Quantum Computing Benchmarks in Light of Quantum Error Correction

Yaakov S. Weinstein

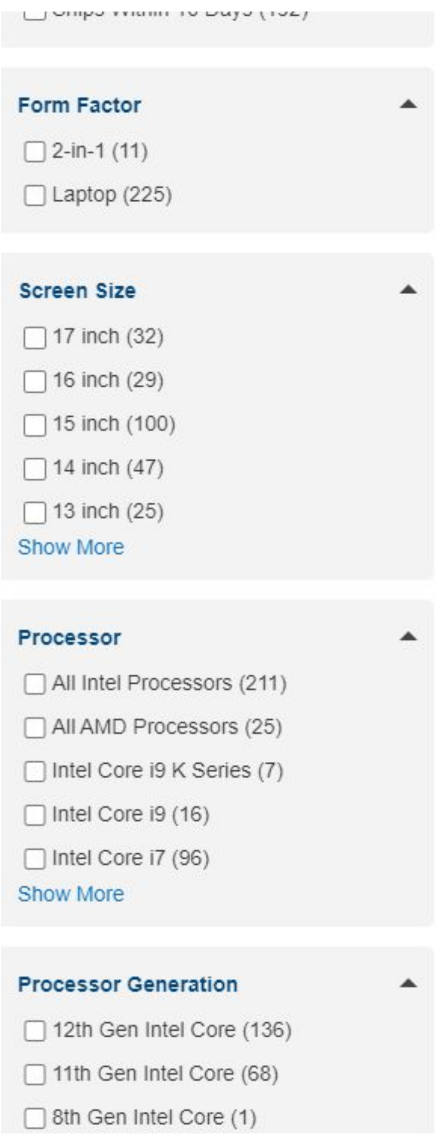
The MITRE Corporation, Quantum Technologies Group

Editor-in-Chief Quantum Information Processing

weinstein@mitre.org

SOLVING PROBLEMS FOR A SAFER WORLD™

How does a non-initiate buy a laptop?



□ Ships Within 10 Days (152)

Form Factor

□ 2-in-1 (11)

□ Laptop (225)

Screen Size

□ 17 inch (32)

□ 16 inch (29)

□ 15 inch (100)

□ 14 inch (47)

□ 13 inch (25)

Show More

Processor

□ All Intel Processors (211)

□ All AMD Processors (25)

□ Intel Core i9 K Series (7)

□ Intel Core i9 (16)

□ Intel Core i7 (96)

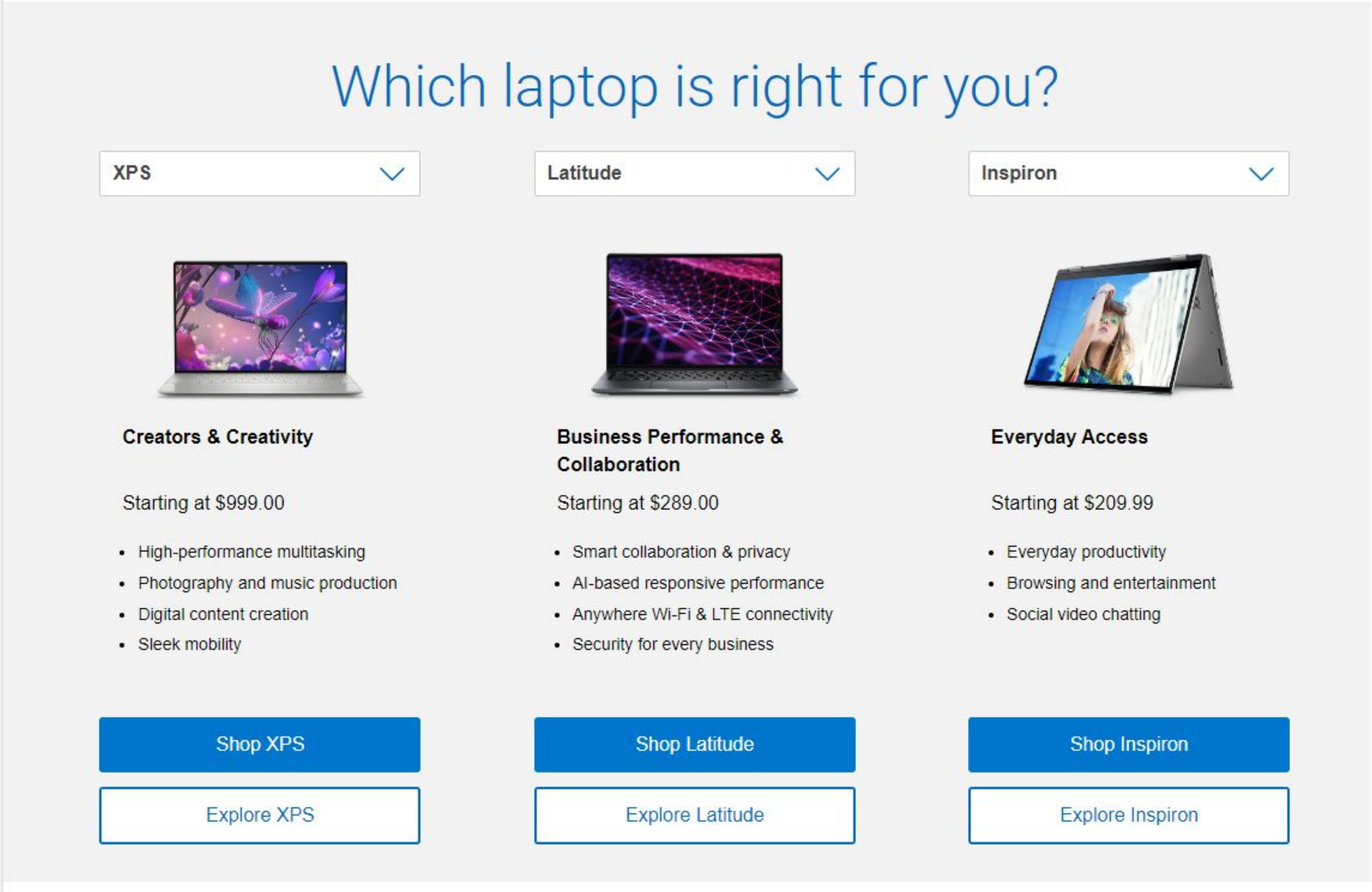
Show More

Processor Generation

□ 12th Gen Intel Core (136)

□ 11th Gen Intel Core (68)

□ 8th Gen Intel Core (1)



Which laptop is right for you?

XPS

Creators & Creativity

Starting at $999.00

- High-performance multitasking
- Photography and music production
- Digital content creation
- Sleek mobility

Shop XPS

Explore XPS

Latitude

Business Performance & Collaboration

Starting at $289.00

- Smart collaboration & privacy
- AI-based responsive performance
- Anywhere Wi-Fi & LTE connectivity
- Security for every business

Shop Latitude

Explore Latitude

Inspiron

Everyday Access

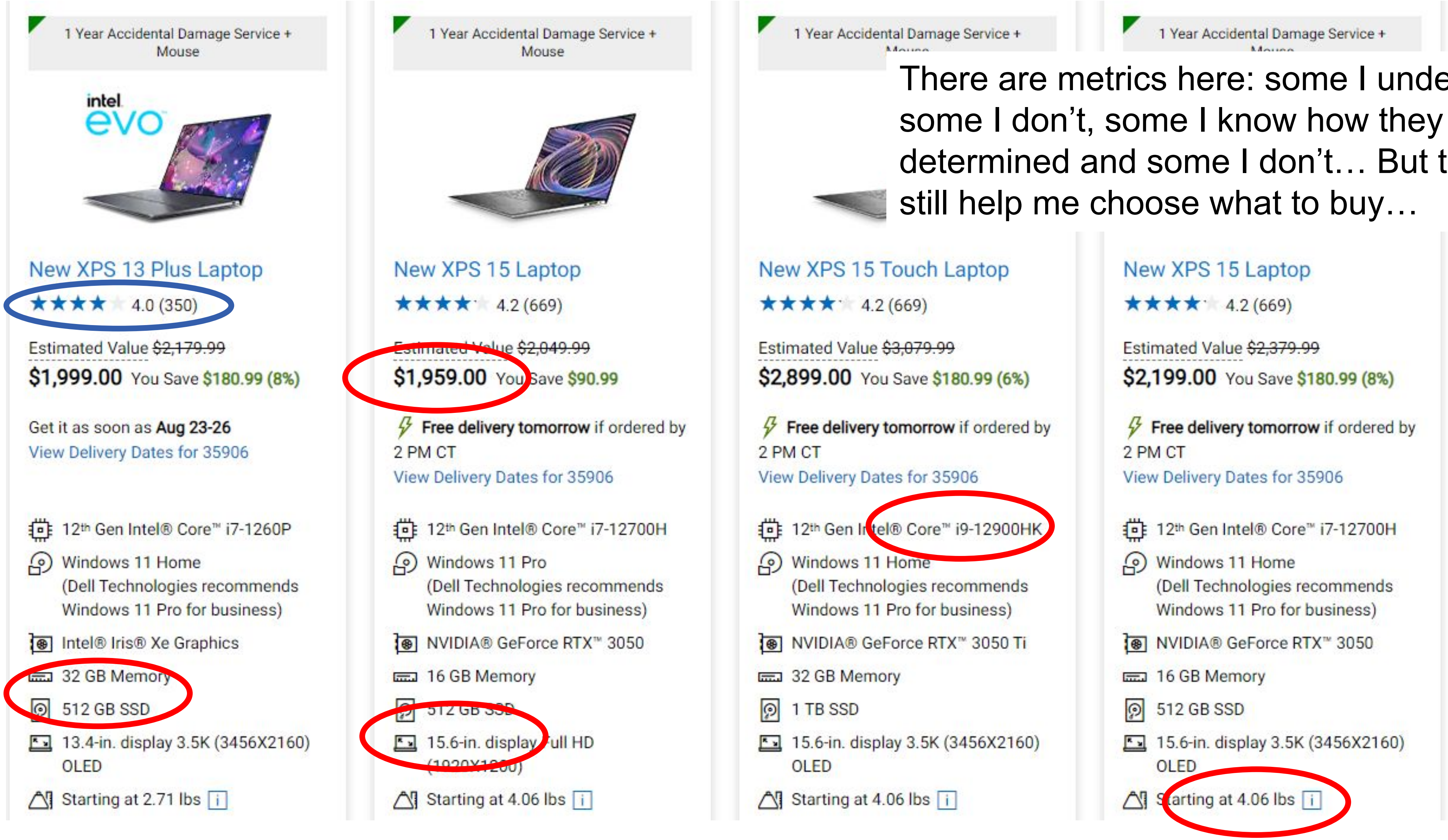
Starting at $209.99

- Everyday productivity
- Browsing and entertainment
- Social video chatting

Shop Inspiron

Explore Inspiron

How does a non-initiate buy a laptop?



1 Year Accidental Damage Service + Mouse

New XPS 13 Plus Laptop

★★★★ 4.0 (350)

Estimated Value ~~$2,179.99~~

$1,999.00 You Save $180.99 (8%)

Get it as soon as Aug 23-26

View Delivery Dates for 35906

- 12th Gen Intel® Core™ i7-1260P
- Windows 11 Home (Dell Technologies recommends Windows 11 Pro for business)
- Intel® Iris® Xe Graphics
- 32 GB Memory
- 512 GB SSD
- 13.4-in. display 3.5K (3456X2160) OLED
- Starting at 2.71 lbs

1 Year Accidental Damage Service + Mouse

New XPS 15 Laptop

★★★★ 4.2 (669)

Estimated Value ~~$2,049.99~~

$1,959.00 You Save $90.99

Free delivery tomorrow if ordered by 2 PM CT

View Delivery Dates for 35906

- 12th Gen Intel® Core™ i7-12700H
- Windows 11 Pro (Dell Technologies recommends Windows 11 Pro for business)
- NVIDIA® GeForce RTX™ 3050
- 16 GB Memory
- 512 GB SSD
- 15.6-in. display Full HD (1920X1200)
- Starting at 4.06 lbs

1 Year Accidental Damage Service + Mouse

New XPS 15 Touch Laptop

★★★★ 4.2 (669)

Estimated Value ~~$3,079.99~~

$2,899.00 You Save $180.99 (6%)

Free delivery tomorrow if ordered by 2 PM CT

View Delivery Dates for 35906

- 12th Gen Intel® Core™ i9-12900HK
- Windows 11 Home (Dell Technologies recommends Windows 11 Pro for business)
- NVIDIA® GeForce RTX™ 3050 Ti
- 32 GB Memory
- 1 TB SSD
- 15.6-in. display 3.5K (3456X2160) OLED
- Starting at 4.06 lbs

1 Year Accidental Damage Service + Mouse

New XPS 15 Laptop

★★★★ 4.2 (669)

Estimated Value ~~$2,379.99~~

$2,199.00 You Save $180.99 (8%)

Free delivery tomorrow if ordered by 2 PM CT

View Delivery Dates for 35906

- 12th Gen Intel® Core™ i7-12700H
- Windows 11 Home (Dell Technologies recommends Windows 11 Pro for business)
- NVIDIA® GeForce RTX™ 3050
- 16 GB Memory
- 512 GB SSD
- 15.6-in. display 3.5K (3456X2160) OLED
- Starting at 4.06 lbs

There are metrics here: some I understand some I don't, some I know how they were determined and some I don't… But they can still help me choose what to buy…

Quantum computing benchmarks and metrics

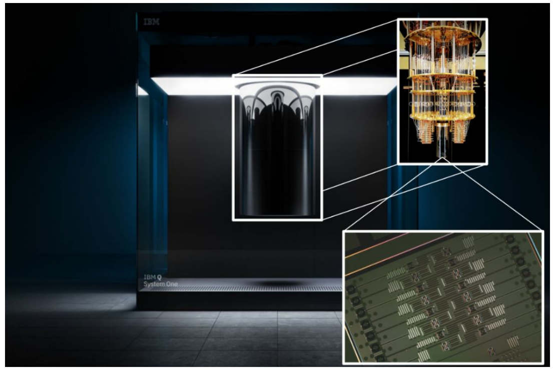
Objective

- Identify or develop benchmarks and metric(s) for evaluating quantum computing devices both in the near-term (i.e., noisy and limited), as well as long term (universal, fault-tolerant/error corrected quantum computers)
- Use metric(s) to evaluate current existing quantum computing platforms for which data is available
- Estimate current rate of technological progress over time (e.g., to allow initial estimate 5 or 10-year projections)

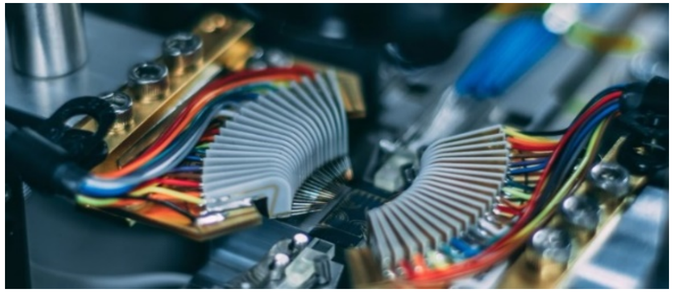
Outcomes

Ability to:

- Easily perform cost-benefit analysis for any specific commercial quantum computing device of interest
- Directly compare competing quantum computing platforms
- Quantitatively track the progress in this emerging field to better predict and anticipate long-term disruptions that quantum computing will create across a variety of applications



vs.



Outline

- **What's the difference between a benchmark and a metric?**
- Benchmarks
  - Survey commonly used benchmarks and how good they are with respect to desired criteria
  - Speculate on how benchmarks would be modified when error correction is used
  - Thought experiment on Quantum Process Tomography
- Metrics
  - Survey commonly used metrics and how good they are with respect to desired criteria
  - Speculate on how metrics would be modified when error correction is used
  - Analysis of Quantum Volume

Quantum Computers: Benchmarks and Metrics

Benchmark

- Procedure used to test a specific quantum computer
- Provides data about specific aspect(s) of the system in question
- Needed to characterize or validate performance of given device

Metric

- Number characterizing given quantum computer
- Figure of merit allowing comparison between devices or tracking progress over time
- May be measured using specific benchmark or act as more general framework

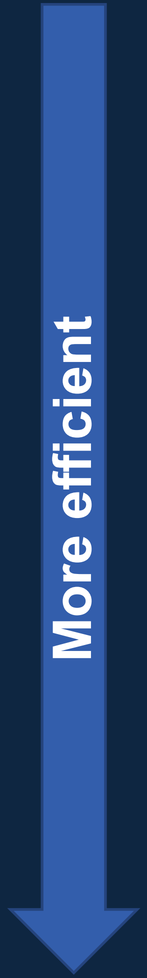
A topic allowing for collaboration between academia, industry, standards bodies, and government

Benchmarking

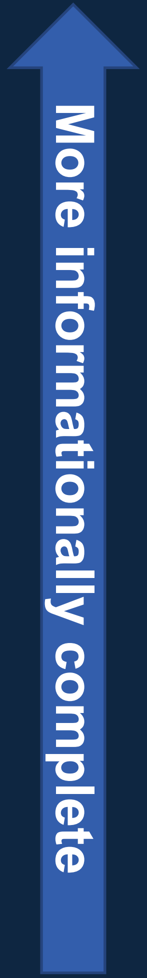
Evaluating Quantum Benchmarks

- Is it system/hardware agnostic?
- Efficiency/how well does it scale?
- What assumptions/models does evaluation rely on?
- What type and how much information does the benchmark give about system
  - e.g. Simple pass/fail vs quantitative number(s)
- How complete a picture is the information gathered? Is anything ignored?

Common Quantum Benchmarking Types


More efficient

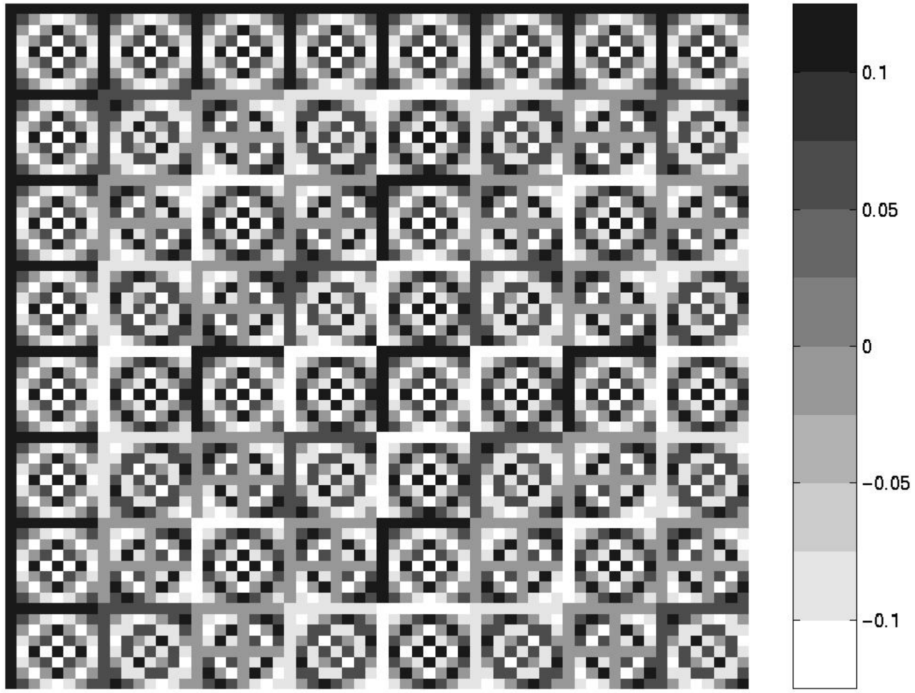
- **Tomographic Benchmarks** (i.e., Quantum State, Process, or Gate Set Tomography)
  - Most informationally complete but very inefficient
- **Compressed/Adaptive Learning Based Methods**
  - Process can be more efficient than tomography through use of assumptions about system or state
- **Quantum Fidelity Estimation**
  - Gives less information than tomography (i.e., measuring fidelity directly or through a fidelity witness)
- **Entropic Sampling Benchmarking** (e.g., Cross-entropy Benchmarking)
  - Based on sampling the output from a series of random circuits
  - Method used in evaluating the device in Google's 'Quantum Supremacy' claim
- **Randomized Benchmarking**
  - Estimation of the rates of certain types of errors via applying a sequence of random gates
  - Very commonly used due to the insensitivity to State Preparation And Measurement (SPAM) errors
- **Application Based Benchmarks**
  - Test based on problem size of a high-level application (e.g., Shor, some quantum chemistry problem)


More informationally complete

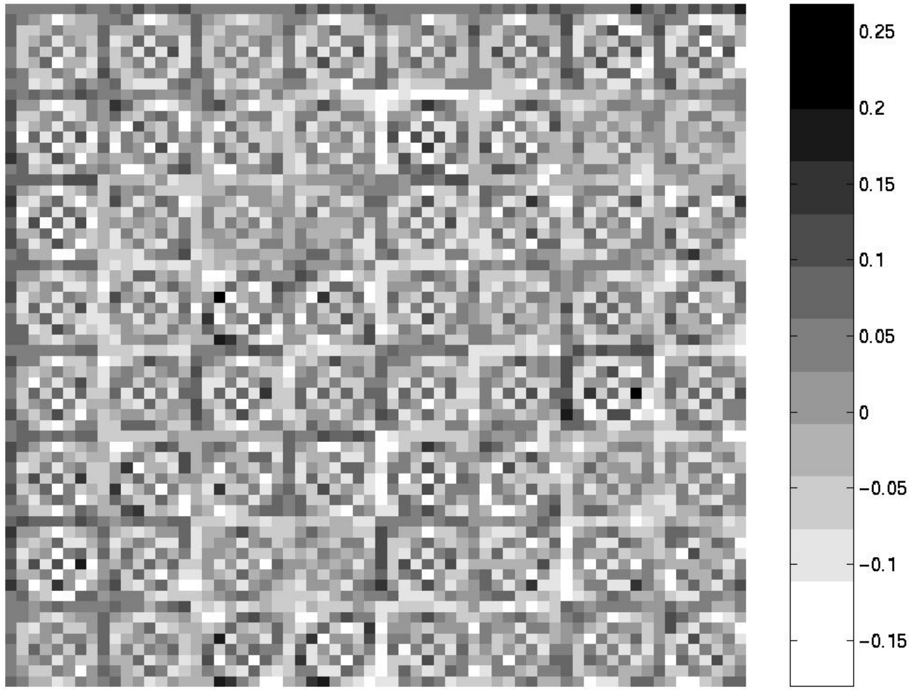
Tomographic Benchmarks

- Repeated measurements together with a computational tomographic reconstruction
- Can be used to characterize an unknown quantum state, process, or set of processes
- Hardware agnostic, inefficient, and provides most information

theory

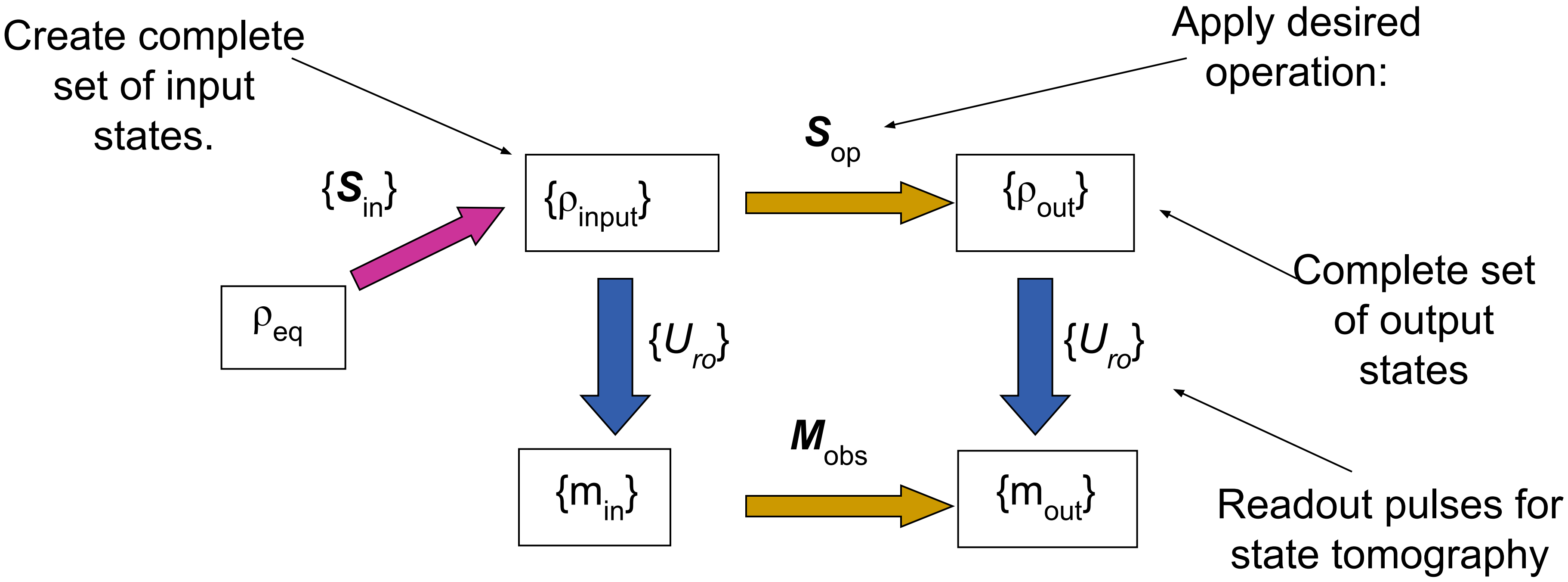


experiment



Quantum Process Tomography

Quantum Process Tomography (QPT) is the experimental procedure to completely determine the evolution of an open quantum system.



Create complete set of input states.

Apply desired operation:

$\boldsymbol{S}_{op}$

$\{\boldsymbol{S}_{in}\}$

$\{\rho_{input}\}$

$\{\rho_{out}\}$

ρ_{eq}

$\{U_{ro}\}$

$\{U_{ro}\}$

Complete set of output states

$\boldsymbol{M}_{obs}$

$\{m_{in}\}$

$\{m_{out}\}$

Readout pulses for state tomography

$$m_{xx} \equiv tr(U_{ro}\rho_{xx}\sigma_{-})$$

Reconstruction of observables

Superoperators and Kraus Operators

- Isolated quantum system undergoes unitary evolution as described by the Schroedinger equation.
- Actual quantum systems are affected by the outside environment
- Phenomenon such as relaxation and other types of decoherence cannot be described by a simple unitary operator

Superoperator N^2 X N^2

$$\begin{bmatrix} \quad S \quad \end{bmatrix}\begin{bmatrix} \rho_{initial} \end{bmatrix} = \begin{bmatrix} \rho_{final} \end{bmatrix}$$

Density matrices as N^2 X 1 vectors

Superoperator form… preserves Hermiticity, trace, and is completely positive

Kraus form… If the evolution is unitary there is only one Kraus operator.

$$\rho_{out} = \sum_k A_k \rho_{in} A_k^+$$

$$1 = \sum_k A_k A_k^+$$

Logical Qubit or Physical Qubits?

- What happens if you encode your information via error correction?
- Let’s assume a [[7,1,3]] QECC, what happens?
- What are we trying to learn?
- Can we use tomography to learn about the encoding process?

Superoperator
4 X 4 or 16384
X 16384

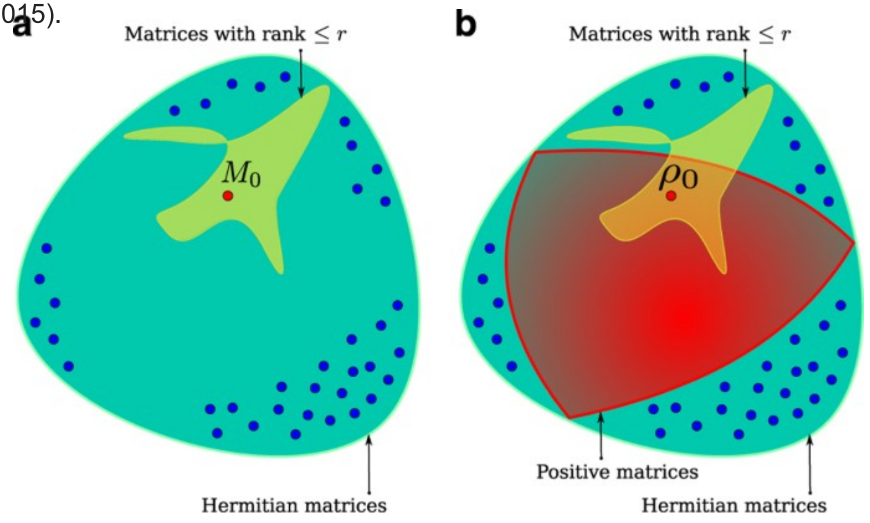
$$\begin{bmatrix} \quad S \quad \end{bmatrix}\begin{bmatrix} \rho_{initial} \end{bmatrix} = \begin{bmatrix} \rho_{final} \end{bmatrix}$$

Density matrices
as 4 X 1 or
16384 X 1
vectors

- Our main objective may be to simply ensure we have the correct state going into the next gate
- However, we can learn more about the decoherence processes if we look at the entire process, even the errors as they appear in other parts of the superoperator representation

Compressed/Adaptive Learning Based Methods

Kalev, A., Kosut, R. & Deutsch, I. Quantum tomography protocols with positivity are compressed sensing protocols. *npj Quantum Inf* **1,** 15018 (2015).



- Rather than performing an exhaustive deterministic set of tests (tomography), one can choose a random (e.g. compressive tomography) or adaptive (e.g. learning based) sub-set of tests
- The quality and completeness of the information gained strongly dependent on quality of model assumptions

✔ System/hardware agnostic

✔ Efficiency much better than tomography, e.g. PAC learning can sometimes scale linear in measurement (though not in computational reconstruction)

× Assumptions: Direct tradeoff in efficiency vs. assumptions underlying model or applicable circumstances

- Information gained (measured + assumed) less than but potentially comparable to tomography
- Type of information may be general (like tomography) or look more like model parameter fitting

Usage: More efficient alternative to tomography for e.g. larger sub-systems or iterative component testing

Quantum Fidelity Estimation/Witnessing

- Fidelity estimation measures how close a quantum state or process is to the ‘ideal’ case rather than performing a full tomographic characterization
- For certain special types of desired states (e.g. certain highly entangled systems), a much more efficient fidelity witness can be used to estimate if the system is close to being in such a state

✔ System/hardware agnostic (but witnesses work for only special circumstances)

- Efficiency? More efficient than tomography, but typically still exponential scaling for fidelity estimation
- Assumptions: State/process is close to ideal
- Information gained: How close is a state/process to the assumed ideal

Usage: Sanity check for state prep of NISQ devices (e.g. can device create a specific highly entangled state)

Entropic Sampling Benchmarking

- A type of fidelity estimation which measures the entropy in the output of a series of randomly applied quantum circuits
- This method was used by Google to claim quantun supremacy, even though the measured circuits responsible for these claims showed very low fidelity

✔ System/hardware agnostic

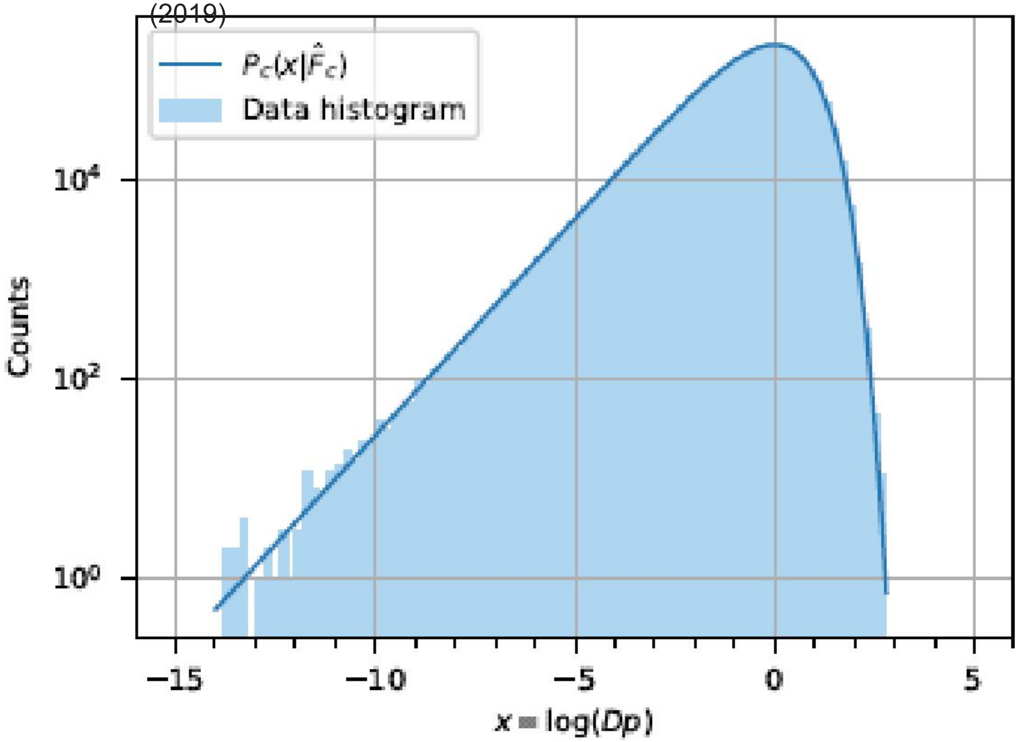
- Efficiency: Easy to test, exponentially hard to verify

× Assumptions: Amplifies certain types of noise while averaging out/ignoring other types (e.g. coherent vs incoherent noise)

- Information gained: average overall fidelity of a certain class of circuits randomly chosen

Usage: Inferring computationally intractable quantum computations even for low fidelity circuits (e.g. shortcut for claiming quantum supremacy)

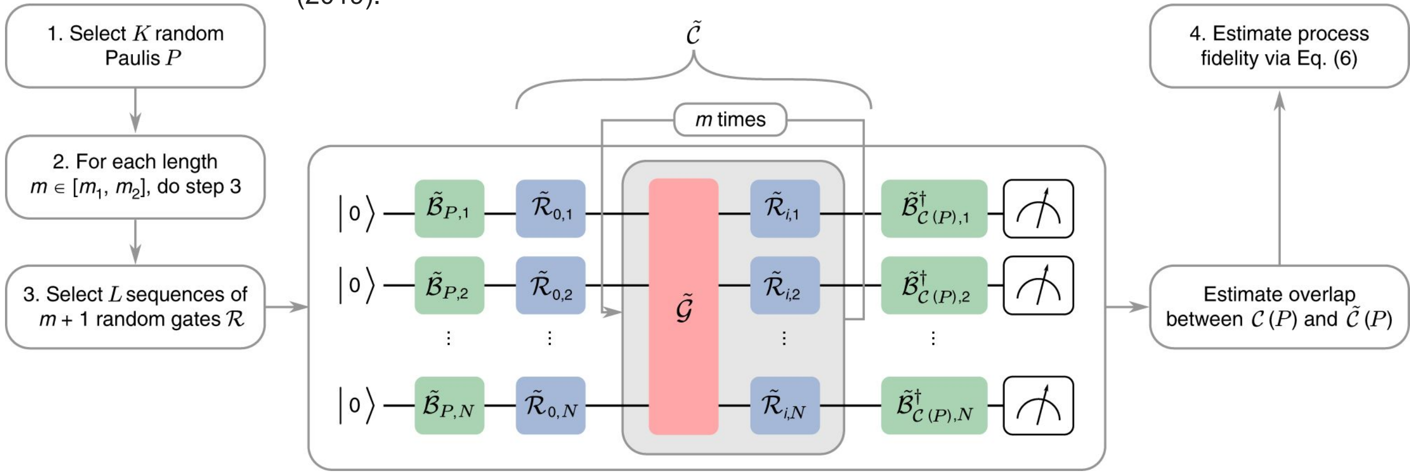
Arute, F., Arya, K., Babbush, R. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574,** 505–510 (2019)



MITRE

Randomized Benchmarking

- Randomized benchmarking refers to the idea that a set of subroutines is chosen randomly from some set
- Long periods of random circuits can be chosen to amplify certain types of noise, allowing for precise measurement of the average effect of such noise
- Robust against state preparation & measurement (SPAM) errors



1. Select K random Paulis P

2. For each length $m \in [m_1, m_2]$, do step 3

3. Select L sequences of $m + 1$ random gates $\mathcal{R}$

Estimate overlap between $\mathcal{C}\,(P)$ and $\tilde{\mathcal{C}}\,(P)$

4. Estimate process fidelity via Eq. (6)

$\tilde{\mathcal{C}}$
m times
$|0\rangle$
$\tilde{\mathcal{B}}_{P,1}$
$\tilde{\mathcal{B}}_{P,2}$
$\tilde{\mathcal{B}}_{P,N}$
$\tilde{\mathcal{R}}_{0,1}$
$\tilde{\mathcal{R}}_{0,2}$
$\tilde{\mathcal{R}}_{0,N}$
$\tilde{\mathcal{G}}$
$\tilde{\mathcal{R}}_{i,1}$
$\tilde{\mathcal{R}}_{i,2}$
$\tilde{\mathcal{R}}_{i,N}$
$\tilde{\mathcal{B}}^{\dagger}_{\mathcal{C}\,(P),1}$
$\tilde{\mathcal{B}}^{\dagger}_{\mathcal{C}\,(P),2}$
$\tilde{\mathcal{B}}^{\dagger}_{\mathcal{C}\,(P),N}$

Erhard, A., Wallman, J.J., Postler, L. *et al.* Characterizing large-scale quantum computers via cycle benchmarking. *Nat Commun* **10,** 5347 (2019).

✔ System/hardware agnostic

✔ Scales efficiently with simple post processing

× Assumptions: Amplifies certain types of noise while averaging out/ignoring other types (e.g. coherent vs incoherent noise)

× Gives information only about certain specific classes of errors

Usage: Measuring incoherent sources of noise in NISQ systems

Application Based Benchmarks

- For sufficiently advanced devices, benchmarking can be fully abstracted to the application level
- Test based on a computational problem that can be easily verified (e.g. Shor or certain quantum chemistry simulations)
- Benchmarking at this level most resembles classical computing benchmarking (e.g. LINPACK)

✔ System/hardware agnostic (with minimum resource requirement)

✔ Efficient

- Assumption: Application benchmark is representative of intended device usage

✔ Information most directly related to usage (can device do a specific type of computation)

Usage: Gauging progress in advanced (especially post-NISQ) devices

Metrics

Evaluating Quantum Metrics

- Is it universal?
  - e.g. only for Noisy intermediate scale quantum (NISQ) systems or will this work for large fault tolerant systems?
- How easy to compute/verify
  - i.e. either with access to the machine or estimation based on system specs
- How can it be gamed/how easy is it for metric to be distorted?
- What does the metric(s) tell us?
  - e.g. General-purpose metric vs application specific
- What does the metric(s) not tell us? What aspects are overlooked/ignored?

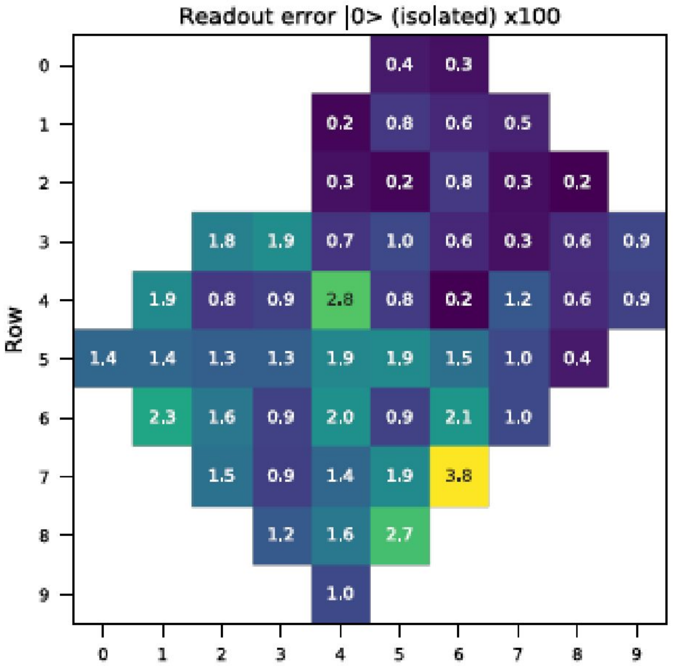
List of Common Quantum Metrics

Ordered from least to most mature

- **Metrics Based on System Spec Lists**
  - May include some set of coherence times, gate speeds/clock frequencies, error rates, initialization/measurement fidelities (SPAM errors), qubit numbers, qubit connectedness/topologies
- **Quantum Volume** (introduced by IBM)
  - Circuit space-time size of largest random square circuit device can perform
- **Generalized Volumetric Metrics**
  - Generalization of the quantum volume to a more general framework of quantifying devices in terms of both size and depth of circuits device can perform
- **Application Based Metrics**
  - Metric based on the size of the largest problem of a certain class of problems that device can successfully solve (e.g. largest number factored by Shor’s algorithm or largest Ising model that can be effectively simulated)

List of System Specs

- A list of one or more features of the system such as qubit numbe and topography, available quantum gates, error rates, coherent times, qubit layout, etc.
- Easy to game by reporting good features and de-emphasizing ba (e.g. only stating qubit number)
- Inconsistent reporting between vendors makes comparisons difficult



Row	0	1	2	3	4	5	6	7	8	9
0						0.4	0.3			
1					0.2	0.8	0.6	0.5		
2					0.3	0.2	0.8	0.3	0.2	
3			1.8	1.9	0.7	1.0	0.6	0.3	0.6	0.9
4		1.9	0.8	0.9	2.8	0.8	0.2	1.2	0.6	0.9
5	1.4	1.4	1.3	1.3	1.9	1.9	1.5	1.0	0.4	
6		2.3	1.6	0.9	2.0	0.9	2.1	1.0		
7			1.5	0.9	1.4	1.9	3.8			
8				1.2	1.6	2.7				
9					1.0					

Arute, F., Arya, K., Babbush, R. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574,** 505–510 (2019)

× Not universal (no standard set of system variables)

- Some features easier to compute/verify than others (e.g. number of qubits vs full noise characterization)

✔ Very easy to game, especially if only some system details are reported

× Only tells us those details that are reported. Generally gives a very incomplete picture for evaluating and comparing system performance

Application Based Metrics*

- Outcome of specific application-based benchmark
- Test based on a specific computational problem representing a desired application area (e.g. a chemistry type problem if quantum simulation is desired application)
- Metrics at this level are the most mature and most closely resembles classical computing benchmarking (e.g. LINPACK as a test of linear algebra capabilities)

- Universal only for mature systems
  - × Not applicable for early NISQ devices
- Easy to test/interpret
- Quality of the metric depends on how close the test benchmarking is to the problem of interest. Can be misleading if not representative
- Information gained is how well can the system do a set of representative test problems

Quantum Volume

- Metric proposed by IBM
- Describes largest random square circuit that can be computed with high fidelity (e.g. n qubits to n logical gate depth gives a quantum volume of 2^n)*
- IBM, Honeywell, and IonQ have all announced quantum volumes for their systems

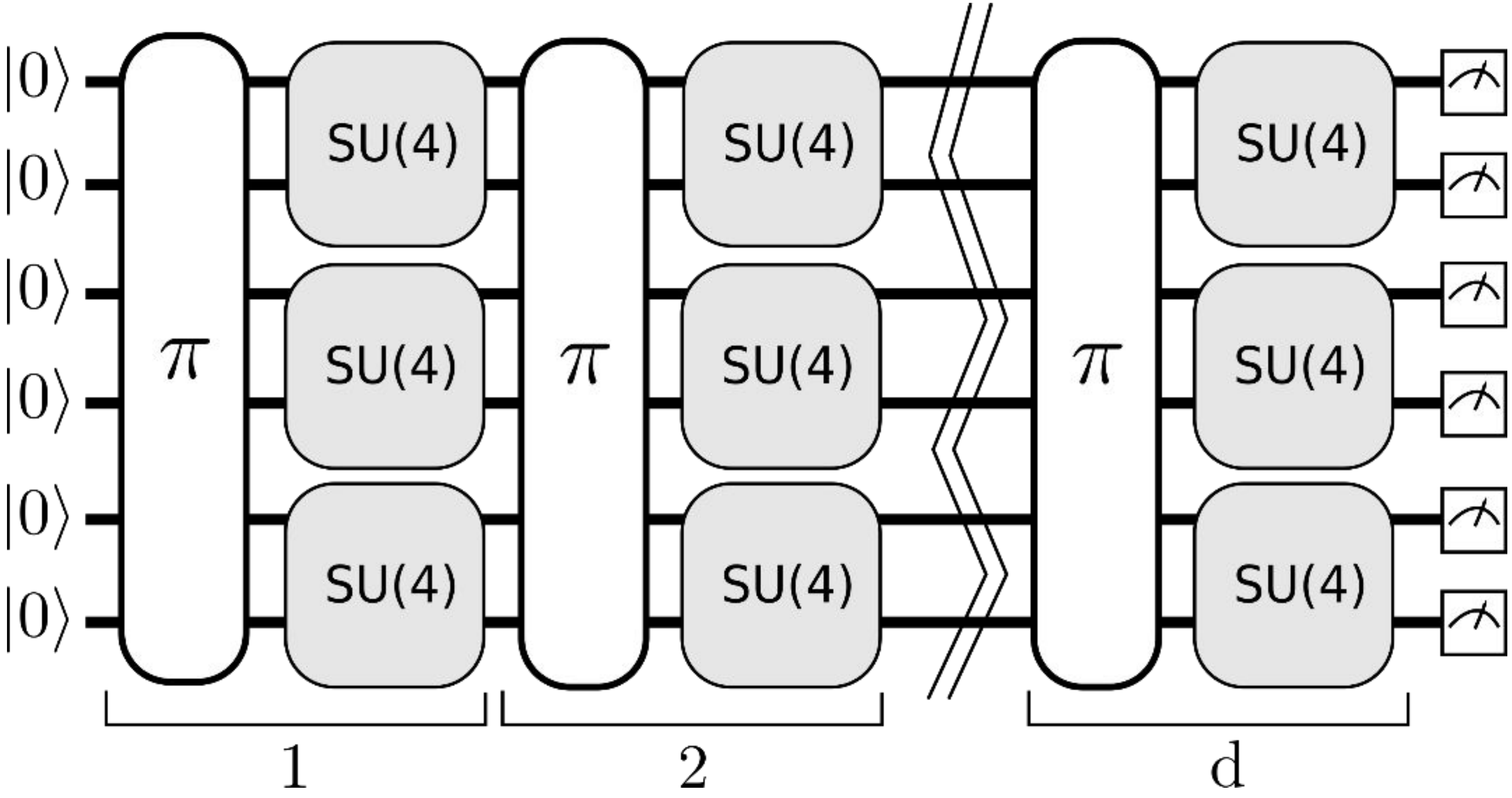
✔ Easy to count number of qubits, harder to measure accuracy

✔ Hard to game. Many important features influence this metric (e.g. qubit number, error rates, qubit topology, etc).

- Metric tells us how a system performs on a certain class of random square circuits, an important performance metric in the NISQ era.
  - × But many important applications are not square (i.e. need much greater depths than qubit number)

*Original definition would have given a quantum volume of n^2

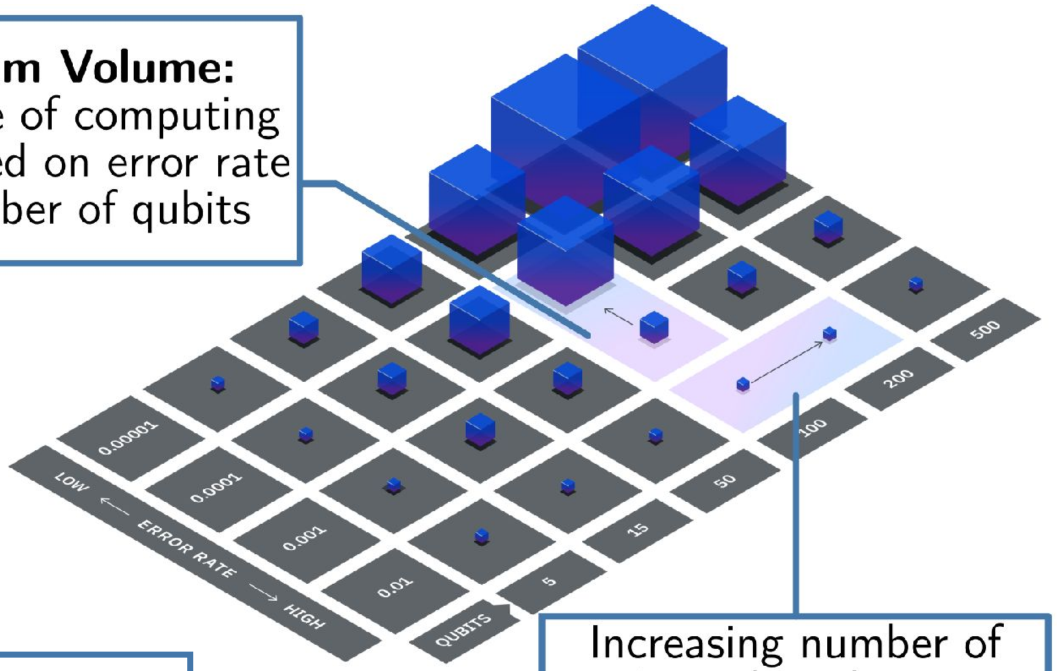
Quantum Volume Determination



- Number of arbitrary 2-qubit gates that can be implemented in a row before “failure”
- Failure needs to be defined
- Ability to implement gate between arbitrary 2 qubits will depend on connectivity of the qubits
- Hence Quantum Volume will be dependent on connectivity

Benchmarking Quantum Computers

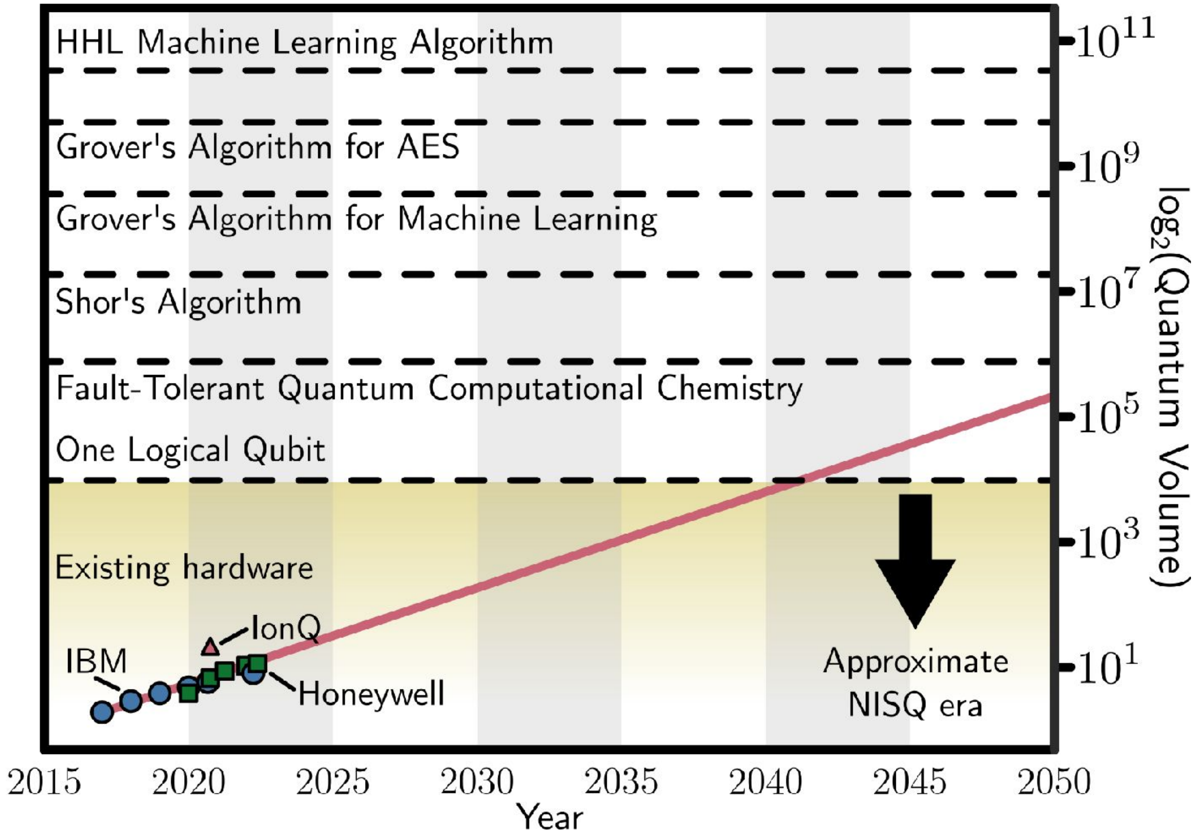
Quantum Volume:
A measure of computing power based on error rate and number of qubits



Increasing number of noisy qubits does not always increase computing power

$$V_Q = 2^{\min(n,d)}$$

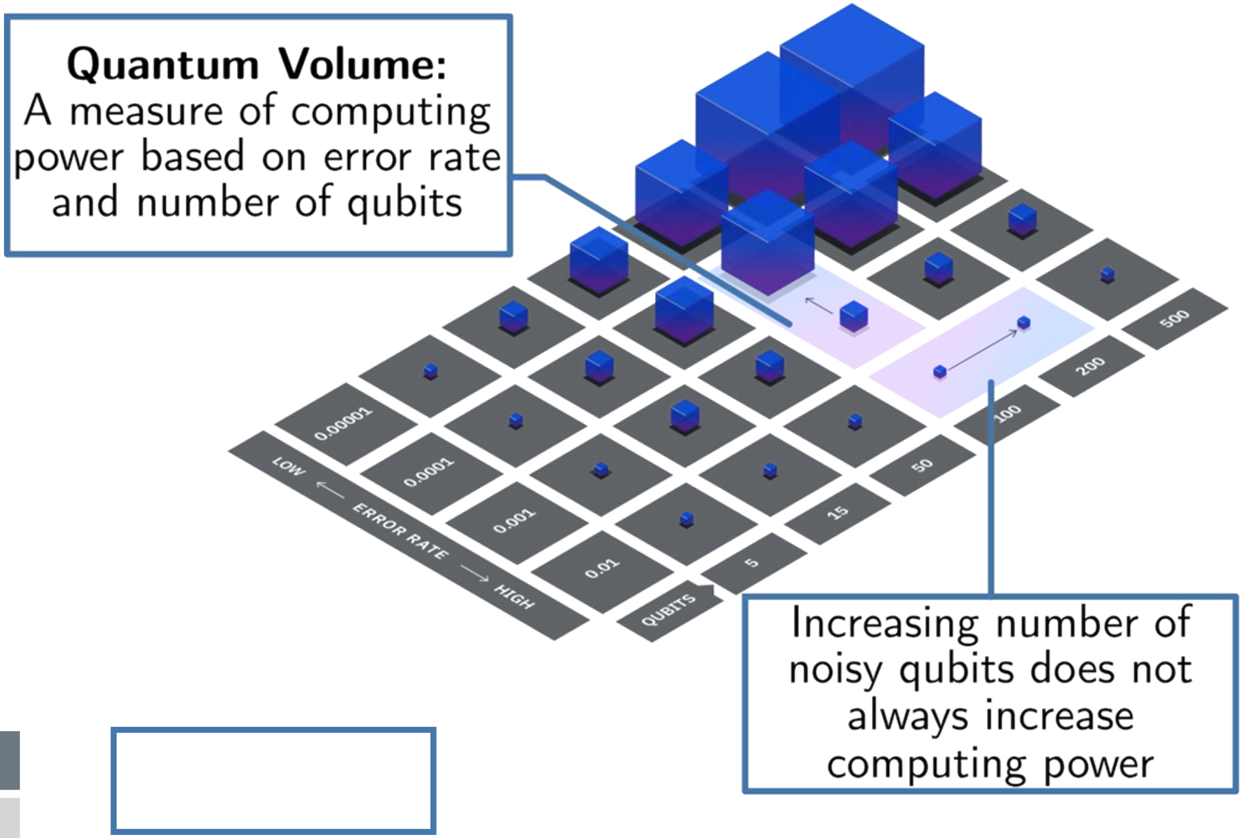
- Hardware independent
- Should be straightforward to verify
- Is the requirement for arbitrary SU(4) too stringent?
  - Is the gate set to broad since error corrected gate sets are much more limited?
  - Is arbitrary connectivity too hard?



Industry	Current V_Q	Planned
IonQ	4 million	
Honeywell	4,096	10x every year
IBM	256	Double every year
Google	256 (estimate)	1 million low noise qubits by 2030

Current: Quantum Volume

- Metric proposed by IBM. IBM, Honeywell, and IonQ have all announced quantum volumes for their systems
- Describes largest random square circuit that can be computed with high fidelity (but most applications will require greater circuit depth than width)
- Relatively small improvements can lead to drastically larger numbers due to the exponentiation in the definition, overemphasizing differences in capability
- Error correction is not part of model (applicable to current NISQ hardware only)



Quantum Volume: A measure of computing power based on error rate and number of qubits

Increasing number of noisy qubits does not always increase computing power

Industry	Current V_Q	Planned
IonQ	~4 million*	
Honeywell	4,096	10x every year
IBM	256	Double every year
Google	256*	1 million low noise qubits by 2030

*Estimate only, not benchmarked

Current: Quantum Volume

- Metric proposed by IBM. IBM, Honeywell, and IonQ have all announced quantum volumes for their systems
- Describes largest random square circuit that can be computed with high fidelity (but most applications will require greater circuit depth than width)
- Relatively small improvements can lead to drastically larger numbers due to the exponentiation in the definition, overemphasizing differences in capability
- Error correction is not part of model (applicable to current NISQ hardware only)

Solution: Generalized Volumetric Metrics

See arXiv:2207.02315

Solution: Integrate model of error correction into metric

Industry	Current V_Q	Planned
IonQ	~4 million*	
Honeywell	4,096	10x every year
IBM	256	Double every year
Google	256*	1 million low noise qubits by 2030

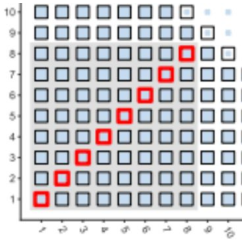
*Estimate only, not benchmarked

"An Improved Volumetric Metric for Quantum Computers via more Representative Quantum Circuit Shapes." arXiv:2207.02315 (2022)

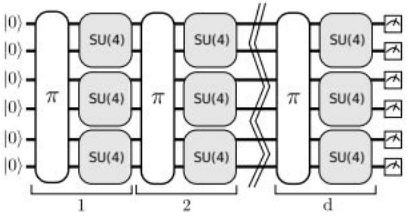
Thrust 1: Realistic "Volumetric" Shapes

Current Metric: Physical Systems -> Quantum Volume

- Measures the largest "square circuit" which a QC can successfully run, using a model circuit
- Square circuits are not representative of real quantum algorithms



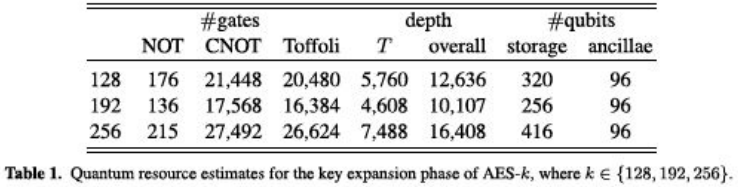
From Blume-Kohout and Young, arXiv:1904.05546v4



From Cross et.al., arXiv:1811.12926v2

Current Metric: Algorithms -> Qubits and T-Depth

- Theoretical number of logical qubits and operations required for a given algorithm
- Assumes perfect gates and qubits, error correction ignored



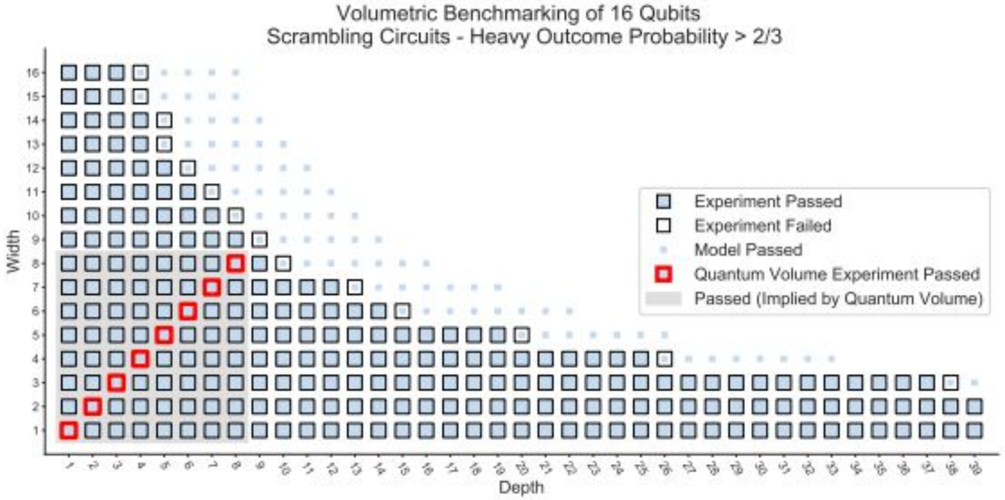
	#gates				depth		#qubits	
	NOT	CNOT	Toffoli	T	overall		storage	ancillae
128	176	21,448	20,480	5,760	12,636		320	96
192	136	17,568	16,384	4,608	10,107		256	96
256	215	27,492	26,624	7,488	16,408		416	96

Table 1. Quantum resource estimates for the key expansion phase of AES-k, where $k \in \{128, 192, 256\}$.

From Grassl et. al., arXiv:1512.04965v1

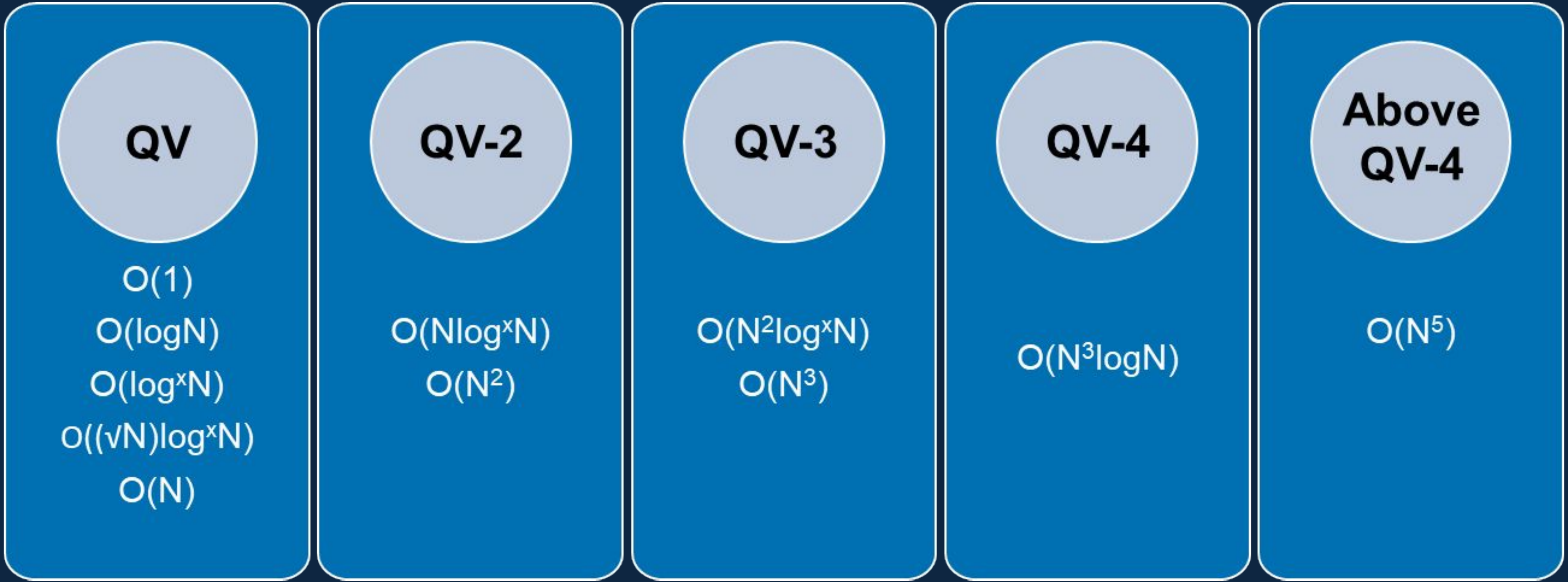
Our Goal: Merge the two metrics

- Use a "volumetric" framework to measure the effectiveness of physical systems
- Use one or more circuit "shapes" based on actual algorithm requirements
- Perform a literature survey to try and group common circuit shapes together



From Blume-Kohout and Young, arXiv:1904.05546v4

Quantum Volume Class x

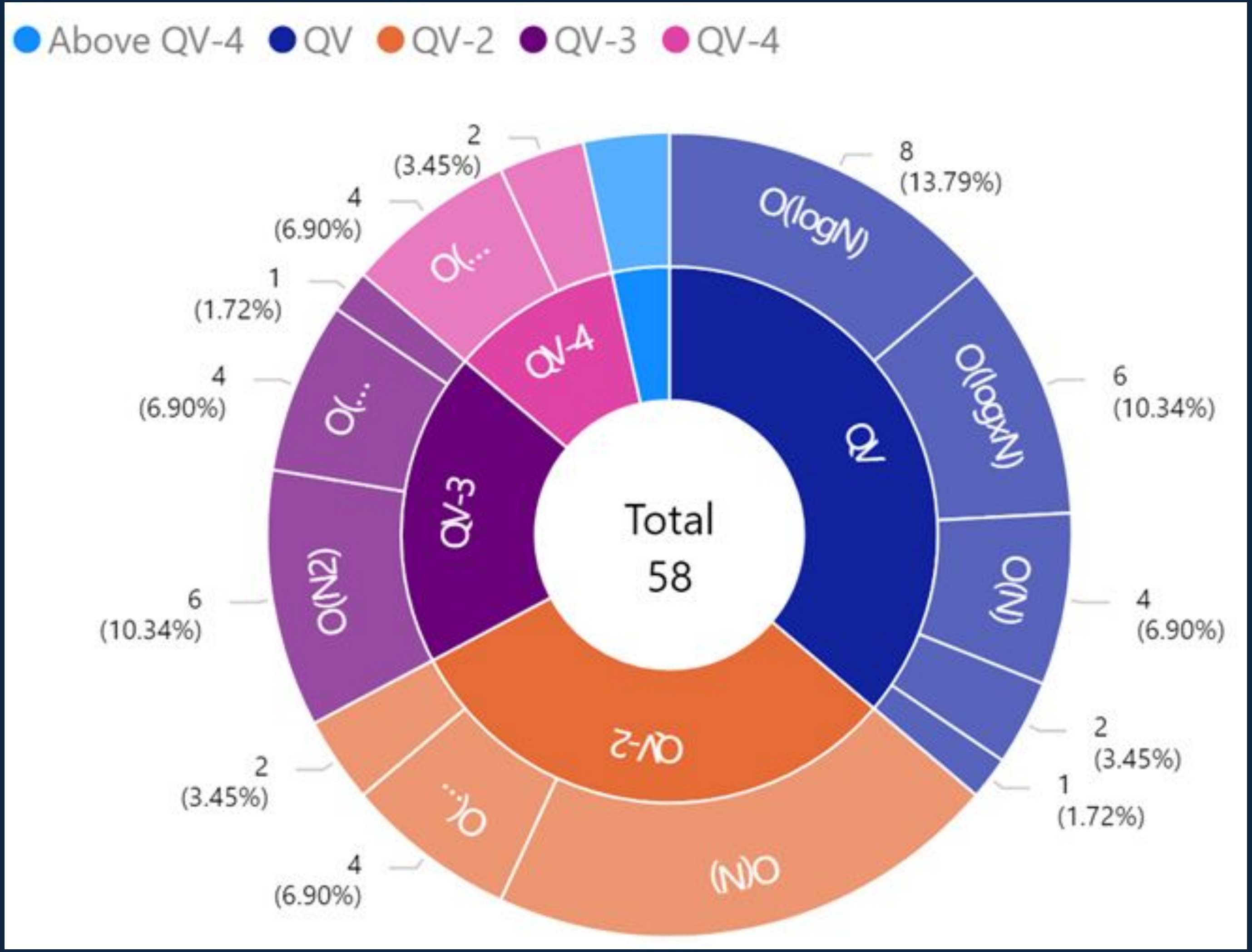


QV	QV-2	QV-3	QV-4	Above QV-4
$O(1)$ $O(\log N)$ $O(\log^x N)$ $O((\sqrt{N})\log^x N)$ $O(N)$	$O(N\log^x N)$ $O(N^2)$	$O(N^2\log^x N)$ $O(N^3)$	$O(N^3\log N)$	$O(N^5)$

- QV – $N \times N$
- QV-2 – $N \times N^2$
- QV-3 – $N \times N^3$
- QV-4 – $N \times N^4$

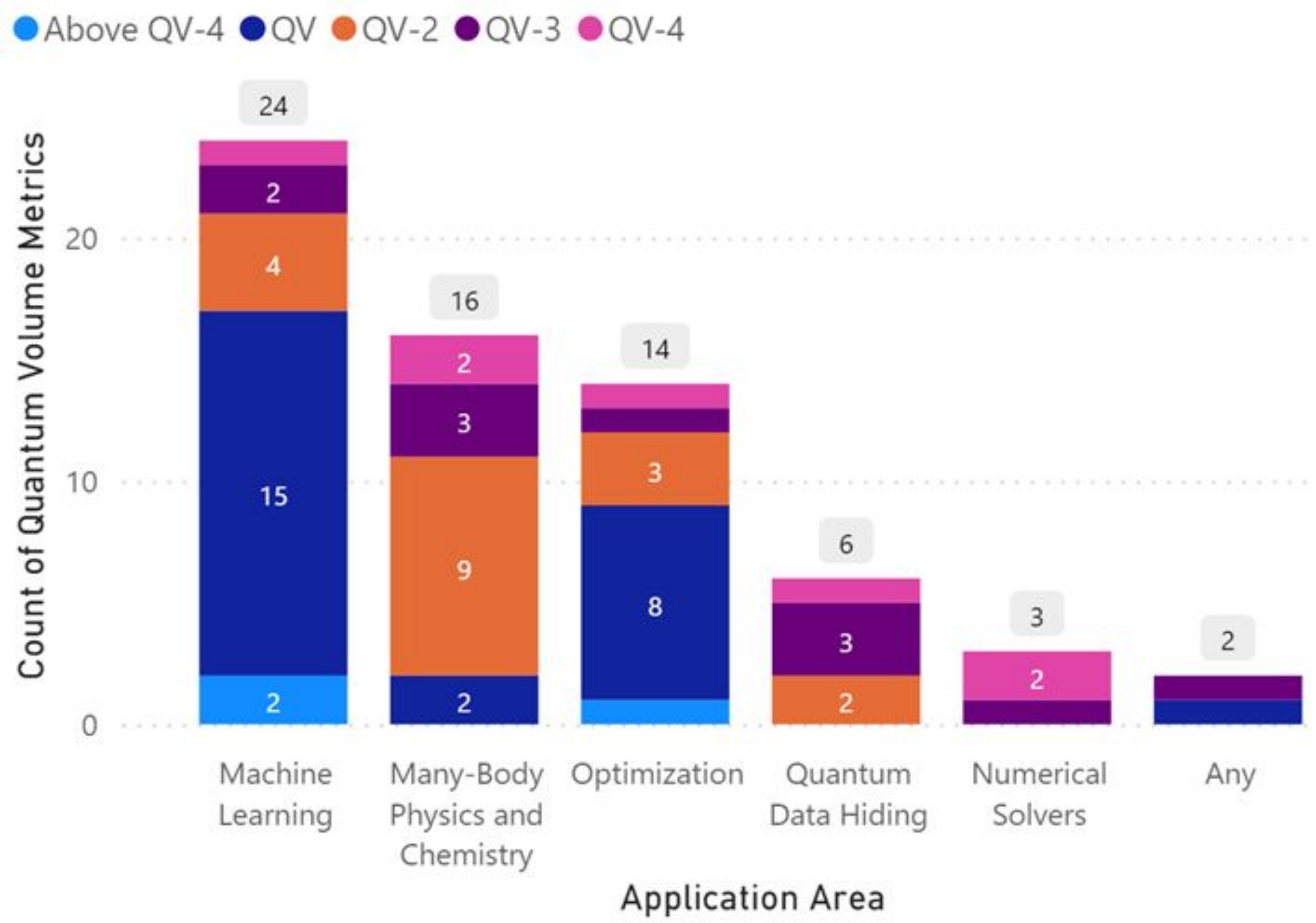
MITRE

Algorithm Survey



- Only a minority of quantum algorithms require only the same number or (or less) gates than number of qubits.
- Most quantum algorithms require far more gates than qubits. Therefore, a more metric that gives more weight to the number of gates is necessary.
- How much more to weight gates than qubits depends on the algorithms

Which Algorithms for Which QV?

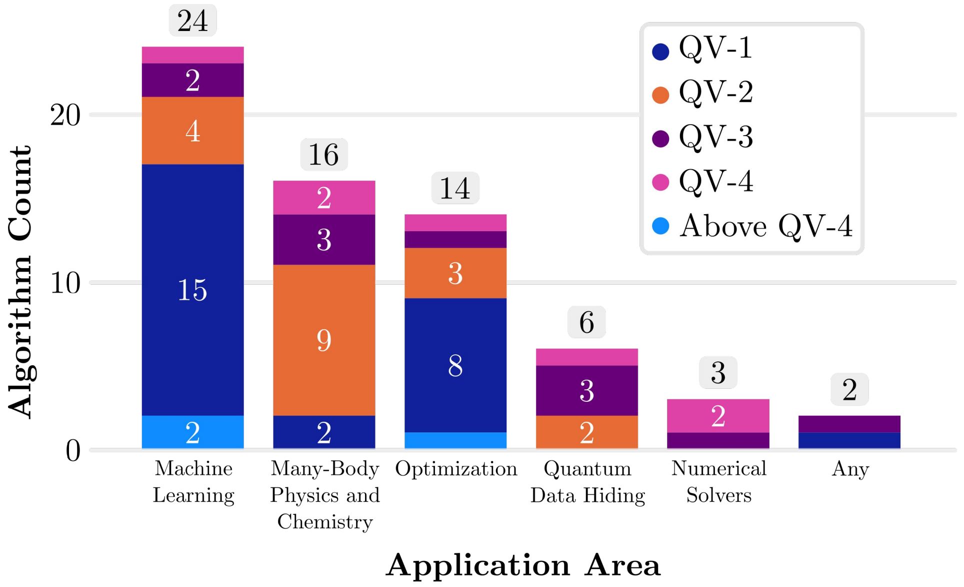


- Machine learning algorithms generally require the least number of gates per qubit.
- Numerical solvers (such as Shor’s algorithm) are most gate intensive

“An Improved Volumetric Metric for Quantum Computers via more Representative Quantum Circuit Shapes.” arXiv:2207.02315 (2022)

Realistic “Volumetric” Shapes

-



QV-1	QV-2	QV-3	QV-4	Above QV-4
$\mathcal{O}(1)$ $\mathcal{O}(\text{polylog}(n))$ $\mathcal{O}(\sqrt{n}\,\text{polylog}(n))$ $\mathcal{O}(n\,\text{polylog}(n))$ $\mathcal{O}(n)$	$\mathcal{O}(n\,\text{polylog}(n))$ $\mathcal{O}(n^2)$	$\mathcal{O}(n^2\,\text{polylog}(n))$ $\mathcal{O}(n^3)$	$\mathcal{O}(n^3\log(n))$	$\mathcal{O}(n^5)$

Quantum Volume (QV) Classes	Algorithm Count (%)
Quantum Volume (QV)	21 (36.21%)
Quantum Volume Class 2 (QV-2)	18 (31.03%)
Quantum Volume Class 3 (QV-3)	11 (18.97%)
Quantum Volume Class 4 (QV-4)	6 (10.34%)
Above QV-4	2 (3.45%)

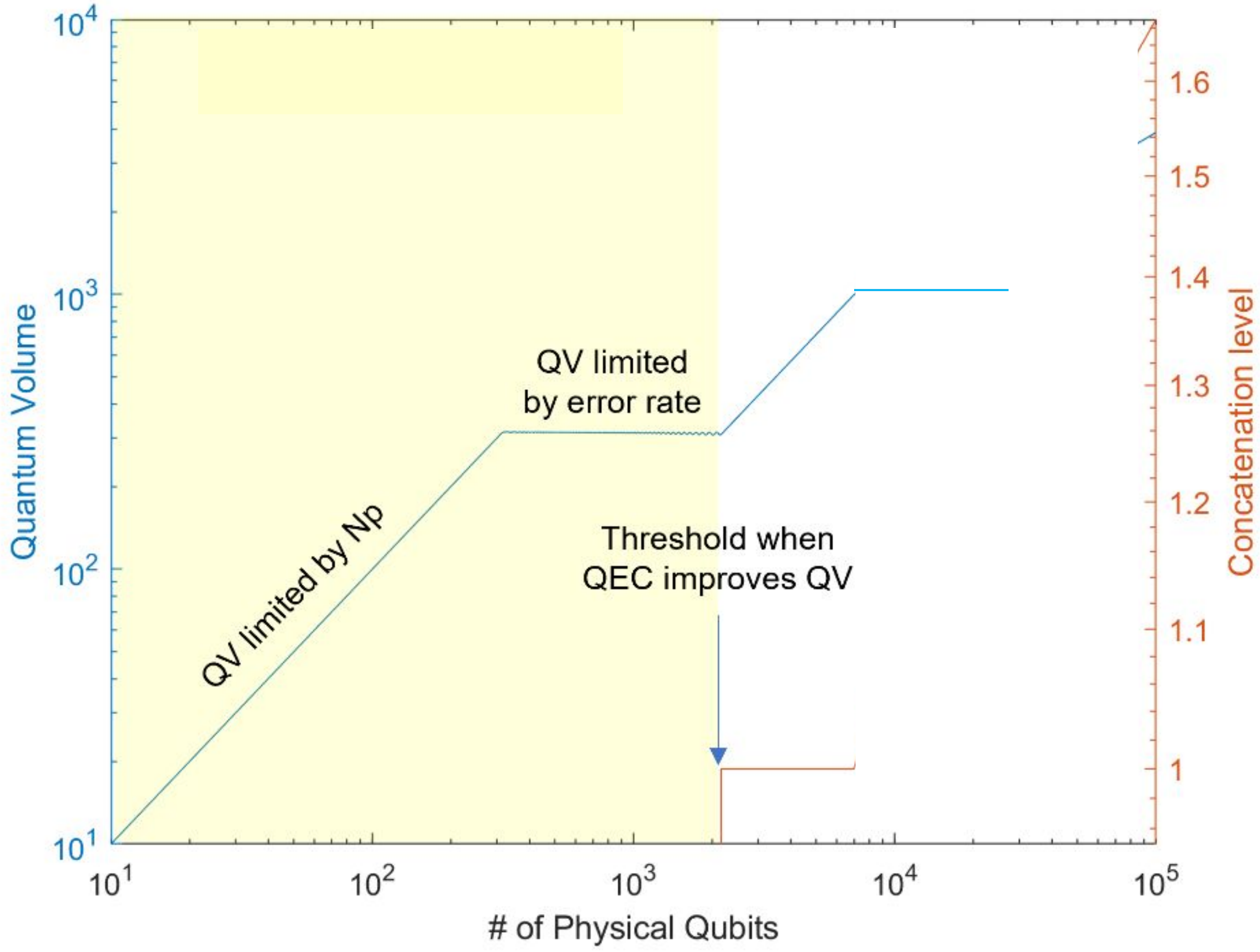
97% of algorithms

MITRE

What happens when we add quantum error correction?

- Widespread assumption that when error correction becomes viable the trend will skew towards higher quantum volume: is this true?
- Harder to perform gates on encoded (logical) qubits (more physical qubits to control)
- Harder to couple together logical qubits
- More subject to correlated errors?
- We need a LOT more physical qubits

QEC resource estimation

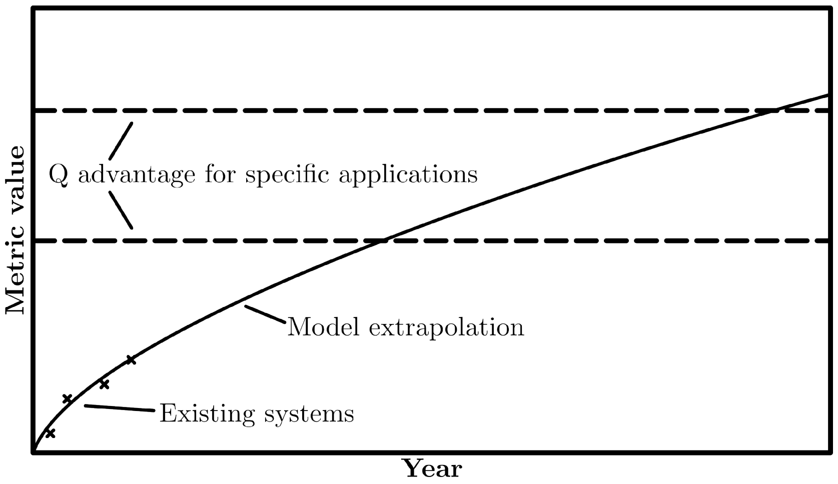


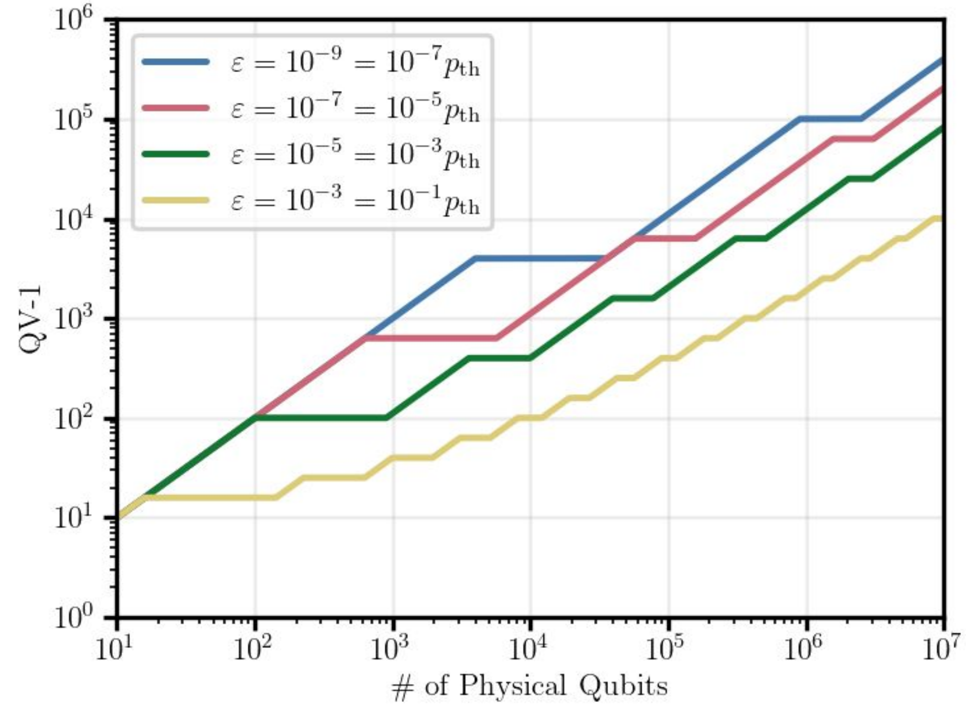
Quantum volume (QV-1, blue trace, left vertical axis) for a physical error rate 1e-5 for a concatenated Steane QEC approach, plotted versus the number of physical qubits.
The right vertical axis shows the associated concatenation level (orange trace). The shaded area represents parameter space for which QEC does not improve QV-1.

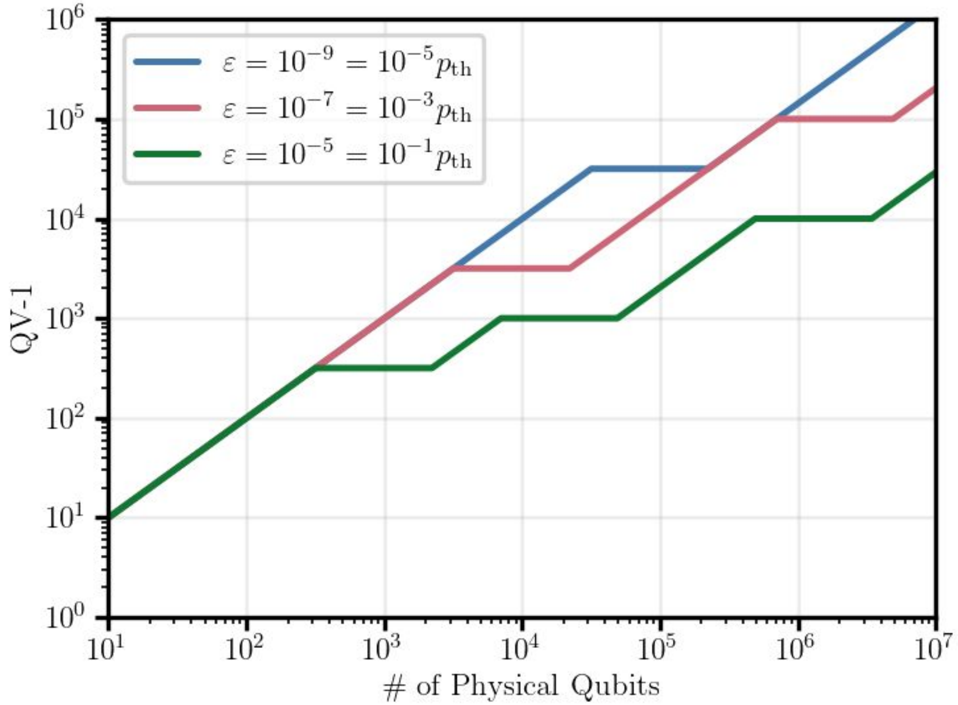
Quantum Volume vs. number of physical qubits (N_p):

1) QV limited by number of qubits (since error rate is low), QV = N_p
2) At 316 qubits (error rate = 1/ N_p^2) errors (10^{-5}) limit QV
3) Adding more qubits doesn’t help because we don’t yet have enough qubits for encoding into our desired QEC code (the [[7,1,3]] code
4) At 2212 qubits we can start encoding and again QV will increase until there are so many logical qubits that 10^{-5} error rate is too large for more gates to be performed
5) Repeat

QEC resource estimation







Conclusions and Further Thoughts

- Metrics and benchmarks provide a common language for evaluating quantum computers
    - Even the non-initiate can determine something from metrics and can demand certain performance without expert familiarity with how a metric is determined
    - Collaborative goals for academia, industry, and government
    - Hardware agnostic, allows identity of trends, etc.
- The advent of practical error correction influences the development and utilization of metrics and benchmarks
    - Benchmarks may be updated
    - Attempt to formulate metrics for this new era

Quantum Information Processing

Proudly Announces:

Geography based Topical Collection – Latin America

Submitted papers will undergo the same comprehensive peer-review as all QIP papers

Include supplementary information (including video) about your research interests

✔ Highlight research interests to entice collaborators and students in your geographic area!

Rafael Sotelo
Universidad de Montevideo, Montevideo, Uruguay
Renato Portugal
National Laboratory of Scientific Computing (LNCC) of the Ministry of Science and Technology, Brazil