systems	H	Powerful model	Simulation
• Decoherence-free subsystems	H_S	Inherent noise mitigation	Limited to system
• Holonomic QC	$\Gamma(t), a_{i,n}$	Control robustness	Limited to system

NISQ Algorithms ("Quantum machine learning")

Example: Variational Quantum Circuits



Zero-noise Extrapolation

Idea:

1. Let τ quantify the noise-level in a circuit.
2. Intentionally increase circuit depth, $\tau' = \lambda\tau$
3. Fit a function $f(\lambda; a_0, a_1, \ldots)$
4. Extrapolate at "zero noise" $\lambda = 0$

(A. Kandala et al. arXiv:1805.04492)
(Y. Li et al. arXiv:1611.09301)

Zero-noise Extrapolation

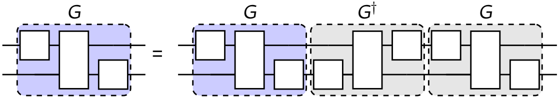


Figure: Global folding strategy of gates.

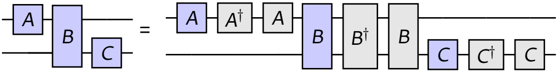


Figure: Local folding strategy of gates.

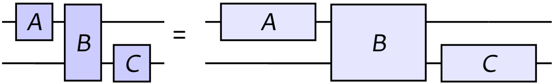


Figure: Pulse stretching of gates (K. Temme et al. arXiv:1612.02058).

Zero-noise Extrapolation

Pros:

- Do not need detailed knowledge of a noise model

Cons:

- May suffer from large bias – a weak-formulated function *f*, such as a low-degree polynomial, may not achieve the zero-noise limit.
- For circuit with short depth, the lowest error points may be dominated by noise and perform worse than the unmitigated result.

Clifford data regression

Idea:

- Use Clifford circuits to classically simulate
- Compare simulation with quantum computer results
- Fit a regression model to recover noiseless result

(P. Czarnik et al. arXiv:2005.10189)

Clifford data regression - Gottesman–Knill theorem

- *U stabilizes* $|\psi\rangle$ if $U|\psi\rangle = |\psi\rangle$
- Clifford group: $\{V \text{ s.t. } VP_iV^\dagger = P_j\}$
- A circuit consisting of only Clifford gates can be simulated efficiently

Clifford data regression

1. Generate data set: $\{\langle A\rangle_i^{\text{noisy}}, \langle A\rangle_i^{\text{exact}}\}$
2. Fit $\langle A\rangle^{\text{exact}} = f\left(\langle A\rangle^{\text{noisy}}, \vec{a}\right)$
3. Correct $\langle A\rangle^{\text{noisy}}$ with fitted function parameters.

Clifford Data Regression

Pros:

- Do not need detailed knowledge of a noise model
- The error mitigation self-tunes with respect to a given quantum computer

Cons:

- Test circuits are mostly Clifford approximations, which may not completely span the available Hilbert space
- Need a balance between the number of Clifford gates, non-Clifford gates, and classical computational complexity

Probabilistic Error Cancellation

Idea:

1. $\mathcal{G}_i = \sum_\alpha \nu_{i,\alpha} \mathcal{G}'_{i,\alpha}, \quad \nu_{i,\alpha} \in \mathbb{R}$
2. Quasi-probability: $\sum_\alpha \nu_{i,\alpha} = 1, \quad \gamma_i = \sum_\alpha |\nu_{i,\alpha}| \geq 1.$
3. $\mathcal{U} = \sum_{\vec{\alpha}} \nu_{\vec{\alpha}} \mathcal{U}'_{\vec{\alpha}}$
4. $\langle A \rangle_{\text{ideal}} = \operatorname{tr}[A\mathcal{U}] = \sum_{\vec{\alpha}} \nu_{\vec{\alpha}} \operatorname{tr}\left[A\mathcal{U}'_{\vec{\alpha}}\right] = \sum_{\vec{\alpha}} \nu_{\vec{\alpha}} \langle A_{\vec{\alpha}} \rangle_{\text{noisy}}$

(A. Mari et al. arXiv:2108.02237)
(J. Sun et al. arXiv:2001.04891)
(S. Zhang. arXiv:1905.10135)

Probabilistic Error Cancellation

Pros:

- Provides an unbiased estimation of expectation values by utilizing gate tomography

Cons:

- Sampling $\langle A_{\vec{\alpha}} \rangle_{\text{noisy}}$ can be costly
- Complete tomographic knowledge of gates in large-scale qubit systems is infeasible

https://volya.xyz/blog/qce2022