Quantum AI

Quantum models and data through a precomputation lens

Jarrod McClean
@JarrodMcClean
Senior Staff Research Scientist

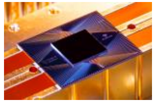
Google

Quantum machine learning & power of data

Computationally limited problems - Simple inputs, known computational procedure

x

Key to factor
Hamiltonian to simulate
...



Compute ~ n

Compute ~ exp(n)

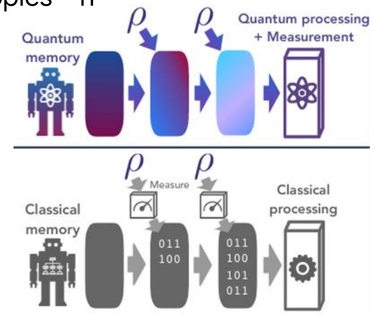
Data limited problems - Limited by availability of data, no computation possible to overcome lack of data

ρ (limited copies)

Transduced quantum state
Analog simulation state
Output of computation
...

Copies ~ n

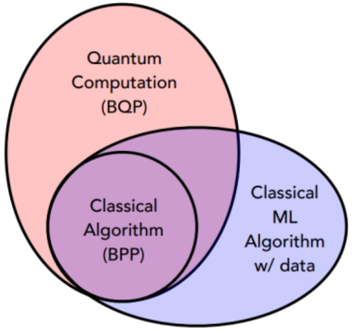
Compute ~ ?



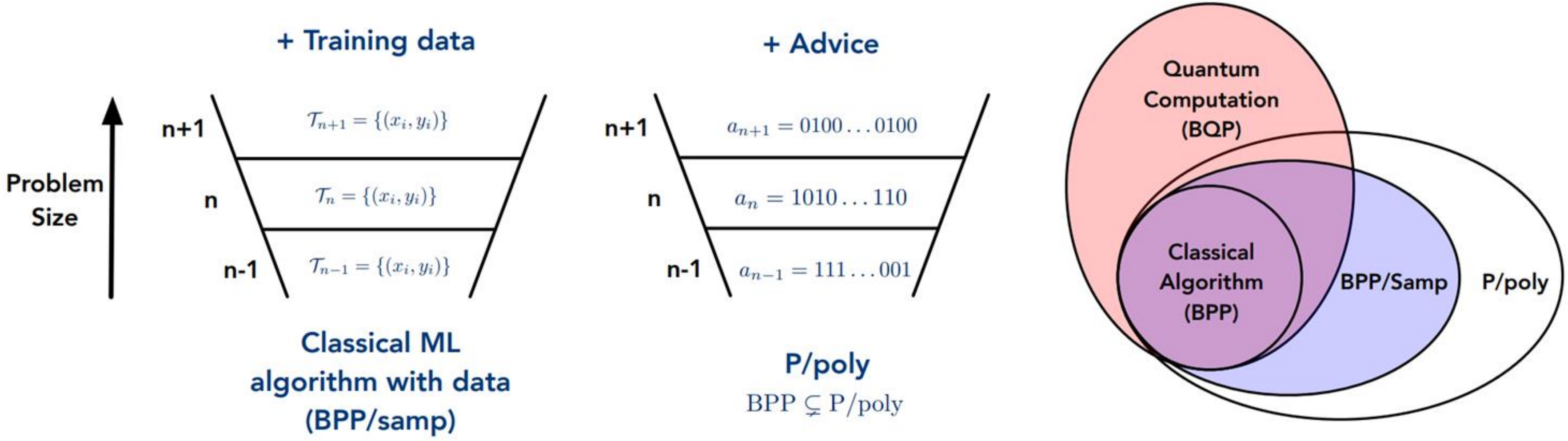
Copies ~ exp(n)

Compute ~ ?

Data assisted problems - Known computational procedure, complexity can change with available data (~advice)

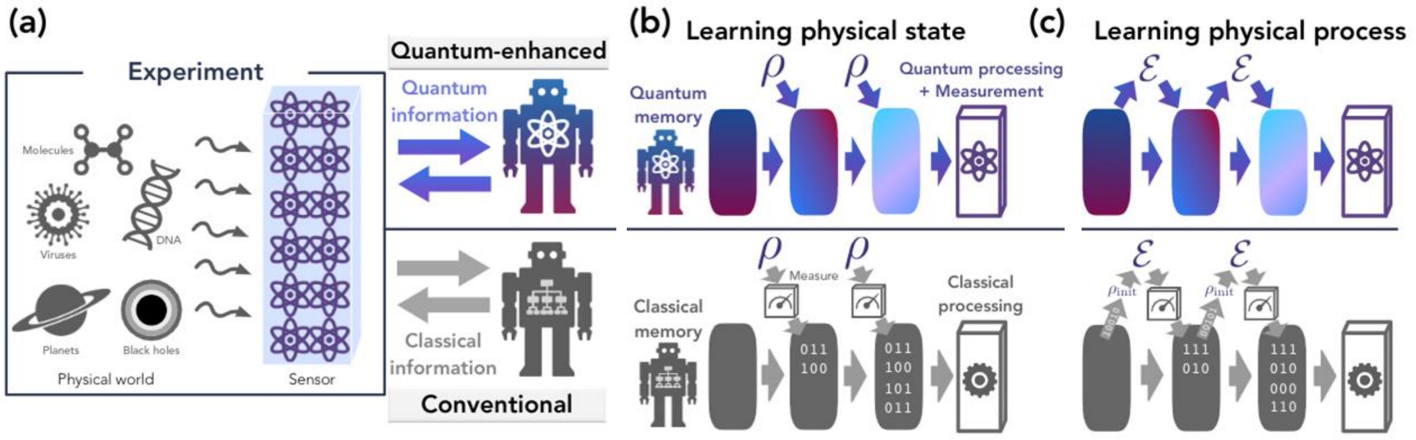


The power of data in quantum machine learning*



* Hsin-Yuan (Robert) Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, **Jarrod R. McClean** “Power of data in quantum machine learning” Nature Communications, Vol.12, No. 2631 (2021)

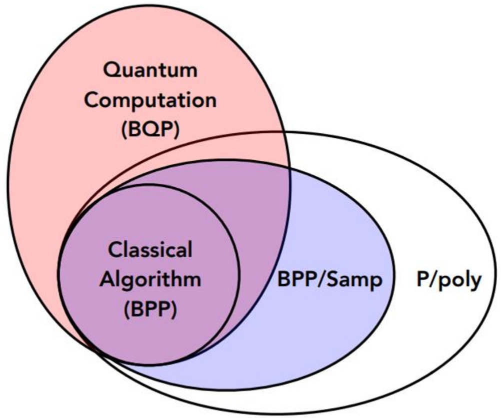
Quantum memory and quantum-enhanced experiments

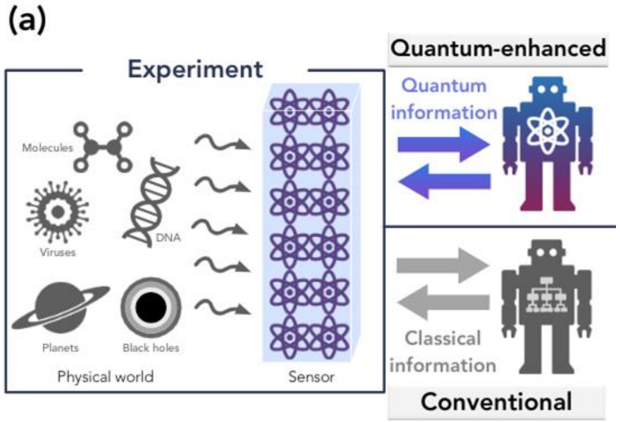


Quantum memory, **classical model**- Exponential advantage with exactly 2 copies on 2 different tasks and efficient classical compute

Quantum advantage in learning from experiments
Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean
Science 376, 6598 (2022)

Do we need fully quantum models for quantum data?





With and without quantum memory, we find an advantage using classical data extracted from quantum computations with classical models. One might be tempted to call these **semi-classical learning models**.

Which models we might use are **fully quantum models** and what are they good for?

The cost of a quantum algorithm

We usually try to quantify the most limited resource

- Gate complexity
- T-gates in fault-tolerance
- Native 2-qubit gates on a NISQ machine
- Circuit repetitions
- Circuit depth

The cost of a quantum algorithm

We usually try to quantify the most limited resource

- Gate complexity
- T-gates in fault-tolerance
- Native 2-qubit gates on a NISQ machine
- Circuit repetitions
- Circuit depth

What if the limited resource is the wall-clock time...

The cost of a quantum algorithm

We usually try to quantify the most limited resource

- Gate complexity
- T-gates in fault-tolerance
- Native 2-qubit gates on a NISQ machine
- Circuit repetitions
- Circuit depth

What if the limited resource is the wall-clock time...

...after a problem is fully specified?

For example...

Real-time scheduling

- “Ride-share” companies routing drivers
- Airlines responding to inclement weather

Time-sensitive decisions in finance

- Re-evaluating trading strategies based on breaking news
- Ticker-tape trading

Can we ever benefit by doing some work ahead of time?

Magic state resevoir

What if we precompute a large number of magic states?

- In many architectures, we will need them to implement non-Clifford gates.
- Producing them is expensive.
- Let's do it ahead of time.

Magic state resevoir

What if we precompute a large number of magic states?

- In many architectures, we will need them to implement non-Clifford gates.
- Producing them is expensive.
- Let's do it ahead of time.

Unfortunately

- Storing them is also expensive when quantum memory involves active error correction.
- Most interesting algorithms require far more magic states than qubits.
- It might be better to use extra space for parallelization instead of resource state storage.

Quantum precomputation

Quantum precomputation

Algorithms without precomputation

$$\mathcal{A} : x, y, \rho \to z, \sigma$$

- Know $\mathcal{A}$ and x ahead of time
- But not y and ρ

Quantum precomputation

Algorithms without precomputation

$$\mathcal{A} : x, y, \rho \rightarrow z, \sigma$$

- Know $\mathcal{A}$ and x ahead of time
- But not y and ρ
- Wait for y and ρ, then execute $\mathcal{A}$

Quantum precomputation

Algorithms without precomputation

$$\mathcal{A} : x, y, \rho \to z, \sigma$$

- Know $\mathcal{A}$ and x ahead of time
- But not y and ρ
- Wait for y and ρ, then execute $\mathcal{A}$

Algorithms with precomputation

- Implement the same map
- Use knowledge of $\mathcal{A}$ and x to prepare ahead of time

Quantum precomputation

Algorithms without precomputation

$$\mathcal{A} : x, y, \rho \to z, \sigma$$

- Know $\mathcal{A}$ and x ahead of time
- But not y and ρ
- Wait for y and ρ, then execute $\mathcal{A}$

Algorithms with precomputation

- Implement the same map
- Use knowledge of $\mathcal{A}$ and x to prepare ahead of time
- Use the prepared resources to execute $\mathcal{P}$

$$\mathcal{P} : \bar{x}(\mathcal{A}, x), |\Gamma(\mathcal{A}, x)\rangle, y, \rho \to z, \sigma$$

Cost in the precomputation model

We only count the cost of $\mathcal{P}$

- Use the same accounting for $\mathcal{A}$ and $\mathcal{P}$
- For example, the number of gates

Cost in the precomputation model

We only count the cost of $\mathcal{P}$

- Use the same accounting for $\mathcal{A}$ and $\mathcal{P}$
- For example, the number of gates

We “charge for” identity operations

- Count single-qubit identity operations as gates
- Reflects the actual cost model of many devices
- Implicitly adds a cost to making $|\Gamma(\mathcal{A}, x)\rangle$ large

Cost in the precomputation model

We only count the cost of $\mathcal{P}$

- Use the same accounting for $\mathcal{A}$ and $\mathcal{P}$
- For example, the number of gates

We “charge for” identity operations

- Count single-qubit identity operations as gates
- Reflects the actual cost model of many devices
- Implicitly adds a cost to making $|\Gamma(\mathcal{A}, x)\rangle$ large

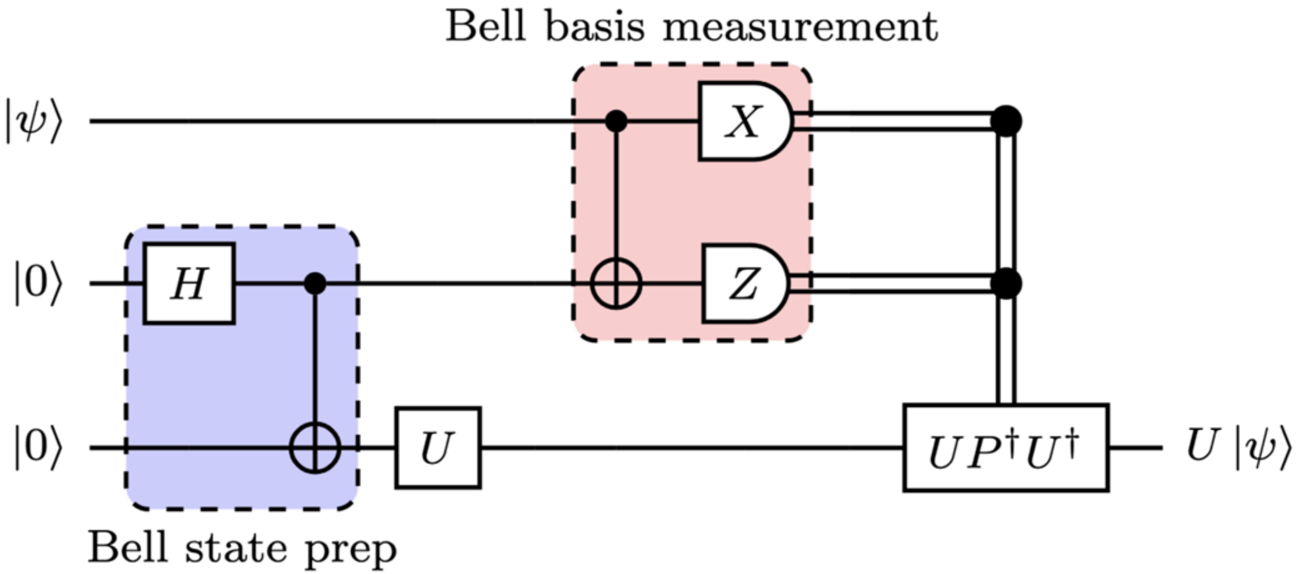
$\bar{x}(\mathcal{A}, x)$ and $|\Gamma(\mathcal{A}, x)\rangle$ must be efficient to prepare

- Polynomial quantum and classical complexity
- Different from quantum advice

A simple example

A simple example

Gate teleportation



A simple example

Teleporting an n-qubit Clifford U

$$\mathcal{A} : \rho \to U \rho U^\dagger$$

A simple example

Teleporting an n-qubit Clifford U

$$\mathcal{A} : \rho \to U \rho U^\dagger$$

- Apply U ahead of time to some Bell pairs to get $|\Gamma(U)\rangle$
- Consume $|\Gamma(U)\rangle$ in $\mathcal{O}(1)$ time to apply U up to a Pauli correction

A simple example

Teleporting an n-qubit Clifford U

$$\mathcal{A} : \rho \to U\rho U^\dagger$$

- Apply U ahead of time to some Bell pairs to get $|\Gamma(U)\rangle$
- Consume $|\Gamma(U)\rangle$ in $\mathcal{O}(1)$ time to apply U up to a Pauli correction

Comparing the cost

- $\mathcal{O}(n^2)$ gates the normal way
- $\mathcal{O}(n)$ gates in the precomputation model

Are there more interesting examples?

Density matrix exponentiation[2]

[2] Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. Nat. Phys. 10, 631–633 (2014).

Density matrix exponentiation[2]

Encoding a Hamiltonian in a state

- Consume copies of ρ to approximate $e^{-it\rho}$

[2] Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. Nat. Phys. 10, 631–633 (2014).

Density matrix exponentiation[2]

Encoding a Hamiltonian in a state

- Consume copies of ρ to approximate $e^{-it\rho}$
- Requires $\approx t^2/\epsilon$ copies for error ϵ

[2] Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. Nat. Phys. 10, 631–633 (2014).

Density matrix exponentiation[2]

Encoding a Hamiltonian in a state

- Consume copies of ρ to approximate $e^{-it\rho}$
- Requires $\approx t^2/\epsilon$ copies for error ϵ
- Useful for us if $\mathrm{tr}\,[H + c\mathbb{I}]$ not too large

[2] Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. Nat. Phys. 10, 631–633 (2014).

Density matrix exponentiation[2]

Encoding a Hamiltonian in a state

- Consume copies of ρ to approximate $e^{-it\rho}$
- Requires $\approx t^2/\epsilon$ copies for error ϵ
- Useful for us if $\mathrm{tr}\,[H + c\mathbb{I}]$ not too large

Precomputing reflection operators

$$R = \mathbb{I} - 2\,|b\rangle\langle b| = e^{-i\pi|b\rangle\langle b|}$$

- We only need $\tilde{\mathcal{O}}(1/\epsilon)$ copies of $|b\rangle$ to implement R up to error ϵ

[2] Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. Nat. Phys. 10, 631–633 (2014).

Density matrix exponentiation[2]

Encoding a Hamiltonian in a state

- Consume copies of ρ to approximate $e^{-it\rho}$
- Requires $\approx t^2/\epsilon$ copies for error ϵ
- Useful for us if $\mathrm{tr}\left[H + c\mathbb{I}\right]$ not too large

Precomputing reflection operators

$$R = \mathbb{I} - 2\left|b\right\rangle\!\left\langle b\right| = e^{-i\pi\left|b\right\rangle\!\left\langle b\right|}$$

- We only need $\tilde{\mathcal{O}}(1/\epsilon)$ copies of $\left|b\right\rangle$ to implement R up to error ϵ
- We can implement q reflections using $\tilde{\mathcal{O}}(q^2/\epsilon)$ copies
- The complexity of preparing $\left|b\right\rangle$ doesn't matter

---

[2] Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. Nat. Phys. 10, 631–633 (2014).

Example: linear systems of equations[3,4]

Quantum algorithms for linear systems

[3] Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. Phys. Rev. Lett. 103, 150502 (2009).
[4] Childs, A. M., Kothari, R. & Somma, R. D. Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision. SIAM J. Comput. 46, 1920–1950 (2017).

Example: linear systems of equations[3,4]

Quantum algorithms for linear systems

- Given oracle access to A, $|b\rangle$, prepare $|x\rangle \propto A^{-1}|b\rangle$

[3] Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. Phys. Rev. Lett. 103, 150502 (2009).
[4] Childs, A. M., Kothari, R. & Somma, R. D. Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision. SIAM J. Comput. 46, 1920–1950 (2017).

Example: linear systems of equations[3,4]

Quantum algorithms for linear systems

- Given oracle access to A, $|b\rangle$, prepare $|x\rangle \propto A^{-1} |b\rangle$
- Requires $\tilde{O}(\kappa)$ queries to A and $|b\rangle$, where κ is the condition number
- Uses state preparation and reflection about $|b\rangle$

[3] Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. Phys. Rev. Lett. 103, 150502 (2009).
[4] Childs, A. M., Kothari, R. & Somma, R. D. Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision. SIAM J. Comput. 46, 1920–1950 (2017).

Example: linear systems of equations[3,4]

Quantum algorithms for linear systems

- Given oracle access to A, $|b\rangle$, prepare $|x\rangle \propto A^{-1}|b\rangle$
- Requires $\tilde{O}(\kappa)$ queries to A and $|b\rangle$, where κ is the condition number
- Uses state preparation and reflection about $|b\rangle$

In the precomputation model

- We can precompute the reflections about $|b\rangle$.

[3] Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. Phys. Rev. Lett. 103, 150502 (2009).
[4] Childs, A. M., Kothari, R. & Somma, R. D. Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision. SIAM J. Comput. 46, 1920–1950 (2017).

Example: linear systems of equations[3,4]

Quantum algorithms for linear systems

- Given oracle access to A, $|b\rangle$, prepare $|x\rangle \propto A^{-1}|b\rangle$
- Requires $\tilde{O}(\kappa)$ queries to A and $|b\rangle$, where κ is the condition number
- Uses state preparation and reflection about $|b\rangle$

In the precomputation model

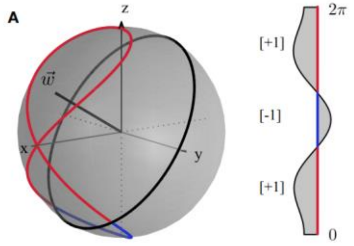
- We can precompute the reflections about $|b\rangle$.
- Same dependence on A, but requires $\mathcal{O}(\kappa^2/\epsilon)$ copies of $|b\rangle$

[3] Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. Phys. Rev. Lett. 103, 150502 (2009).
[4] Childs, A. M., Kothari, R. & Somma, R. D. Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision. SIAM J. Comput. 46, 1920–1950 (2017).

What models require a "quantum program"

$$\mathcal{A} : x, y, \rho, \rho_t \to z, \sigma \qquad \mathcal{P} : \bar{x}(\mathcal{A}, x, \rho_t), |\Gamma(\mathcal{A}, x, \rho_t)\rangle, y, \rho \to z, \sigma$$

Another situation where we can allow quantum data to define a model is a kernel method



(Quadratic) Quantum kernel[1]

$$K(x, z) = |\langle \Phi(x) | \Phi(z) \rangle|^2$$

(training data & test data in quantum state form)

Projected quantum kernel[2]

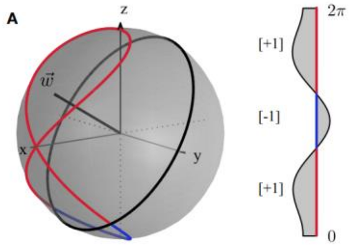
Classical shadows → nonlinear function → K(x, z)

(training data just classical bit strings, test data first reduced to bit strings)

Both quadratic and projected kernel efficiently learn discrete log problem (given right quantum embedding)

1 - Havlicek et al, Nature. vol. 567, pp. 209-212 (2019)
2 - Huang et al, Nature Communications 12, Article number: 2631 (2021)

Fully quantum vs semi-classical kernel methods



(Quadratic) Quantum kernel[1]

$$K(x, z) = |\langle \Phi(x) | \Phi(z) \rangle|^2$$

(training data & test data in quantum state form)

Projected quantum kernel[2]

Classical shadows → nonlinear function → K(x, z)

(training data just classical bit strings, test data first reduced to bit strings)

Classical shadows on two unknown states cannot be used to efficiently estimate fidelity [Anshu 22] hence there must be functions quadratic quantum kernel computes the projected quantum kernel cannot

This separation **may** vanish if training and test data is built from polynomial complexity circuits [Zhao 23]

On the other hand, no linear observable can determine the topological phase [Huang 22], while projected kernels can, so there *may* be some functions the projected kernel can naturally estimate the quadratic kernel cannot

[Huang 22] - "Provably efficient machine learning for quantum many-body problems"
Huang, Kueng, Torlai, Albert, Preskill - Science 377.6613 (2022)
[Anshu 22]- "Distributed quantum inner product estimation"
Anshu, Landau, Liu (STOC 2022), pp. 44-51
[Zhao 23] - "Learning quantum states and unitaries of bounded gate complexity",
Zhao, Lewis, Kannan, Quek, Huang, Caro arXiv:2310.19882 (2023)

A case where currently quantum models matter

Given N copies of a n qubit quantum state ρ, when is it **efficient** to determine M Pauli operators e.g. $P_j = X \otimes Z \otimes \ldots \otimes Y$

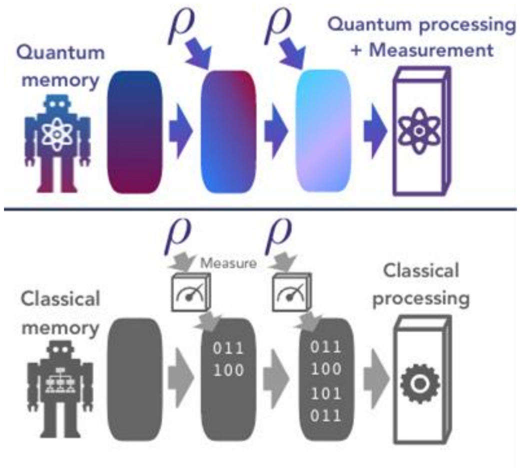
Efficient -
Number of copies N scales like ~log M poly(n),
Computation scales like ~M log M poly(n)

Classical shadows (single copy, semi-classical) -
k-local Pauli's only, magnitudes, and signs

2-copy bell sketches (two copy memory, semi-classical) -
All Pauli's, magnitudes only, efficient methods for signs **unknown**

Small quantum memory (log copy memory, quantum) -
All Pauli's, efficient magnitudes and signs

Open question - Does a computationally efficient semi-classical model with quantum memory exist for this task or is it essentially fully quantum?



Huang, Kueng, and Preskill. "Information-theoretic bounds on quantum advantage in machine learning." Physical Review Letters 126.19 (2021): 190505.

Making a gamble with classical data

Quantum data is interesting for future discovery of the universe (recall the impact of CCD cameras on telescopes - see "The Perfect Theory"), but most data we work with today, even from quantum systems, seems **classical.**

There are a few pieces of evidence that QC might help for **classical data** (sampling hard distributions, learning problems based on discrete log, linear algebra routines, ...) but a lot of pieces of evidence that it will be hard to achieve in practice

Immediate path for everyone opening Nielsen and Chuang

1. Stick N features of classical data into Log N qubits
2. Read about Holevo's bound limiting you to Log N bits of information out, get sad

$$\vec{x} \rightarrow |x\rangle = \sum_k^N x^k |k\rangle$$

Naive amplitude encoding + expected values limits you to quadratic functions on data - pretty weak models

Rotation based encoding and calling data multiple times (Data re-uploading) can get trig functions and higher degree polynomials - but can be hard to design in some cases [Schuld et al 2020]

In fault tolerance, about as easy to encode data in a way that allows non-linear functions over compact intervals for each feature by applying a unitary to the state

$$\vec{x} \rightarrow |x\rangle = \sum_j^M \sum_k^N e^{-2\pi i j x^k / M} |j\rangle |k\rangle$$

Quantum AI

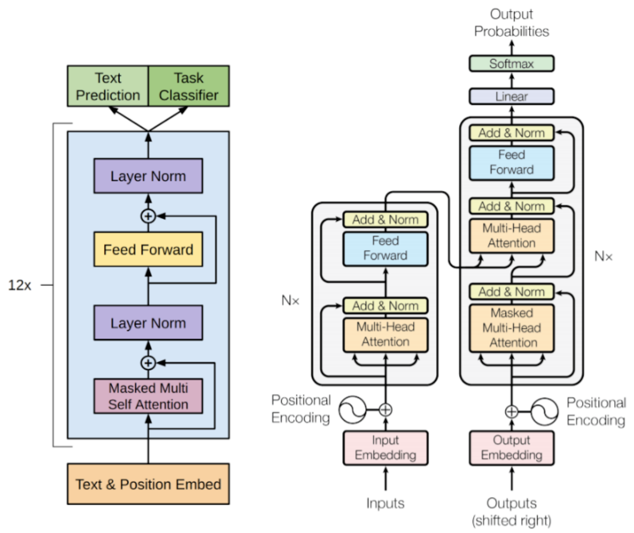
Even when amplitude encoding or feature-by-feature is enough - loading data always scales with ~N, many successful classical algorithms scale like N or < N – advantage lost?

Current premier models

Bard - 137 billion parameters (~10^{11})

GPT 3 - 175 billion parameters (~10^{11})

GPT 4 - 1.76 Trillion parameters (~10^{12}) (Speculated)



When # of key parameters like weights or # of data points ~ 10^{12} any scalings worse than linear can be catastrophic and determines architecture / algorithm success or failure.

The same may be true in the quantum case independent of the large Hilbert space dimension

Quantum communication complexity

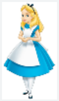
x y

Alice has x, Bob has y, want to compute f(x,y), and the cost is only counted in terms of bits or qubits exchanged

$$\vec{x} \to |x\rangle = \sum_{k}^{N} x^k |k\rangle$$

In spite of Holevo's bound stating n qubits contain n bits of information, exponential quantum communication advantages are known. Sending log n qubits in place of n bits

Raz problem



Inputs	$\vec{x}$	$M = UZ_0U \to \{M_0, M_1\}$	V
Size	N	N x N	N x N

Quantum - Log N qubits
Classical - Ω(√N) bits

Output 0(1) if Vx is close to $M_0(M_1)$

Quantum AI

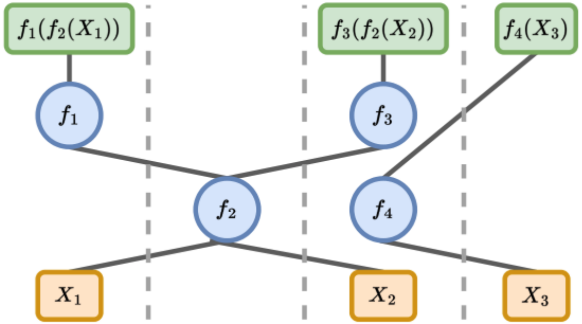
Some existing work in ML-like problems

Exponential communication advantaged shown for linear regression problem
Montanaro A, Shao C. Quantum communication complexity of linear regression arXiv:2210.01601 (2022)

Distributed quantum networks & models

Conjecture / opinion (controversial?) **-**

- The ability to make good quantum networks will be roughly coincident with really good quantum memories.
- Requirements beyond quantum memory - heralded transduction fidelity, rate, entanglement distillation - relatively modest.
- What we lack is **compelling, end-to-end costed out applications** to help motivate their development.



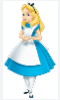
For an expressive class of functions, we find an exponential quantum communication advantage in the problem of inference and gradient determination.

Quantum AI

Gilboa , McClean “Exponential Quantum Communication Advantage in Distributed Learning”arXiv:2310.07136 (2023)

Small communication advantage with ‘simple’

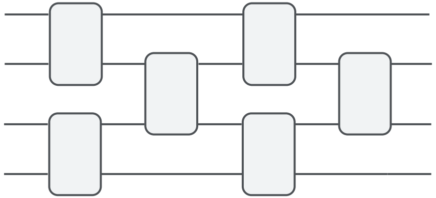
Raz problem



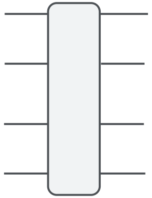
Inputs	$\vec{x}$	$M = UZ_0U \rightarrow \{M_0, M_1\}$	V
Size	N	N x N	N x N

Quantum - Log N qubits
Classical - Ω(√N) bits

Output 0(1) if Vx is close to $M_0(M_1)$



If M or V generated from polylog(N) ~ poly(n) and white box, it suffices to send circuit description.

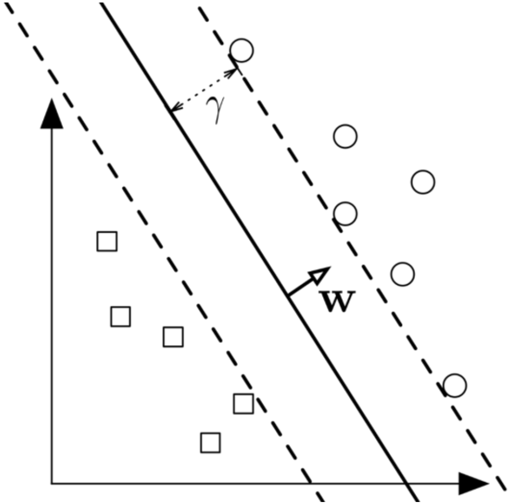


If M or V generated from polylog(N) ~ poly(n) and blackbox, it suffices to send Clifford classical shadows.

Lack of communication advantage, does not preclude computational advantage though.

- Suppose M or U was pseudo-random.
- Using black box + classical communication approach requires exponential classical computation under cryptographic assumptions

Problem structure can degrade advantage

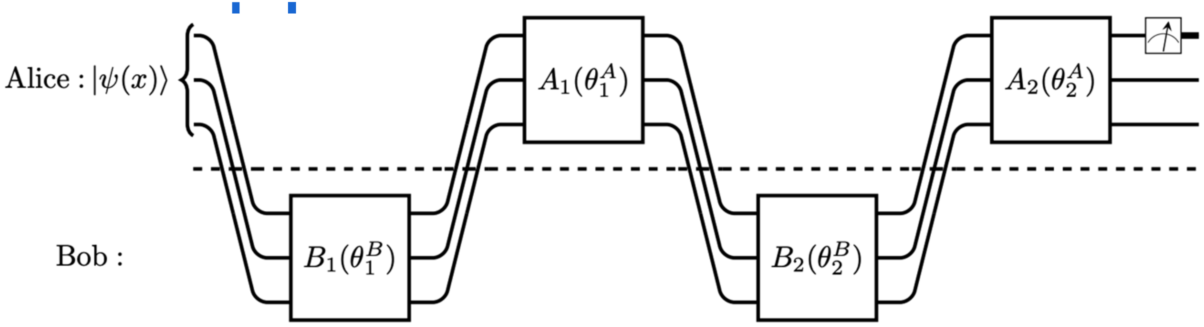


Problem 5 (Distributed Linear Classification). *Alice and Bob are given* $x, y \in S^N$, *with the promise that* $|x \cdot y| \geq \gamma$ *for some* $0 \leq \gamma \leq 1$. *Their goal is to determine the sign of* $x \cdot y$.

Lemma 7. *The quantum communication complexity of Problem 5 is* $\Omega\left(\sqrt{N/\max(1, \lceil \gamma N \rceil)}\right)$. *The randomized classical communication complexity of Problem 5 is* $O(\min(N, 1/\gamma^2))$.

If the classifier is really good ($\gamma \sim 1$), classical = quantum = O(1)

Exponential advantage in more general PQC models



$$|\varphi(\Theta, x)\rangle \equiv \left(\prod_{\ell=L}^{1} A_\ell(\theta_\ell^A, x) B_\ell(\theta_\ell^B, x)\right) |\psi(x)\rangle ,$$

$$\mathcal{L}(\Theta, x) \equiv \langle\varphi(\Theta, x)| \mathcal{P}_0 |\varphi(\Theta, x)\rangle ,$$

Problem 1 (Distributed Inference). *Alice and Bob each compute an estimate of $\langle\varphi| \mathcal{P}_0 |\varphi\rangle$ up to additive error ε.*

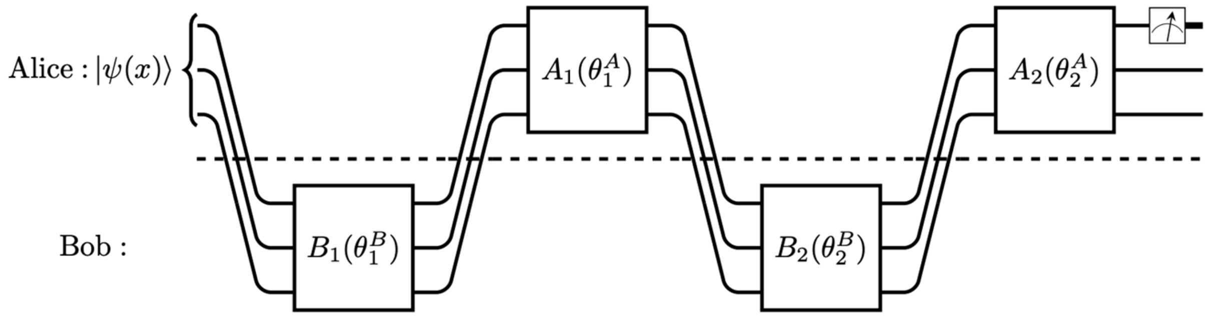
Problem 2 (Distributed Gradient Estimation). *Alice computes an estimate of $\nabla_A \langle\varphi| \mathcal{P}_0 |\varphi\rangle$, while Bob computes an estimate of $\nabla_B \langle\varphi| Z_0 |\varphi\rangle$, up to additive error ε in L^∞.*

Lemma 3. *i) The classical communication complexity of Problem 1 and Problem 2 is $\Omega(\max(\sqrt{N}, L))$.*

Lemma 1. *Problem 1 can be solved by communicating $O(\log N)$ qubits over $O(L/\varepsilon^2)$ rounds.*

Lemma 2. *Problem 2 can be solved with probability greater than $1-\delta$ by communicating $\tilde{O}(\log N(\log P)^2 \log(L/\delta)/\varepsilon^4)$ qubits over $O(L^2)$ rounds. The time and space complexity of the algorithm is $\sqrt{P}\ L\ \mathrm{poly}(N, \log P, \varepsilon^{-1}, \log(1/\delta))$.*

Model expressivity and privacy



$$|\varphi(\Theta, x)\rangle \equiv \left(\prod_{\ell=L}^{1} A_\ell(\theta_\ell^A, x) B_\ell(\theta_\ell^B, x)\right) |\psi(x)\rangle\,,$$

$$\mathcal{L}(\Theta, x) \equiv \langle \varphi(\Theta, x)| \mathcal{P}_0 |\varphi(\Theta, x)\rangle\,,$$

Lemma 9. *Let f be a p-times continuously differentiable function with period 1, and denote by $\hat{f}_{:M}$ the vector of the first M Fourier components of f. If $\left\|\hat{f}_{:M}\right\|_1 = 1$ then there exists a circuit of the form eq. (3.1) over $O(\log M)$ qubits such that*

$$\|\mathcal{L} - f\|_\infty \leq \frac{C}{M^{p-1/2}} \qquad \text{(F.4)}$$

Corollary 1. *If Alice and Bob are implementing the quantum algorithm for gradient estimation described in Lemma 2, and all the communication between Alice and Bob is intercepted by an attacker, the attacker cannot extract more than $\tilde{O}(L^2(\log N)^2(\log P)^2 \log(L/\delta)/\varepsilon^4)$ bits of classical information about the inputs to the players.*

Expressivity - a double edged sword

Lemma 9. *Let f be a p-times continuously differentiable function with period 1, and denote by $\hat{f}_{:M}$ the vector of the first M Fourier components of f. If $\left\|\hat{f}_{:M}\right\|_1 = 1$ then there exists a circuit of the form eq. (3.1) over $O(\log M)$ qubits such that*

$$\|\mathcal{L} - f\|_\infty \leq \frac{C}{M^{p-1/2}} \tag{F.4}$$

Recall, if we want this approximation converge for each feature dimension separately, we can use the relatively easy to prepare state |x> using Log(N) Log(M) qubits

$$\vec{x} \to |x\rangle = \sum_j^M \sum_k^N e^{-2\pi i j x^k / M} |j\rangle |k\rangle$$

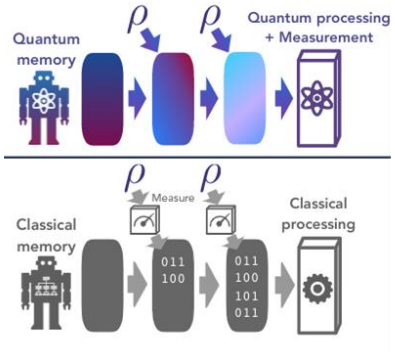
Only contains quadratic cross-feature terms, not complete across all N dimensions

Counting arguments tell you for this you need the multi-dimensional fourier state using N Log (M) qubits, which is the same complexity to just send the full state x classically → **no communication advantage**

Summary and outlook

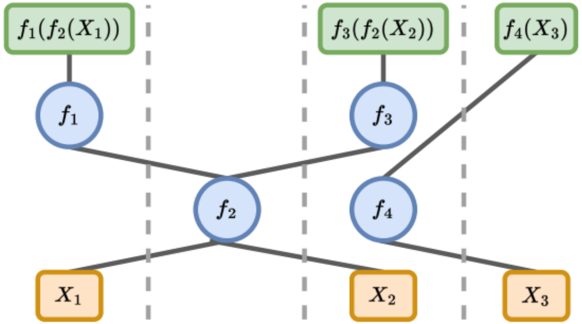
Quantum data

Punchline - Quantum devices today allow us to build classical impressive classical models, but fully understanding when a quantum model is required and what it accelerates remains an interesting open question



Classical data

Punchline(s) - Data changes the landscape of quantum advantage. If we can accept a future where stored quantum data states are stable and computers networked, we may find significant communication and privacy advantages in taking advantage of quantum encodings.

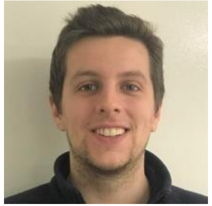


Acknowledgements

Bill Huggins

Dar Gilboa

Hsin-Yuan (Robert) Huang



Michael Broughton

Accelerating Quantum Algorithms with Precomputation
Huggins, McClean
arXiv:2305.09638 (2023)

Exponential Quantum Communication Advantage in Distributed Learning
Gilboa , McClean
arXiv:2310.07136 (2023)

Quantum advantage in learning from experiments
Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean
Science 376, 6598 (2022)

Quantum AI