Hamiltonian learning: recent progress and open problems

Yu Tong

Institute for Quantum Information and Matter, Caltech

October, 2023

This talk is based on

- Hsin-Yuan Huang, **Yu Tong**, Di Fang, Yuan Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

- Haoya Li, **Yu Tong**, Hongkang Ni, Tuvia Gefen, Lexing Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

Scope of the talk

- Learning the Hamiltonian from time-evolution, and with focus on the Heisenberg limit and the role of quantum control.

[1] Anshu, Arunachalam, Kuwahara, Soleimanifar, 2020, *Sample-efficient learning of interacting quantum systems*.

[2] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.

[3] Qi, Ranard, 2017, *Determining a local Hamiltonian from a single eigenstate*.

[4] Anshu, Arunachalam, 2023, *A survey on the complexity of learning quantum states*.

Scope of the talk

- Learning the Hamiltonian from time-evolution, and with focus on the Heisenberg limit and the role of quantum control.
- Not covered: learning the Hamiltonian from the Gibbs state or the ground state.[1,2,3,4]

[1] Anshu, Arunachalam, Kuwahara, Soleimanifar, 2020, *Sample-efficient learning of interacting quantum systems*.
[2] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.
[3] Qi, Ranard, 2017, *Determining a local Hamiltonian from a single eigenstate*.
[4] Anshu, Arunachalam, 2023, *A survey on the complexity of learning quantum states*.

The problem

We have an N-qubit quantum system evolving under a Hamiltonian H. We are allowed to interact with the system. The goal is to have a complete characterization of H classically. We may have some prior knowledge of H.

The problem

We have an N-qubit quantum system evolving under a Hamiltonian H. We are allowed to interact with the system. The goal is to have a complete characterization of H classically. We may have some prior knowledge of H.

- Interaction: we can prepare a (simple) initial state, apply (simple) unitaries during time evolution, and measure in some (simple) basis.

$$\text{Measure} \leftarrow U_r e^{-iHt_r} \cdots U_2 e^{-iHt_2} U_1 e^{-iHt_1} |\Phi\rangle .$$

The problem

We have an N-qubit quantum system evolving under a Hamiltonian H. We are allowed to interact with the system. The goal is to have a complete characterization of H classically. We may have some prior knowledge of H.

- Interaction: we can prepare a (simple) initial state, apply (simple) unitaries during time evolution, and measure in some (simple) basis.

$$\text{Measure} \leftarrow U_r e^{-iHt_r} \cdots U_2 e^{-iHt_2} U_1 e^{-iHt_1} |\Phi\rangle .$$

- Characterization: $H = \sum_{P\in\{I,X,Y,Z\}^{\otimes N}} \lambda_P P$. Want to learn all λ_P.

The problem

We have an N-qubit quantum system evolving under a Hamiltonian H. We are allowed to interact with the system. The goal is to have a complete characterization of H classically. We may have some prior knowledge of H.

- Interaction: we can prepare a (simple) initial state, apply (simple) unitaries during time evolution, and measure in some (simple) basis.

$$\text{Measure} \leftarrow U_r e^{-iHt_r} \cdots U_2 e^{-iHt_2} U_1 e^{-iHt_1} |\Phi\rangle .$$

- Characterization: $H = \sum_{P \in \{I,X,Y,Z\}^{\otimes N}} \lambda_P P$. Want to learn all λ_P.
- Prior knowledge: only a known $(\mathrm{poly}(N)$-sized) subset of λ_P's are non-zero, $|\lambda_P| \leq 1$. E.g., geometrically local.

The problem

We have an N-qubit quantum system evolving under a Hamiltonian H. We are allowed to interact with the system. The goal is to have a complete characterization of H classically. We may have some prior knowledge of H.

- Interaction: we can prepare a (simple) initial state, apply (simple) unitaries during time evolution, and measure in some (simple) basis.

$$\text{Measure} \leftarrow U_r e^{-iHt_r} \cdots U_2 e^{-iHt_2} U_1 e^{-iHt_1} \left|\Phi\right\rangle .$$

- Characterization: $H = \sum_{P \in \{I,X,Y,Z\}^{\otimes N}} \lambda_P P$. Want to learn all λ_P.
- Prior knowledge: only a known $(\mathrm{poly}(N)$-sized) subset of λ_P's are non-zero, $|\lambda_P| \leq 1$. E.g., geometrically local.
- Restriction: we cannot apply control-e^{-iHt} or e^{iHt}.

Measuring the cost

- We can get the Hamiltonian by learning the unitary $e^{-iH\tau}$ for a small τ. Requires $e^{\mathcal{O}(N)}\epsilon^{-1}$ queries to $e^{-iH\tau}$.[5]

[5] Haah, Kothari, O'Donnell, Tang, 2023, *Query-optimal estimation of unitary channels in diamond distance*.

Measuring the cost

- We can get the Hamiltonian by learning the unitary $e^{-iH\tau}$ for a small τ. Requires $e^{\mathcal{O}(N)}\epsilon^{-1}$ queries to $e^{-iH\tau}$.[5]
- But we want the "cost" to be at most $\mathrm{poly}(N)$.

[5] Haah, Kothari, O'Donnell, Tang, 2023, *Query-optimal estimation of unitary channels in diamond distance*.

Measuring the cost

- We can get the Hamiltonian by learning the unitary $e^{-iH\tau}$ for a small τ. Requires $e^{\mathcal{O}(N)}\epsilon^{-1}$ queries to $e^{-iH\tau}$.[5]
- But we want the "cost" to be at most $\mathrm{poly}(N)$.
- Need to define the cost.

[5] Haah, Kothari, O'Donnell, Tang, 2023, *Query-optimal estimation of unitary channels in diamond distance*.

Measuring the cost

- We can get the Hamiltonian by learning the unitary $e^{-iH\tau}$ for a small τ. Requires $e^{\mathcal{O}(N)}\epsilon^{-1}$ queries to $e^{-iH\tau}$.[5]
- But we want the "cost" to be at most $\mathrm{poly}(N)$.
- Need to define the cost.
  - Query complexity? $e^{-0.01iH}$ vs $e^{-1000iH}$.

[5] Haah, Kothari, O'Donnell, Tang, 2023, *Query-optimal estimation of unitary channels in diamond distance*.

Measuring the cost

- We can get the Hamiltonian by learning the unitary $e^{-iH\tau}$ for a small τ. Requires $e^{\mathcal{O}(N)}\epsilon^{-1}$ queries to $e^{-iH\tau}$.[5]

- But we want the "cost" to be at most $\mathrm{poly}(N)$.

- Need to define the cost.
  - Query complexity? $e^{-0.01iH}$ vs $e^{-1000iH}$.
  - We use total evolution time: if we use e^{-iHt_1}, e^{-iHt_2}, ..., $e^{-iHt_{N_{\mathrm{exp}}}}$, then the total evolution time is $t_1 + t_2 + \cdots t_{N_{\mathrm{exp}}}$.

[5] Haah, Kothari, O'Donnell, Tang, 2023, *Query-optimal estimation of unitary channels in diamond distance*.

Measuring the cost

- We can get the Hamiltonian by learning the unitary $e^{-iH\tau}$ for a small τ. Requires $e^{\mathcal{O}(N)}\epsilon^{-1}$ queries to $e^{-iH\tau}$.[5]
- But we want the "cost" to be at most $\mathrm{poly}(N)$.
- Need to define the cost.
  - Query complexity? $e^{-0.01iH}$ vs $e^{-1000iH}$.
  - We use total evolution time: if we use e^{-iHt_1}, e^{-iHt_2}, ..., $e^{-iHt_{N_{\mathrm{exp}}}}$, then the total evolution time is $t_1 + t_2 + \cdots t_{N_{\mathrm{exp}}}$.
  - We also need to make sure that the number of experiments N_{exp} and the number of unitaries are not too large.

[5] Haah, Kothari, O'Donnell, Tang, 2023, *Query-optimal estimation of unitary channels in diamond distance*.

Connection with quantum metrology

- **Quantum metrology:** high-precision estimation of a few physical parameters. Asymptotic convergence governed by the quantum Fisher information.

  ....

Connection with quantum metrology

- **Quantum metrology:** high-precision estimation of a few physical parameters. Asymptotic convergence governed by the quantum Fisher information.

  ....

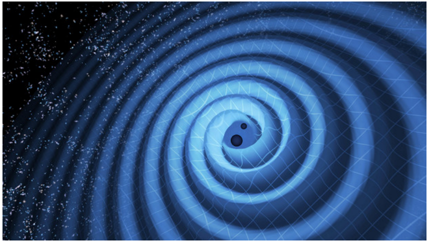


Figure: Image credit: LIGO/T. Pyle

Connection with quantum metrology

- **Quantum metrology:** high-precision estimation of a few physical parameters. Asymptotic convergence governed by the quantum Fisher information.

  ....

- **Hamiltonian learning:** Estimation of many parameters. Non-asymptotic (without good prior information).

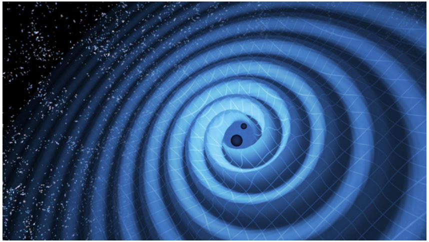


Figure: Image credit: LIGO/T. Pyle

A brief history

- **Heuristic algorithms** based on optimization and Bayesian inference.[6,7]

---

[6] Granade, Ferrie, Wiebe, Cory, 2012, *Robust online Hamiltonian learning*.

[7] Wiebe, Granade, Ferrie, Cory, 2014, *Hamiltonian Learning and Certification Using Quantum Resources*.

[8] Wang, Paesani, Santagati, et al., 2017, *Experimental quantum Hamiltonian learning*.

[9] Hangleiter, Roth, Eisert, Roushan, 2021, *Precise Hamiltonian identification of a superconducting quantum processor*.

A brief history

- **Heuristic algorithms** based on optimization and Bayesian inference.[6,7]
- **Experimental implementation:** single spin (NV center),[8] non-interacting boson (superconducting qubits).[9]

[6] Granade, Ferrie, Wiebe, Cory, 2012, *Robust online Hamiltonian learning*.

[7] Wiebe, Granade, Ferrie, Cory, 2014, *Hamiltonian Learning and Certification Using Quantum Resources*.

[8] Wang, Paesani, Santagati, et al., 2017, *Experimental quantum Hamiltonian learning*.

[9] Hangleiter, Roth, Eisert, Roushan, 2021, *Precise Hamiltonian identification of a superconducting quantum processor*.

Learning all ($\mathcal{O}(N)$) parameters to precision ϵ with probability at least $1-\delta$.

- **Provably efficient algorithms (perturbative):**

[10] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.

[11] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[12] Gu, Cincio, Coles, 2022, *Practical Black Box Hamiltonian Learning*.

[13] Yu, Sun, Han, Yuan, 2022, *Robust and Efficient Hamiltonian Learning*.

Learning all ($\mathcal{O}(N)$) parameters to precision ϵ with probability at least $1 - \delta$.

- **Provably efficient algorithms (perturbative):**
  - cluster expansion ($\mathcal{O}(\epsilon^{-2} \log(N/\delta))$),[10]

[10] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.

[11] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[12] Gu, Cincio, Coles, 2022, *Practical Black Box Hamiltonian Learning*.

[13] Yu, Sun, Han, Yuan, 2022, *Robust and Efficient Hamiltonian Learning*.

Learning all ($\mathcal{O}(N)$) parameters to precision ϵ with probability at least $1-\delta$.

- **Provably efficient algorithms (perturbative):**
  - cluster expansion ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$),[10]
  - derivative estimation ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$, Lindbladian),[11]

[10] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.

[11] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[12] Gu, Cincio, Coles, 2022, *Practical Black Box Hamiltonian Learning*.

[13] Yu, Sun, Han, Yuan, 2022, *Robust and Efficient Hamiltonian Learning*.

Learning all ($\mathcal{O}(N)$) parameters to precision ϵ with probability at least $1-\delta$.

- **Provably efficient algorithms (perturbative):**
  - cluster expansion ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$),[10]
  - derivative estimation ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$, Lindbladian),[11]
  - better scaling with degree ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$),[12]

---

[10] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.

[11] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[12] Gu, Cincio, Coles, 2022, *Practical Black Box Hamiltonian Learning*.

[13] Yu, Sun, Han, Yuan, 2022, *Robust and Efficient Hamiltonian Learning*.

Learning all ($\mathcal{O}(N)$) parameters to precision ϵ with probability at least $1-\delta$.

- **Provably efficient algorithms (perturbative):**
  - cluster expansion ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$),[10]
  - derivative estimation ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$, Lindbladian),[11]
  - better scaling with degree ($\mathcal{O}(\epsilon^{-2}\log(N/\delta))$),[12]
  - Pauli channel estimation ($\mathcal{O}(\epsilon^{-4}\log(N/\delta))$, SPAM-robust).[13]

---

[10] Haah, Kothari, Tang, 2021, *Optimal learning of quantum Hamiltonians from high-temperature Gibbs states*.

[11] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[12] Gu, Cincio, Coles, 2022, *Practical Black Box Hamiltonian Learning*.

[13] Yu, Sun, Han, Yuan, 2022, *Robust and Efficient Hamiltonian Learning*.

- **Provably efficient algorithms (Heisenberg limit):**
  - Hamiltonian reshaping with random Pauli operators ($\mathcal{O}(\epsilon^{-1}\log(N/\delta))$, SPAM-robust),[14]

[14] Huang, **Tong**, Fang, Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

[15] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

[16] Li, **Tong**, Ni, Gefen, Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

- **Provably efficient algorithms (Heisenberg limit):**
  - Hamiltonian reshaping with random Pauli operators ($\mathcal{O}(\epsilon^{-1}\log(N/\delta))$, SPAM-robust),[14]
  - Connection between quantum control and the Heisenberg limit,[15]

[14] Huang, **Tong**, Fang, Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

[15] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

[16] Li, **Tong**, Ni, Gefen, Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

- **Provably efficient algorithms (Heisenberg limit):**
  - Hamiltonian reshaping with random Pauli operators ($\mathcal{O}(\epsilon^{-1}\log(N/\delta))$, SPAM-robust),[14]
  - Connection between quantum control and the Heisenberg limit,[15]
  - Random gaussian unitaries ($\mathcal{O}(\epsilon^{-1}\log(N/\delta))$, boson).[16]

[14] Huang, **Tong**, Fang, Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

[15] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

[16] Li, **Tong**, Ni, Gefen, Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

The perturbative approach

- The Hamiltonian
$$H = \sum_{P \in \{I,X,Y,Z\}^{\otimes N}} \lambda_P P.$$

The perturbative approach

- The Hamiltonian

$$H = \sum_{P \in \{I,X,Y,Z\}^{\otimes N}} \lambda_P P.$$

- **Key observation:** e^{-iHt} is almost linear in H when t is small.

The perturbative approach

- The Hamiltonian

$$H = \sum_{P \in \{I,X,Y,Z\}^{\otimes N}} \lambda_P P.$$

- **Key observation:** e^{-iHt} is almost linear in H when t is small.

- Start from state ρ, evolve for time t, and measure observable O. The time derivative is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Tr}[\rho e^{iHt} O e^{-iHt}]|_{t=0} = i\mathrm{Tr}[\rho[H,O]] = i\mathrm{Tr}[H[O,\rho]].$$

The perturbative approach

- The Hamiltonian

$$H = \sum_{P \in \{I,X,Y,Z\}^{\otimes N}} \lambda_P P.$$

- **Key observation:** e^{-iHt} is almost linear in H when t is small.

- Start from state ρ, evolve for time t, and measure observable O. The time derivative is

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Tr}[\rho e^{iHt} O e^{-iHt}]|_{t=0} = i\mathrm{Tr}[\rho[H,O]] = i\mathrm{Tr}[H[O,\rho]].$$

- Choose ρ (Pauli eigenstate) and O (Pauli) so that $[O,\rho] = \frac{i}{2^{N-1}} P$.

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Tr}[\rho e^{iHt} O e^{-iHt}]|_{t=0} = -2\lambda_P.$$

- Derivatives can be estimated accurately using polynomial interpolation. Many derivatives can be estimated simultaneously using classical shadows.[17,18,19]

[17] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[18] Levy, Luo, Clark, 2021, *Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers*.

[19] Kunjummen, Tran, Carney, Taylor, 2021, *Shadow process tomography of quantum channels*.

- Derivatives can be estimated accurately using polynomial interpolation. Many derivatives can be estimated simultaneously using classical shadows.[17,18,19]

- Estimating $\mathrm{Tr}[\rho e^{iHt} O e^{-iHt}]$ through sampling and taking average. Error$\sim 1/\sqrt{N_s}$, where N_s is the number of samples.

[17] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[18] Levy, Luo, Clark, 2021, *Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers*.

[19] Kunjummen, Tran, Carney, Taylor, 2021, *Shadow process tomography of quantum channels*.

- Derivatives can be estimated accurately using polynomial interpolation. Many derivatives can be estimated simultaneously using classical shadows.[17,18,19]
- Estimating $\mathrm{Tr}[\rho e^{iHt} O e^{-iHt}]$ through sampling and taking average. Error$\sim 1/\sqrt{N_s}$, where N_s is the number of samples.
- Total evolution time $T \sim N_s$. $T = \mathcal{O}(\epsilon^{-2})$. The **standard quantum limit (SQL)**.

[17] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.

[18] Levy, Luo, Clark, 2021, *Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers*.

[19] Kunjummen, Tran, Carney, Taylor, 2021, *Shadow process tomography of quantum channels*.

- Derivatives can be estimated accurately using polynomial interpolation. Many derivatives can be estimated simultaneously using classical shadows.[17,18,19]
- Estimating $\mathrm{Tr}[\rho e^{iHt} O e^{-iHt}]$ through sampling and taking average. Error$\sim 1/\sqrt{N_s}$, where N_s is the number of samples.
- Total evolution time $T \sim N_s$. $T = \mathcal{O}(\epsilon^{-2})$. The **standard quantum limit (SQL)**.
- **The Heisenberg limit:** $T = \epsilon^{-1}$, and N_s can be $\mathcal{O}(\log(\epsilon^{-1}))$.

[17] Stilck-França, Markovich, Dobrovitski, 2022, *Efficient and robust estimation of many-qubit Hamiltonians*.
[18] Levy, Luo, Clark, 2021, *Classical Shadows for Quantum Process Tomography on Near-term Quantum Computers*.
[19] Kunjummen, Tran, Carney, Taylor, 2021, *Shadow process tomography of quantum channels*.

- The perturbative approach **cannot** achieve the Heisenberg limit.

- The perturbative approach **cannot** achieve the Heisenberg limit.
- With $t = \mathcal{O}(1)$,

- The perturbative approach **cannot** achieve the Heisenberg limit.
- With $t = \mathcal{O}(1)$,
  - Each experiment outcome distribution has Fisher information $\mathcal{O}(1)$.

- The perturbative approach **cannot** achieve the Heisenberg limit.
- With $t = \mathcal{O}(1)$,
  - Each experiment outcome distribution has Fisher information $\mathcal{O}(1)$.
  - Need the Fisher information of all experiments to be ϵ^{-2} (By Cramer-Rao bound).

- The perturbative approach **cannot** achieve the Heisenberg limit.
- With $t = \mathcal{O}(1)$,
  - Each experiment outcome distribution has Fisher information $\mathcal{O}(1)$.
  - Need the Fisher information of all experiments to be ϵ^{-2} (By Cramer-Rao bound).
- We need $\Omega(\epsilon^{-2})$ experiments to get to ϵ standard deviation (for non-adaptive and unbiased estimation).

- The perturbative approach **cannot** achieve the Heisenberg limit.
- With $t = \mathcal{O}(1)$,
  - Each experiment outcome distribution has Fisher information $\mathcal{O}(1)$.
  - Need the Fisher information of all experiments to be ϵ^{-2} (By Cramer-Rao bound).
- We need $\Omega(\epsilon^{-2})$ experiments to get to ϵ standard deviation (for non-adaptive and unbiased estimation).
- The proof can be extended to the adaptive and biased case.

- The perturbative approach **cannot** achieve the Heisenberg limit.
- With $t = \mathcal{O}(1)$,
  - Each experiment outcome distribution has Fisher information $\mathcal{O}(1)$.
  - Need the Fisher information of all experiments to be ϵ^{-2} (By Cramer-Rao bound).
- We need $\Omega(\epsilon^{-2})$ experiments to get to ϵ standard deviation (for non-adaptive and unbiased estimation).
- The proof can be extended to the adaptive and biased case.
- Reaching the Heisenberg limit requires something **qualitatively different**.

The Heisenberg limit: an example

Consider time-dependent signal $S(t)$, $t \geq 0$

$$S(t) = e^{i\theta t} + g, \quad g \sim \mathcal{N}(\mu, \sigma^2 I).$$

We want to estimate $\theta \in (-1, 1]$ to precision ϵ.

[20] Kimmel, Low, Yoder, 2015, *Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation*.

The Heisenberg limit: an example

Consider time-dependent signal $S(t)$, $t \geq 0$

$$S(t) = e^{i\theta t} + g, \quad g \sim \mathcal{N}(\mu, \sigma^2 I).$$

We want to estimate $\theta \in (-1, 1]$ to precision ϵ.

- We can let $t = \pi$, average out the noise, and estimate θ with $\mathcal{O}(\epsilon^{-2})$ samples.

[20] Kimmel, Low, Yoder, 2015, *Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation*.

The Heisenberg limit: an example

Consider time-dependent signal $S(t)$, $t \geq 0$

$$S(t) = e^{i\theta t} + g, \quad g \sim \mathcal{N}(\mu, \sigma^2 I).$$

We want to estimate $\theta \in (-1, 1]$ to precision ϵ.

- We can let $t = \pi$, average out the noise, and estimate θ with $\mathcal{O}(\epsilon^{-2})$ samples.
- I will outline a method that uses (ignoring the $\log\log$ factor)[20]
  1. $\mathcal{O}(\log(\epsilon^{-1}))$ samples,
  2. $\mathcal{O}(\epsilon^{-1})$ **total evolution time**.

[20] Kimmel, Low, Yoder, 2015, *Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation*.

The Heisenberg limit: an example

Consider time-dependent signal $S(t)$, $t \geq 0$

$$S(t) = e^{i\theta t} + g, \quad g \sim \mathcal{N}(\mu, \sigma^2 I).$$

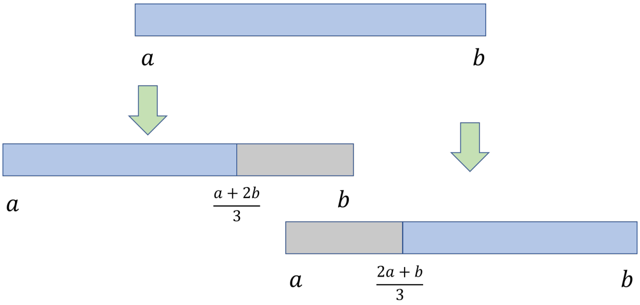
We want to estimate $\theta \in (-1, 1]$ to precision ϵ.

- We can let $t = \pi$, average out the noise, and estimate θ with $\mathcal{O}(\epsilon^{-2})$ samples.
- I will outline a method that uses (ignoring the $\log\log$ factor)[20]
  1. $\mathcal{O}(\log(\epsilon^{-1}))$ samples,
  2. $\mathcal{O}(\epsilon^{-1})$ **total evolution time**.
- Suppose our samples are $S(t_1), S(t_2), \cdots, S(t_{N_s})$, then the total evolution time is $t_1 + t_2 + \cdots + t_{N_s}$.

[20] Kimmel, Low, Yoder, 2015, *Robust Calibration of a Universal Single-Qubit Gate-Set via Robust Phase Estimation*.

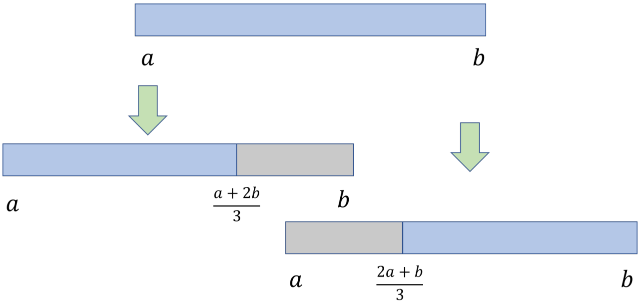
Suppose we know $a \leq \theta \leq b$. We want to determine

1. $a \leq \theta \leq \frac{a+2b}{3}$,
2. or $\frac{2a+b}{3} \leq \theta \leq b$.



Suppose we know $a \le \theta \le b$. We want to determine

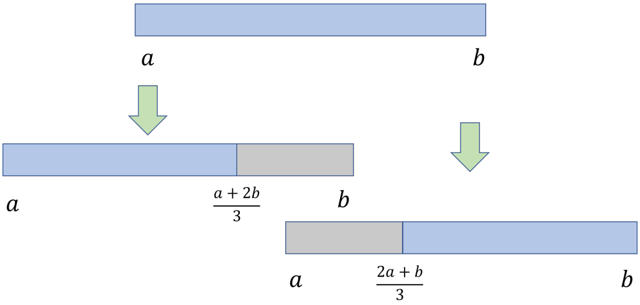
1. $a \le \theta \le \frac{a+2b}{3}$,
2. or $\frac{2a+b}{3} \le \theta \le b$.



- We can then update $a \leftarrow a$, $b \leftarrow (1/3)a + (2/3)b$, or $a \leftarrow (2/3)a + (1/3)b$, $b \leftarrow b$.

Suppose we know $a \leq \theta \leq b$. We want to determine

1. $a \leq \theta \leq \frac{a+2b}{3}$,
2. or $\frac{2a+b}{3} \leq \theta \leq b$.



- We can then update $a \leftarrow a$, $b \leftarrow (1/3)a + (2/3)b$, or $a \leftarrow (2/3)a + (1/3)b$, $b \leftarrow b$.
- We can **reduce the uncertainty by** $1/3$ at each step. $\mathcal{O}(\log(\epsilon^{-1}))$ steps are needed for ϵ precision.

We look at the value of

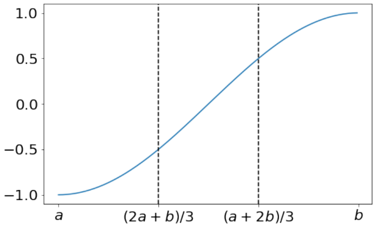
$$f_{a,b}(\theta) = \sin\left(\frac{\pi}{b-a}\left(\theta - \frac{a+b}{2}\right)\right) = \operatorname{Im}\left\langle S(t^*)\right\rangle e^{-i\frac{(a+b)\pi}{2(b-a)}},$$

where $t^* = \frac{\pi}{b-a}$.

We look at the value of

$$f_{a,b}(\theta) = \sin\left(\frac{\pi}{b-a}\left(\theta - \frac{a+b}{2}\right)\right) = \mathrm{Im}\left\langle S(t^*)\right\rangle e^{-i\frac{(a+b)\pi}{2(b-a)}},$$

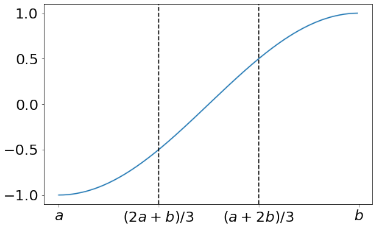
where $t^* = \frac{\pi}{b-a}$.



We look at the value of

$$f_{a,b}(\theta) = \sin\left(\frac{\pi}{b-a}\left(\theta - \frac{a+b}{2}\right)\right) = \mathrm{Im}\left\langle S(t^*)\right\rangle e^{-i\frac{(a+b)\pi}{2(b-a)}},$$

where $t^* = \frac{\pi}{b-a}$.

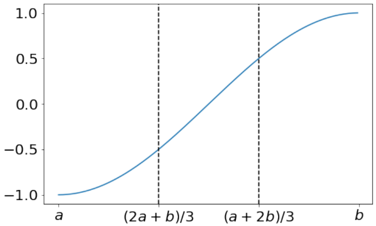


- If $f_{a,b}(\theta) \leq \frac{1}{2}$, then $a \leq \theta \leq \frac{a+2b}{3}$;

We look at the value of

$$f_{a,b}(\theta) = \sin\left(\frac{\pi}{b-a}\left(\theta - \frac{a+b}{2}\right)\right) = \mathrm{Im}\,\langle S(t^*)\rangle\, e^{-i\frac{(a+b)\pi}{2(b-a)}},$$

where $t^* = \frac{\pi}{b-a}$.

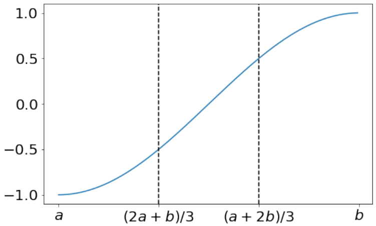


- If $f_{a,b}(\theta) \le \frac{1}{2}$, then $a \le \theta \le \frac{a+2b}{3}$;
- If $f_{a,b}(\theta) \ge -\frac{1}{2}$, then $\frac{2a+b}{3} \le \theta \le b$.

We look at the value of

$$f_{a,b}(\theta) = \sin\left(\frac{\pi}{b-a}\left(\theta - \frac{a+b}{2}\right)\right) = \operatorname{Im}\langle S(t^*)\rangle\, e^{-i\frac{(a+b)\pi}{2(b-a)}},$$

where $t^* = \frac{\pi}{b-a}$.

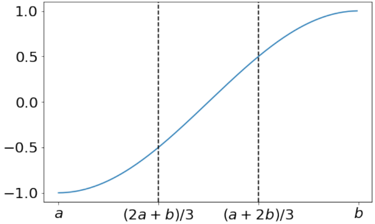


- If $f_{a,b}(\theta) \leq \frac{1}{2}$, then $a \leq \theta \leq \frac{a+2b}{3}$;
- If $f_{a,b}(\theta) \geq -\frac{1}{2}$, then $\frac{2a+b}{3} \leq \theta \leq b$.
- Evaluating $f_{a,b}(\theta)$ to precision $\frac{1}{2}$ is enough.

We look at the value of

$$f_{a,b}(\theta) = \sin\left(\frac{\pi}{b-a}\left(\theta - \frac{a+b}{2}\right)\right) = \operatorname{Im}\langle S(t^*)\rangle\, e^{-i\frac{(a+b)\pi}{2(b-a)}},$$

where $t^* = \frac{\pi}{b-a}$.



- If $f_{a,b}(\theta) \leq \frac{1}{2}$, then $a \leq \theta \leq \frac{a+2b}{3}$;
- If $f_{a,b}(\theta) \geq -\frac{1}{2}$, then $\frac{2a+b}{3} \leq \theta \leq b$.
- Evaluating $f_{a,b}(\theta)$ to precision $\frac{1}{2}$ is enough.
- Can get confidence level $1-\delta'$ with $\mathcal{O}(\log(\delta'^{-1}))$ samples.

- At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.

- At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.
- The cost of the last step is $\mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1}))$.

- At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.
- The cost of the last step is $\mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1}))$.
- The total cost is

$$\mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \times \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots\right) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})).$$

- At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.
- The cost of the last step is $\mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1}))$.
- The total cost is

$$\mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \times \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots\right) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})).$$

- Need $\delta' = \mathcal{O}(\delta / \log(\epsilon^{-1}))$ to ensure that all steps are successful with probability $1 - \delta$.

- At the last search step $b - a \approx (3/2)\epsilon$, and therefore $t^* \approx (2/3)\epsilon^{-1}$.
- The cost of the last step is $\mathcal{O}(t^* \log(\delta'^{-1})) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1}))$.
- The total cost is

$$\mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})) \times \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots\right) = \mathcal{O}(\epsilon^{-1} \log(\delta'^{-1})).$$

- Need $\delta' = \mathcal{O}(\delta / \log(\epsilon^{-1}))$ to ensure that all steps are successful with probability $1 - \delta$.
- Total evolution time is $\mathcal{O}(\epsilon^{-1} \log(\delta^{-1}))$ and the number of samples is $\mathcal{O}(\log(\epsilon^{-1}))$.
- Robust to noise ($|\mu| + \sigma = \mathcal{O}(1)$).

A single-qubit Hamiltonian

- We consider Hamiltonian $H = \theta Z$, and we want to learn the parameter θ from dynamics.

A single-qubit Hamiltonian

- We consider Hamiltonian $H = \theta Z$, and we want to learn the parameter θ from dynamics.
- We start from $|+\rangle$, evolve for time t, and measure in the X basis:

$$\langle + | e^{iHt} X e^{-iHt} | + \rangle = \cos(2\theta t).$$

A single-qubit Hamiltonian

- We consider Hamiltonian $H = \theta Z$, and we want to learn the parameter θ from dynamics.
- We start from $|+\rangle$, evolve for time t, and measure in the X basis:

$$\langle +| e^{iHt} X e^{-iHt} |+\rangle = \cos(2\theta t).$$

- Similarly when measure in the Y basis

$$\langle +| e^{iHt} Y e^{-iHt} |+\rangle = -\sin(2\theta t).$$

A single-qubit Hamiltonian

- We consider Hamiltonian $H = \theta Z$, and we want to learn the parameter θ from dynamics.
- We start from $|+\rangle$, evolve for time t, and measure in the X basis:

$$\langle +| e^{iHt} X e^{-iHt} |+\rangle = \cos(2\theta t).$$

- Similarly when measure in the Y basis

$$\langle +| e^{iHt} Y e^{-iHt} |+\rangle = -\sin(2\theta t).$$

- Combine to get a signal $e^{2i\theta t}+$ noise.

The difficulties of reaching the Heisenberg limit

- Reaching the Heisenberg limit requires long-time evolution.

[21] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

The difficulties of reaching the Heisenberg limit

- Reaching the Heisenberg limit requires long-time evolution.

- Many-body systems thermalize during the time evolution. For a local observable O:

$$\langle O(t)\rangle \approx \frac{1}{Z_\beta}\mathrm{Tr}[Oe^{-\beta H}].$$

[21] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

The difficulties of reaching the Heisenberg limit

- Reaching the Heisenberg limit requires long-time evolution.
- Many-body systems thermalize during the time evolution. For a local observable O:

$$\langle O(t) \rangle \approx \frac{1}{Z_\beta} \mathrm{Tr}[O e^{-\beta H}].$$

- Expectation values stop changing. Evolving for longer does not yield more information.

[21] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

The difficulties of reaching the Heisenberg limit

- Reaching the Heisenberg limit requires long-time evolution.
- Many-body systems thermalize during the time evolution. For a local observable O:
$$\langle O(t)\rangle \approx \frac{1}{Z_\beta}\mathrm{Tr}[Oe^{-\beta H}].$$
- Expectation values stop changing. Evolving for longer does not yield more information.
- Using non-local observables does not help either (under the eigenstate thermalization hypothesis and learning many parameters).[21]

[21] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

Creating conservation laws

- An abundance of local conservation laws can prevent thermalization (e.g., integrable models) or make it very slow (e.g., many-body localization).

Creating conservation laws

- An abundance of local conservation laws can prevent thermalization (e.g., integrable models) or make it very slow (e.g., many-body localization).
- If we can artificially create conservation laws we may use it to get coherent signal at late times.

Hamiltonian reshaping

- Inserting random Pauli operators.[22]

[22] Huang, **Tong**, Fang, Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

Hamiltonian reshaping

- Inserting random Pauli operators.[22]

$$e^{-iHt} = e^{-iH\tau} \cdots e^{-iH\tau} e^{-iH\tau} \to P_r e^{-iH\tau} P_r \cdots P_2 e^{-iH\tau} P_2 P_1 e^{-iH\tau} P_1,$$

where P_j are uniformly randomly drawn from a Pauli subgroup $K \leq G_N$.

[22] Huang, **Tong**, Fang, Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

Hamiltonian reshaping

- Inserting random Pauli operators.[22]

$$e^{-iHt} = e^{-iH\tau} \cdots e^{-iH\tau} e^{-iH\tau} \to P_r e^{-iH\tau} P_r \cdots P_2 e^{-iH\tau} P_2 P_1 e^{-iH\tau} P_1,$$

  where P_j are uniformly randomly drawn from a Pauli subgroup $K \leq G_N$.

- Because ${P_j}^2 = I$,

$$P_r e^{-iH\tau} P_r \cdots P_2 e^{-iH\tau} P_2 P_1 e^{-iH\tau} P_1 = e^{-iP_r H P_r \tau} \cdots e^{-iP_2 H P_2 \tau} e^{-iP_1 H P_1 \tau}.$$

[22] Huang, **Tong**, Fang, Su, 2022, *Learning many-body Hamiltonians with Heisenberg-limited scaling*.

In one time step

$$\rho \mapsto \rho - i\mathbb{E}_{P\sim\mathcal{U}(K)}[PHP, \rho]\tau + \mathcal{O}(\tau^2)$$
$$= \rho - i[H_{\text{effective}}, \rho]\tau + \mathcal{O}(\tau^2),$$

where

$$H_{\text{effective}} = \mathbb{E}_{P\sim\mathcal{U}(K)} PHP = \frac{1}{|K|}\sum_{P\in K} PHP,$$

$$e^{-iHt} \mapsto e^{-iH_{\text{effective}}t}.$$

[23] Campbell, 2018, *A random compiler for fast Hamiltonian simulation*.

In one time step

$$\begin{aligned}\rho \mapsto \rho - i\mathbb{E}_{P\sim\mathcal{U}(K)}[PHP, \rho]\tau + \mathcal{O}(\tau^2) \\ = \rho - i[H_{\text{effective}}, \rho]\tau + \mathcal{O}(\tau^2),\end{aligned}$$

where

$$H_{\text{effective}} = \mathbb{E}_{P\sim\mathcal{U}(K)} PHP = \frac{1}{|K|}\sum_{P\in K} PHP,$$

$$e^{-iHt} \mapsto e^{-iH_{\text{effective}}t}.$$

- This is the same idea underlying the qDRIFT algorithm.[23]

[23] Campbell, 2018, *A random compiler for fast Hamiltonian simulation*.

- The Hamiltonian is transformed through

$$H \mapsto H_{\text{effective}} = \frac{1}{|K|} \sum_{P \in K} P H P.$$

- The Hamiltonian is transformed through

$$H \mapsto H_{\text{effective}} = \frac{1}{|K|} \sum_{P \in K} PHP.$$

- **Every element in K is a conservation law in $H_{\text{effective}}$.** For $Q \in K$,

$$QH_{\text{effective}}Q = \frac{1}{|K|} \sum_{P \in K} QPHPQ = H_{\text{effective}} \implies [Q, H_{\text{effective}}] = 0.$$

- The Hamiltonian is transformed through

$$H \mapsto H_{\text{effective}} = \frac{1}{|K|} \sum_{P \in K} P H P.$$

- **Every element in K is a conservation law in $H_{\text{effective}}$.** For $Q \in K$,

$$Q H_{\text{effective}} Q = \frac{1}{|K|} \sum_{P \in K} Q P H P Q = H_{\text{effective}} \implies [Q, H_{\text{effective}}] = 0.$$

- **The coefficients we want to learn are preserved.** For any Pauli operator $P' \in G_N$,

$$\frac{1}{|K|} \sum_{P \in K} P P' P = \begin{cases} P', & P' \in C_{G_N}(K), \\ 0, & P' \notin C_{G_N}(K). \end{cases} \implies H_{\text{effective}} = \sum_{P \in C_{G_N}(K)} \lambda_P P.$$

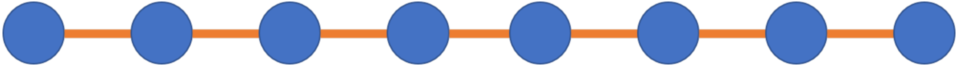


Figure: Every qubit interacts with its neighbors.

- Choose $K = \langle Z_3, Z_6, Z_9, \cdots, X_3, X_6, X_9, \cdots \rangle$.

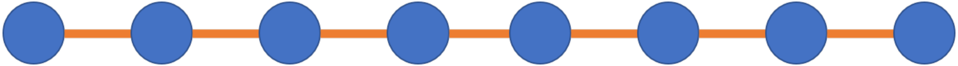


Figure: Every qubit interacts with its neighbors.

- Choose $K = \langle Z_3, Z_6, Z_9, \cdots, X_3, X_6, X_9, \cdots \rangle$.
- $P \in C_{G_N}(K)$ only when it acts trivially on qubits $3, 6, 9, \cdots$.

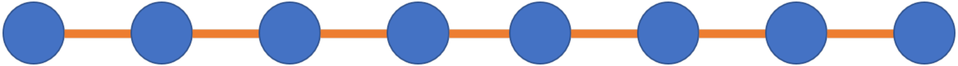


Figure: Every qubit interacts with its neighbors.

- Choose $K = \langle Z_3, Z_6, Z_9, \cdots, X_3, X_6, X_9, \cdots \rangle$.
- $P \in C_{G_N}(K)$ only when it acts trivially on qubits $3, 6, 9, \cdots$.
- If H has only nearest-neighbor interaction, then the system will be **decoupled**.

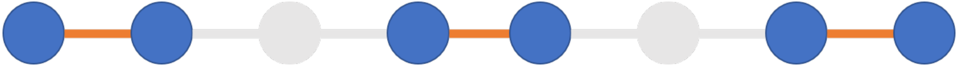


Figure: Suppressing qubits so that the rest are isolated.

- Choose $K = \langle Z_3, Z_6, Z_9, \cdots, X_3, X_6, X_9, \cdots \rangle$.
- $P \in C_{G_N}(K)$ only when it acts trivially on qubits $3, 6, 9, \cdots$.
- If H has only nearest-neighbor interaction, then the system will be **decoupled**.

- We can also use this approach to make the effective Hamiltonian diagonal in a certain basis (e.g., let $\langle X_1, X_2, X_3, \cdots \rangle \subset K$).

[24] Greene, Kjaergaard, Schwartz, et al., 2021, *Error mitigation via stabilizer measurement emulation*.

- We can also use this approach to make the effective Hamiltonian diagonal in a certain basis (e.g., let $\langle X_1, X_2, X_3, \cdots \rangle \subset K$).

- We use conservation laws to decouple the system into non-interacting clusters, each evolving under a Hamiltonian that is diagonal w.r.t a known basis.

[24] Greene, Kjaergaard, Schwartz, et al., 2021, *Error mitigation via stabilizer measurement emulation*.

- We can also use this approach to make the effective Hamiltonian diagonal in a certain basis (e.g., let $\langle X_1, X_2, X_3, \cdots \rangle \subset K$).
- We use conservation laws to decouple the system into non-interacting clusters, each evolving under a Hamiltonian that is diagonal w.r.t a known basis.
- The Hamiltonian coefficients are preserved in the process.

[24] Greene, Kjaergaard, Schwartz, et al., 2021, *Error mitigation via stabilizer measurement emulation*.

- We can also use this approach to make the effective Hamiltonian diagonal in a certain basis (e.g., let $\langle X_1, X_2, X_3, \cdots \rangle \subset K$).

- We use conservation laws to decouple the system into non-interacting clusters, each evolving under a Hamiltonian that is diagonal w.r.t a known basis.

- The Hamiltonian coefficients are preserved in the process.

- Can be generalized to all bounded-degree local Hamiltonians (each term involve $\mathcal{O}(1)$ qubits, and each qubit is involved in $\mathcal{O}(1)$ terms).

[24] Greene, Kjaergaard, Schwartz, et al., 2021, *Error mitigation via stabilizer measurement emulation*.

- We can also use this approach to make the effective Hamiltonian diagonal in a certain basis (e.g., let $\langle X_1, X_2, X_3, \cdots \rangle \subset K$).

- We use conservation laws to decouple the system into non-interacting clusters, each evolving under a Hamiltonian that is diagonal w.r.t a known basis.

- The Hamiltonian coefficients are preserved in the process.

- Can be generalized to all bounded-degree local Hamiltonians (each term involve $\mathcal{O}(1)$ qubits, and each qubit is involved in $\mathcal{O}(1)$ terms).

- Close connection to dynamical decoupling, but more versatile.

[24] Greene, Kjaergaard, Schwartz, et al., 2021, *Error mitigation via stabilizer measurement emulation*.

- We can also use this approach to make the effective Hamiltonian diagonal in a certain basis (e.g., let $\langle X_1, X_2, X_3, \cdots \rangle \subset K$).
- We use conservation laws to decouple the system into non-interacting clusters, each evolving under a Hamiltonian that is diagonal w.r.t a known basis.
- The Hamiltonian coefficients are preserved in the process.
- Can be generalized to all bounded-degree local Hamiltonians (each term involve $\mathcal{O}(1)$ qubits, and each qubit is involved in $\mathcal{O}(1)$ terms).
- Close connection to dynamical decoupling, but more versatile.
- Similar subgroup-based strategy can be used to suppress coherent errors in quantum circuits.[24]

---

[24] Greene, Kjaergaard, Schwartz, et al., 2021, *Error mitigation via stabilizer measurement emulation*.

Based on the Hamiltonian reshaping technique, we propose a Hamiltonian learning protocol that

- Achieves the Heisenberg scaling with $\mathcal{O}(\epsilon^{-1}\log(N/\delta))$ total evolution time;
- Uses $\mathcal{O}(\text{polylog}(\epsilon^{-1})\log(N/\delta))$ experiments;
- Uses only single-qubit Pauli eigenstates, Pauli gates, and single-qubit measurements;
- Is robust against state preparation and measurement (SPAM) error.

Hamiltonian reshaping for bosons

- Let H be a bosonic Hamiltonian, e.g.

$$H = \sum_{\langle i,j \rangle} h_{ij} b_i^\dagger b_j + \sum_i \omega_i n_i + \frac{1}{2} \sum_i \xi_i n_i (n_i - 1),$$

where $b_i^\dagger$ (b_i) are the bosonic creation (annihilation) operators.

[25] Li, **Tong**, Ni, Gefen, Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

Hamiltonian reshaping for bosons

- Let H be a bosonic Hamiltonian, e.g.

$$H = \sum_{\langle i,j \rangle} h_{ij} b_i^\dagger b_j + \sum_i \omega_i n_i + \frac{1}{2} \sum_i \xi_i n_i (n_i - 1),$$

where $b_i^\dagger$ (b_i) are the bosonic creation (annihilation) operators.

- We can apply $e^{i\theta n_i}$ (phase shifter) for $\theta \sim \mathcal{U}([0, 2\pi])$ to enforce local particle number conservation (U(1) symmetry).[25]

[25] Li, **Tong**, Ni, Gefen, Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

Hamiltonian reshaping for bosons

- Let H be a bosonic Hamiltonian, e.g.

$$H = \sum_{\langle i,j \rangle} h_{ij} b_i^\dagger b_j + \sum_i \omega_i n_i + \frac{1}{2} \sum_i \xi_i n_i (n_i - 1),$$

where $b_i^\dagger$ (b_i) are the bosonic creation (annihilation) operators.

- We can apply $e^{i\theta n_i}$ (phase shifter) for $\theta \sim \mathcal{U}([0, 2\pi])$ to enforce local particle number conservation (U(1) symmetry).[25]

- This can be used to isolate parts of the quantum system (no particle can hop to or from site i).

[25] Li, **Tong**, Ni, Gefen, Ying, 2023, *Heisenberg-limited Hamiltonian learning for interacting bosons*.

- Learning off-diagonal terms: apply $e^{i\theta(b_i^\dagger b_j + b_j^\dagger b_i)/2}$ (beam splitter) for $\theta \sim \mathcal{U}([0, 2\pi])$ to conserve $b_i^\dagger b_j + b_j^\dagger b_i$ (similarly for $ib_i^\dagger b_j - ib_j^\dagger b_i$).

- Learning off-diagonal terms: apply $e^{i\theta(b_i^\dagger b_j + b_j^\dagger b_i)/2}$ (beam splitter) for $\theta \sim \mathcal{U}([0, 2\pi])$ to conserve $b_i^\dagger b_j + b_j^\dagger b_i$ (similarly for $ib_i^\dagger b_j - ib_j^\dagger b_i$).

- These $b_i^\dagger b_j + b_j^\dagger b_i$ and $ib_i^\dagger b_j - ib_j^\dagger b_i$ can be made diagonal if we change the single-particle basis.

- Learning off-diagonal terms: apply $e^{i\theta(b_i^\dagger b_j + b_j^\dagger b_i)/2}$ (beam splitter) for $\theta \sim \mathcal{U}([0, 2\pi])$ to conserve $b_i^\dagger b_j + b_j^\dagger b_i$ (similarly for $ib_i^\dagger b_j - ib_j^\dagger b_i$).
- These $b_i^\dagger b_j + b_j^\dagger b_i$ and $ib_i^\dagger b_j - ib_j^\dagger b_i$ can be made diagonal if we change the single-particle basis.
- Based on this, we propose a protocol for learning the Bose-Hubbard-type Hamiltonian with

- Learning off-diagonal terms: apply $e^{i\theta(b_i^\dagger b_j + b_j^\dagger b_i)/2}$ (beam splitter) for $\theta \sim \mathcal{U}([0, 2\pi])$ to conserve $b_i^\dagger b_j + b_j^\dagger b_i$ (similarly for $ib_i^\dagger b_j - ib_j^\dagger b_i$).

- These $b_i^\dagger b_j + b_j^\dagger b_i$ and $ib_i^\dagger b_j - ib_j^\dagger b_i$ can be made diagonal if we change the single-particle basis.

- Based on this, we propose a protocol for learning the Bose-Hubbard-type Hamiltonian with
  - Achieves the Heisenberg scaling with $\mathcal{O}(\epsilon^{-1}\log(N/\delta))$ total evolution time;
  - Uses $\mathcal{O}(\text{polylog}(\epsilon^{-1})\log(N/\delta))$ experiments;
  - Uses only coherent states, random one- or two-mode gaussian unitaries, and homodyne measurements;
  - Is robust against state preparation and measurement (SPAM) error.

Open problems

- Quantum control is necessary, but "how much" control do we need?
  - $e^{-iH\tau} \approx I - iH\tau \mapsto I - iH_{\text{effective}}\tau$: error of order $\mathcal{O}(\tau^2)$.

[26] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

Open problems

- Quantum control is necessary, but "how much" control do we need?
  - $e^{-iH\tau} \approx I - iH\tau \mapsto I - iH_{\text{effective}}\tau$: error of order $\mathcal{O}(\tau^2)$.
  - To reach ϵ accuracy, we need $\tau = \Theta(\epsilon)$. Apply Pauli unitaries very fast.

---

[26] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

Open problems

- Quantum control is necessary, but "how much" control do we need?
  - $e^{-iH\tau} \approx I - iH\tau \mapsto I - iH_{\text{effective}}\tau$: error of order $\mathcal{O}(\tau^2)$.
  - To reach ϵ accuracy, we need $\tau = \Theta(\epsilon)$. Apply Pauli unitaries very fast.
  - Can use 2nd-order Trotter to get $\tau = \Theta(\epsilon^{1/2})$. Higher order requires evolving backward in time.

[26] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

Open problems

- Quantum control is necessary, but "how much" control do we need?
  - $e^{-iH\tau} \approx I - iH\tau \mapsto I - iH_{\text{effective}}\tau$: error of order $\mathcal{O}(\tau^2)$.
  - To reach ϵ accuracy, we need $\tau = \Theta(\epsilon)$. Apply Pauli unitaries very fast.
  - Can use 2nd-order Trotter to get $\tau = \Theta(\epsilon^{1/2})$. Higher order requires evolving backward in time.
  - Evolving up to time T, we need at least $\Omega(T)$ unitaries to be inserted,[26] corresponding to $\tau = \mathcal{O}(1)$.

[26] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

Open problems

- Quantum control is necessary, but "how much" control do we need?
  - $e^{-iH\tau} \approx I - iH\tau \mapsto I - iH_{\text{effective}}\tau$: error of order $\mathcal{O}(\tau^2)$.
  - To reach ϵ accuracy, we need $\tau = \Theta(\epsilon)$. Apply Pauli unitaries very fast.
  - Can use 2nd-order Trotter to get $\tau = \Theta(\epsilon^{1/2})$. Higher order requires evolving backward in time.
  - Evolving up to time T, we need at least $\Omega(T)$ unitaries to be inserted,[26] corresponding to $\tau = \mathcal{O}(1)$.
  - Can we design a protocol to achieve this scaling? Apply unitaries with only constant frequency.

[26] Dutkiewicz, O'Brien, Schuster, 2023, *The advantage of quantum control in many-body Hamiltonian learning*.

- Can we tolerate error during time evolution (other than SPAM)?

[27] Zhou, Zhang, Preskill, Jiang, 2018, *Achieving the Heisenberg limit in quantum metrology using quantum error correction*.

- Can we tolerate error during time evolution (other than SPAM)?
  - Quantum noise will make signal decay, preventing us from reaching the Heisenberg limit.

[27] Zhou, Zhang, Preskill, Jiang, 2018, *Achieving the Heisenberg limit in quantum metrology using quantum error correction*.

- Can we tolerate error during time evolution (other than SPAM)?
  - Quantum noise will make signal decay, preventing us from reaching the Heisenberg limit.
  - Can quantum error correction (QEC) help?

[27] Zhou, Zhang, Preskill, Jiang, 2018, *Achieving the Heisenberg limit in quantum metrology using quantum error correction*.

- Can we tolerate error during time evolution (other than SPAM)?
  - Quantum noise will make signal decay, preventing us from reaching the Heisenberg limit.
  - Can quantum error correction (QEC) help?
  - Only certain Hamiltonian terms can benefit from QEC (Hamiltonian-not-in-Lindblad-span (HNLS) condition).[27]

[27] Zhou, Zhang, Preskill, Jiang, 2018, *Achieving the Heisenberg limit in quantum metrology using quantum error correction*.

- Can we tolerate error during time evolution (other than SPAM)?
  - Quantum noise will make signal decay, preventing us from reaching the Heisenberg limit.
  - Can quantum error correction (QEC) help?
  - Only certain Hamiltonian terms can benefit from QEC (Hamiltonian-not-in-Lindblad-span (HNLS) condition).[27]
  - For terms not in the Lindblad span, can we design a non-asymptotic protocol to learn all of them scalably in the presence of quantum noise?

[27] Zhou, Zhang, Preskill, Jiang, 2018, *Achieving the Heisenberg limit in quantum metrology using quantum error correction*.

Conclusion

- Hamiltonian learning in the Heisenberg limit requires long-time evolution.

Conclusion

- Hamiltonian learning in the Heisenberg limit requires long-time evolution.
- We need control to artificially create conservation laws to put off thermalization.

Conclusion

- Hamiltonian learning in the Heisenberg limit requires long-time evolution.
- We need control to artificially create conservation laws to put off thermalization.
- Open questions remain as to how fast and strong the control needs to be and tolerance of quantum noise.