Los Alamos
NATIONAL LABORATORY

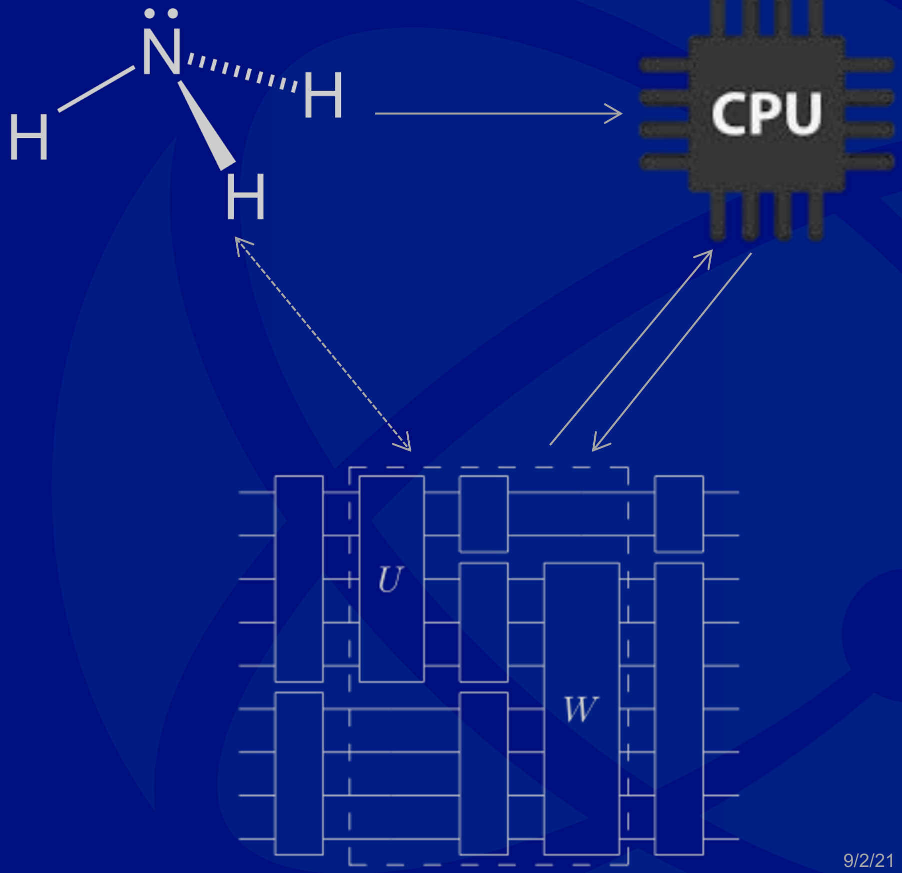
Learning Properties of Physical Systems from Quantum Circuit Data

Andi Gu, mentored by Andrew Arrasmith

Diego García-Martín, Marco Cerezo, Lukasz Cincio and Patrick Coles

Quantum Computer Summer School 2021

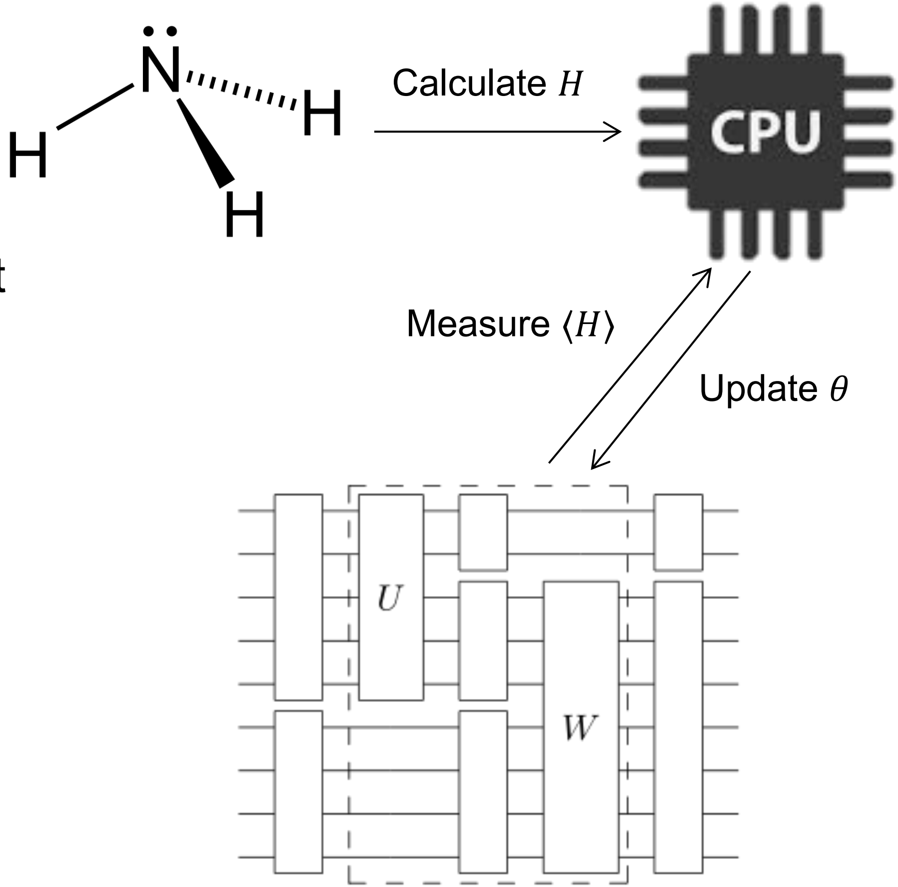
August 13, 2021



Given access to a physical system that we can write a Hamiltonian for, how can **efficiently** characterize system's ground state?

VQE

- Variational principle:
$$\langle\psi|H|\psi\rangle \geq E_0 \quad \forall\psi \in \mathcal{H}$$
- Encode ψ in a quantum circuit – fast evaluation of $C(\theta) \coloneqq \langle\psi(\theta)|H|\psi(\theta)\rangle$
- VQE: $E_0 \sim \min_{\theta} C(\theta)$

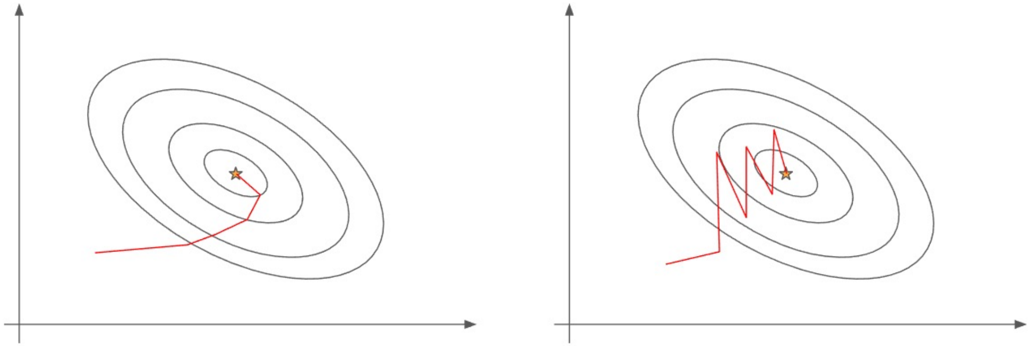


Minimizing $f(\theta)$

- Simple! Use gradient descent: $\theta^{(k+1)} = \theta^{(k)} - \alpha\nabla f$
- Shot noise means we only have access to a noisy estimate of $f(\theta)$

$$E[\mathcal{G}] = \left(\alpha - \frac{L\alpha^2}{2}\right)\|\nabla f\|^2 - \frac{L\alpha^2}{2}\sum \frac{\sigma_i^2}{s_i} = \textbf{Signal} - \textbf{Noise}$$

- How can we get the most bang for our buck?



Gradient Descent   Stochastic Gradient Descent

Los Alamos NATIONAL LABORATORY

Adaptive Shots

- Intuition: want few shots in the beginning, many shots at end
- Figure of merit:

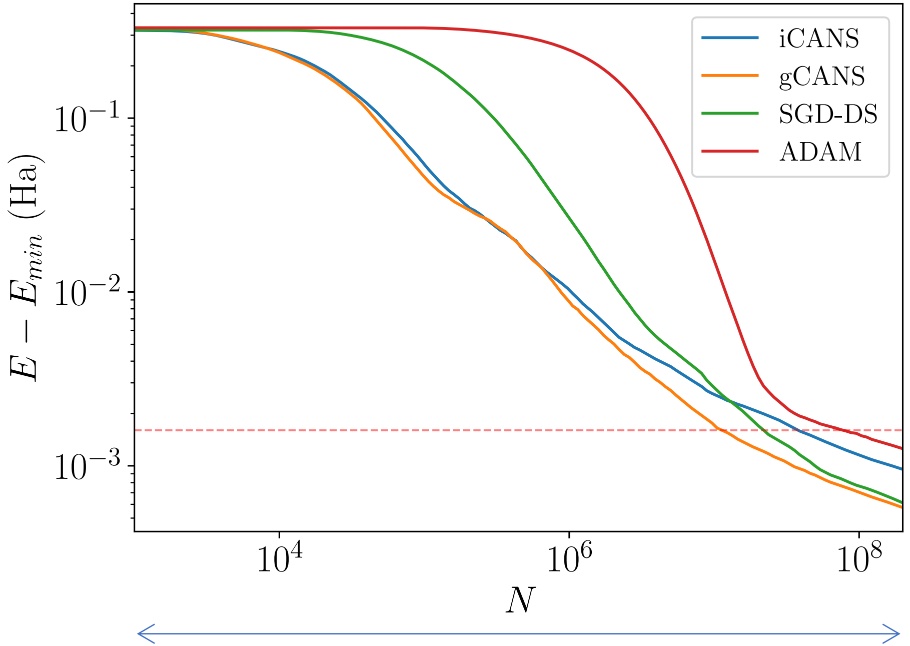
$$\gamma_i := \frac{E[\mathcal{G}_i]}{s_i} \quad \text{vs.} \quad \gamma := \frac{E[\mathcal{G}]}{\sum s_i}$$

Individual CANS vs. ***Global CANS***

(CANS = Coupled Adaptive Number of Shots)

- gCANS rule:

$$s_i = \frac{2L\alpha}{2 - L\alpha} \frac{\sigma_i \sum_j \sigma_j}{\|\nabla f\|_2^2}$$



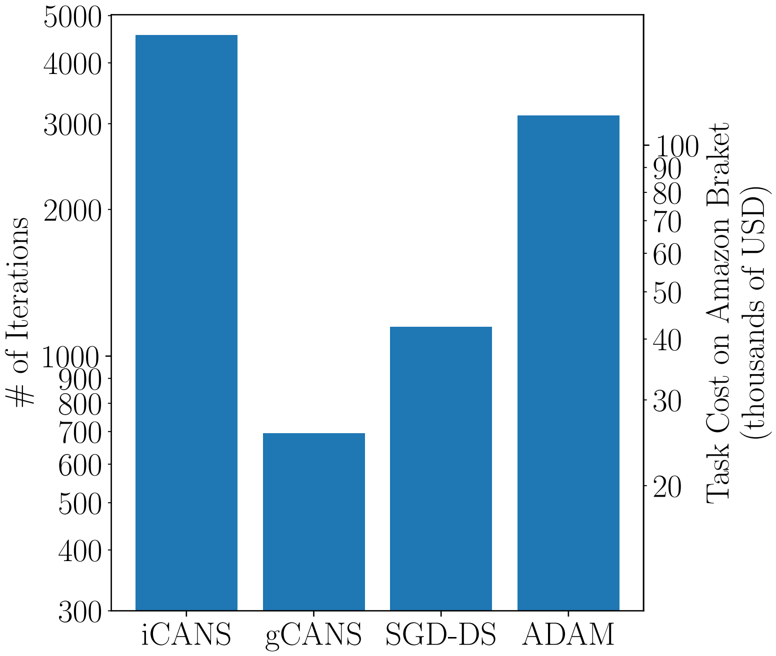
Far from optimum
Fewer shots

Close to optimum
More shots

Los Alamos NATIONAL LABORATORY

Why gCANS?

- Only one free hyperparameter
- Provable convergence guarantees: $f\left(\theta^{(k)}\right) - f^* \sim \epsilon^{-k}$
- Does better with higher learning rate → fewer iterations, fewer circuit compilations

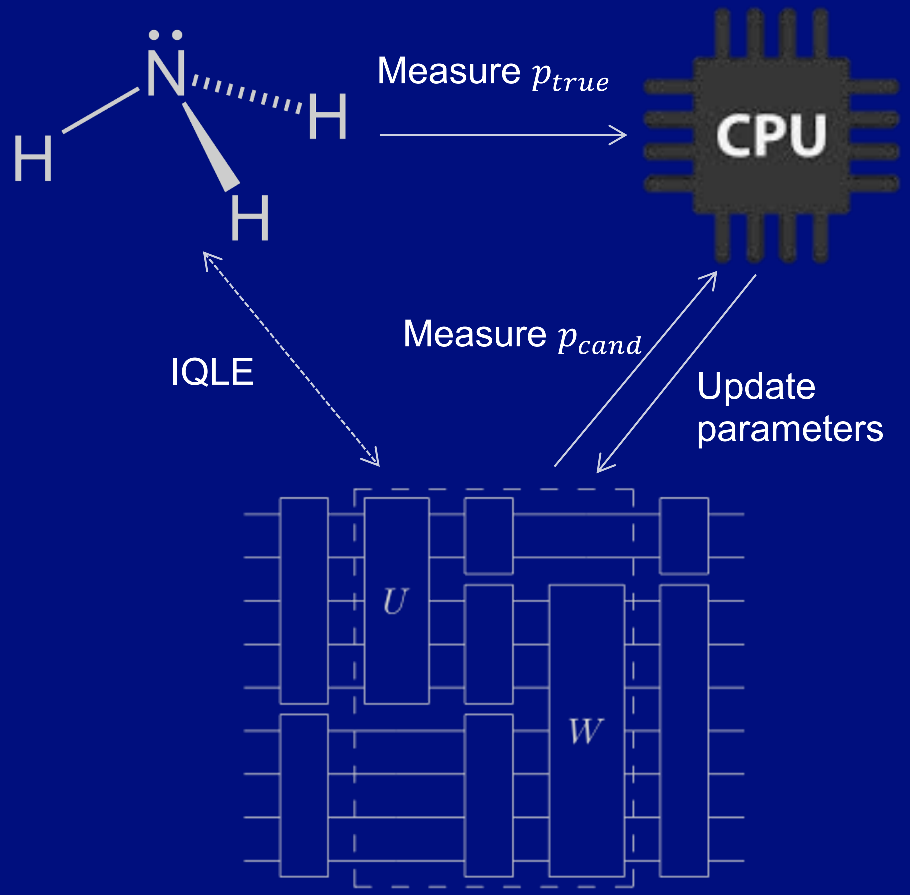


If we can observe the time evolution of a physical system, how can we **characterize** the **underlying Hamiltonian**?

Classical QLE

Idea: use a Trotterized quantum circuit as an ansatz $\widehat{H}(\theta)$ for true Hamiltonian H

1. Pick an initial state ψ_0, Pauli matrix P, evolution time t
2. Evolve $\psi_{true} = e^{-iHt}|\psi_0\rangle$ and $\psi_{cand} = e^{-i\widehat{H}t}|\psi_0\rangle$
3. Record measurement outcomes of P for ψ_{true} and ψ_{cand}
4. Optimize closeness of measurement outcomes



But is it scalable?

- Showed there is a barren plateau for classical flavor of Hamiltonian learning:

$$E_{\psi_0}[\mathcal{L}] \sim \frac{\left\| e^{iHt} e^{-i\hat{H}t} \right\|_F + C}{4^n}; \left\| e^{iHt} e^{-i\hat{H}t} \right\| \sim \mathcal{O}(2^n)$$

- Information extracted about system vanishes with 2^{-n}
- Origin of unfavorable scaling? $E_{\boldsymbol{\psi_0}}[\mathcal{L}]$ causes $\frac{1}{4^{\boldsymbol{n}}}$
- Intuition:

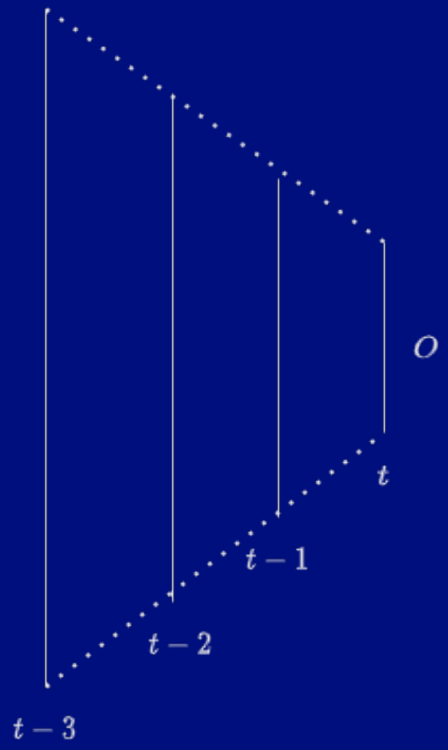
$$\langle \psi_{true} | P | \psi_{true} \rangle - \langle \psi_{cand} | P | \psi_{cand} \rangle \sim \mathcal{O}(N\epsilon)$$

but

$$\langle \psi_{true} | \psi_{cand} \rangle \sim \mathcal{O}(\epsilon^{-N})$$

How to fix QLE?

- Want to restrict ourself to a smaller subset of input states ψ_0
- Does this fail to constrain our learned Hamiltonian? Not for short times
- Study **subsystems** of the Hamiltonian



Next Steps

- For optimizers:
  - Can we design an adaptive optimizer that requires zero a priori knowledge (i.e., **dynamically set learning rate**)?
- For Hamiltonian learning:
  - Is there quantum advantage for short time evolution?
  - Can we circumvent the scaling problem with weak measurement?