Estimating Quantum Hamiltonians via Joint Measurements of Noisy Non-Commuting Observables

Dan McNulty

Center for Theoretical Physics,
Polish Academy of Sciences, Warsaw, Poland

III R. S. Ingarden Memorial Session, 30th November 2022

Joint work with **Filip Maciejewski** and **Michał Oszmaniec**

(arXiv:2206.08912)

Outline

- Introduction and motivation
- Estimating Hamiltonians via joint measurements
- Connections to classical shadows
- Estimating in the presence of physical noise

Motivation: NISQ computing

- Variational quantum algorithms:[1] quantum speedups on near-term quantum computers
- Estimate energies of quantum many-body Hamiltonians (e.g. of molecules)

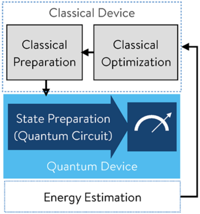


Figure: Variational quantum algorithm for energy estimation with classical optimisation [2]

- The Hamiltonian, to be measured on a quantum computer, can be expressed as

$$H = \sum_{P \in \mathbb{P}_n} \lambda_P P$$

where $\mathbb{P}_n = \{ P = \otimes_{i=1}^n P_i \mid P_i \in \{\mathbb{1}, X, Y, Z\} \}$ and $\lambda_P \in \mathbb{R}$
- Requires measuring non-commuting Pauli strings, i.e., $\mathrm{tr}\,[H\rho] = \sum_{P \in \mathbb{P}_n} \lambda_P \mathrm{tr}\,[P\rho]$

[1] K. Bharti et al, *Rev. Mod. Phys.* **94** 015004 (2022).

[2] B. Bauer et al, *Chem. Rev.* **120**, 12685 (2020).

Existing strategies

Estimating the expectation value of

$$H = \sum_{P \in \mathbb{P}_n} \lambda_P P$$

Two main approaches

- Grouping observables into compatible sets (maps to minimum clique cover problem[3])

$$H = \lambda_1 X \otimes X + \lambda_2 X \otimes \mathbb{1} + \lambda_3 Z \otimes Y + \lambda_4 Y \otimes Z + \lambda_5 \mathbb{1} \otimes Y$$

- Classical shadows [4]

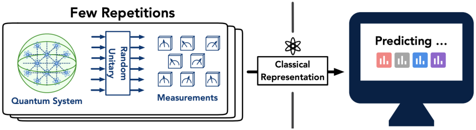


Figure: A classical representation of the state is built using randomised Pauli or Clifford measurements (Fig. reproduced from Ref. 4)

Our approach: Joint measurability

[3] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, *J. Chem. Phys.* **152**, 124114 (2020).

[4] H.-Y. Huang, R. Kueng, and J. Preskill, *Nat. Phys.* **16**, 1050 (2020).

Joint measurability

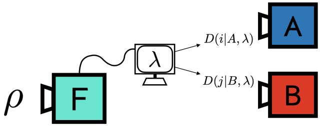
- A POVM M is a collection of PSD operators $\mathrm{M}(i)$ such that $\sum_i \mathrm{M}(i) = \mathbb{1}$

Definition

Two POVMs $\mathrm{A}(i)$ and $\mathrm{B}(j)$ are jointly measurable if they can be obtained from a POVM $\mathrm{F}(\lambda)$ via a stochastic transformation (classical post-processing $D(.|.,\lambda)$),

$$\mathrm{A}(i) = \sum_{\lambda} D(i|\mathrm{A},\lambda)\mathrm{F}(\lambda) \quad \text{and} \quad \mathrm{B}(j) = \sum_{\lambda} D(j|\mathrm{B},\lambda)\mathrm{F}(\lambda)$$

where $0 \leq D(i|A,\lambda) \leq 1$ and $\sum_i D(i|A,\lambda) = 1$.



- Projective measurements: joint measurability $\iff$ commutativity
- For POVMs, joint measurability $\not\Rightarrow$ commutativity
- Non-commuting observables can be measured simultaneously by adding noise

Joint measurability of Pauli observables

- Consider the three non-commuting (qubit) Pauli observables X, Y and Z.
- Their noisy (unsharp) versions

$$\widetilde{\mathrm{M}}_X(\pm) = \frac{1}{2}(\mathbb{1} \pm \eta^x X)\,, \quad \widetilde{\mathrm{M}}_Y(\pm) = \frac{1}{2}(\mathbb{1} \pm \eta^y Y)\,, \quad \widetilde{\mathrm{M}}_Z(\pm) = \frac{1}{2}(\mathbb{1} \pm \eta^z Z)$$

are jointly measurable if [5] and only if [6] $(\eta^x)^2 + (\eta^y)^2 + (\eta^z)^2 \leq 1\,.$

Joint measurement

A parent POVM G is given by

$$\mathrm{G}(x, y, z) = \frac{1}{8}(\mathbb{1} + x\eta^x X + y\eta^y Y + z\eta^z Z)\,,$$

with outcomes $x, y, z \in \{\pm 1\}$.

- $\sum_{z,y} \mathrm{G}(x,y,z) = \frac{1}{8}\sum_{y,z}(\mathbb{1} + x\eta^x X + y\eta^y Y + z\eta^z Z) = \frac{1}{2}(\mathbb{1} + x\eta^x X) = \widetilde{\mathrm{M}}_X(x)$

[5] P. Busch, *Phys. Rev. D* **33**, 2253 (1986).

[6] T. Brougham and E. Andersson, *Phys. Rev. A* **76**, 052313 (2007).

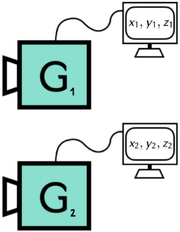
Joint measurability of two-qubit Paulis

Two-qubit parent POVM

$$\begin{aligned}\mathsf{F}(x_1, \ldots, z_2) &= \mathsf{G}_1(x_1, y_1, z_1) \otimes \mathsf{G}_2(x_2, y_2, z_2) \\ &= \frac{1}{8^2}(\mathbb{1} + x_1\eta_1^x X + y_1\eta_1^y Y + z_1\eta_1^z Z) \otimes (\mathbb{1} + x_2\eta_2^x X + y_2\eta_2^y Y + z_2\eta_2^z Z)\end{aligned}$$

Classical post-processing for $X \otimes Y$

Measure $X \otimes Y$, with outcome $\mathbf{s} = x_1 \cdot y_2$ and noise $\eta_1^x \cdot \eta_2^y$,



$$\sum_{\substack{x_1, \ldots, z_2 \\ x_1 \cdot y_2 = \mathbf{s}}} \mathsf{G}_1 \otimes \mathsf{G}_2 = \ldots = \frac{1}{2}(\mathbb{1} + \mathbf{s} \cdot \eta_1^x \eta_2^y X \otimes Y)$$

Joint measurability of Pauli strings

Joint measurement of Pauli strings

A locally biased joint measurement on n-qubits is given by

$$\mathrm{F}(\mathbf{x}_1, \dots, \mathbf{x}_n) := \bigotimes_{i=1}^{n} \mathrm{G}_i(x_i, y_i, z_i)\,,$$

where $\mathrm{G}_i(x_i, y_i, z_i) = \frac{1}{8}(\mathbb{1} + x_i\eta_i^x X + y_i\eta_i^y Y + z_i\eta_i^z Z)$ and $\mathbf{x}_i = (x_i, y_i, z_i)\,.$

- F is a joint measurement of all noisy (unsharp) Pauli strings

  $$\widetilde{\mathrm{M}}_P(s_P) = \frac{1}{2}(\mathbb{1} + s_P\eta_P P)\,,$$

  where $s_P \in \{\pm 1\}$ is the product of local outcomes and η_P the product of local noises
- For example, if $P = X \otimes \mathbb{1} \otimes Z$, then $s_P = x_1 z_3$ and $\eta_P = \eta_1^x \eta_3^z\,.$

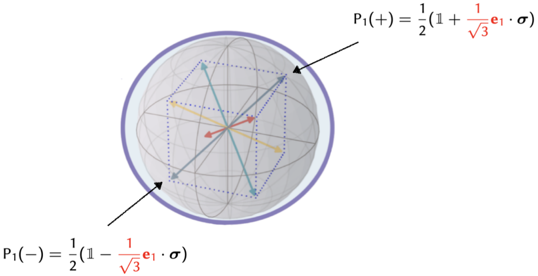
Implementing the joint measurement on a quantum computer

Projective simulability

A POVM is projective simulable if it can be implemented via randomisation of projective measurements & classical post-processing[a]

[a] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, *Phys. Rev. Lett.* **119**, 190501 (2017).

- POVM $\mathrm{G}(x, y, z) = \frac{1}{8}(\mathbb{1} + \frac{1}{\sqrt{3}}(xX + yY + zZ))$, with uniform noise, is simulated by a uniform mixture of four projective measurements onto opposite vertices of a cube



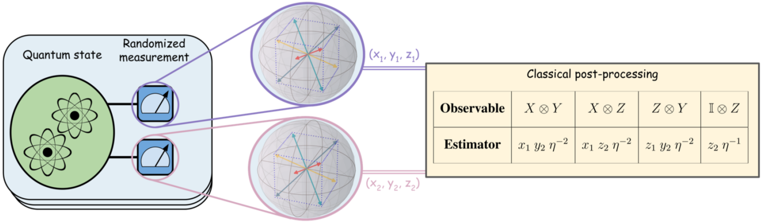
with $\mathbf{e}_1 = (1, 1, 1)$, $\mathbf{e}_2 = (1, 1, -1)$, $\mathbf{e}_3 = (1, -1, 1)$, $\mathbf{e}_4 = (-1, 1, 1)$, and $\boldsymbol{\sigma} = (X, Y, Z)$

- Implementing G corresponds to measuring P_j with probability $p(j) = \frac{1}{4}$
- For example, measuring P_1 and obtaining outcome $+1$ corresponds to $(1, 1, 1)$

Estimating Hamiltonians via joint measurability

Estimating Pauli strings

Estimate, simultaneously, $\mathrm{tr}\,[P\rho]$ for all $P = \otimes_{i=1}^{n} P_i$, where $P_i \in \{\mathbb{1}, X, Y, Z\}$



1. Perform the joint measurement G on each qubit system, obtaining an outcome tuple (x_i, y_i, z_i) for every qubit, where $x_i, y_i, z_i \in \{\pm 1\}$.
2. The outcome of the noisy version of $P = \otimes_{i=1}^{n} P_i$ is the product of local outcomes p_i (equal to either x_i, y_i or z_i corresponding to P_i).
3. An unbiased estimator $\hat{P}$ of $\mathrm{tr}\,[P\rho]$ is obtained by dividing $\prod_i p_i$ by the product of the local noises.

Estimating Hamiltonians

Unbiased estimator of $\mathrm{tr}\,[H\rho]$

For an n-qubit Hamiltonian $H = \sum_P \lambda_P P$, with $\lambda_P \in \mathbb{R}$, a single shot estimator is given by

$$\hat{H} = \sum_P \frac{1}{\eta_P} \lambda_P s_P$$

where s_P is the outcome associated with the unsharp measurement $\widetilde{M}_P(s_P) = \frac{1}{2}(\mathbb{1} + \eta_P s_P P)$

- $\hat{H}$ is unbiased, i.e., $\mathbb{E}[\hat{H}] = \mathrm{tr}\,[H\rho]$
- The number of copies of ρ such that $\mathrm{Prob}(|\hat{H} - \mathrm{tr}\,[H\rho]\,| < \epsilon) > 1 - \delta$, is

$$N = \mathcal{O}\left(\frac{\log(1/\delta)}{\epsilon^2} Var[\hat{H}]\right) ,$$

with $Var[\hat{H}] = \mathbb{E}[\hat{H}^2] - (\mathrm{tr}\,[H\rho])^2$, and $\mathbb{E}[\hat{H}^2] = \sum_{P,Q \in \mathbb{P}_n} \frac{\lambda_P \lambda_Q}{\eta_P \eta_Q} \mathbb{E}[s_P s_Q]$

Variance of estimator

- $\mathbb{E}[s_P s_Q] = \sum_{s_P, s_Q} \mathrm{tr}\left[\widetilde{\mathrm{M}}_{P,Q}(s_P, s_Q)\rho\right] s_P s_Q$
- If P and Q qubit-wise commute, then $\mathbb{E}[s_P s_Q] = \frac{\eta_{PQ}}{\eta_P \eta_Q} \mathrm{tr}\,[PQ\rho]$
- Otherwise, $\mathbb{E}[s_P s_Q] = 0$

Variance of $\hat{H}$

The variance of the estimator $\hat{H}$ is given by

$$Var[\hat{H}] = \sum_{P,Q \in \mathbb{P}_n} \frac{\eta_{PQ} f(P,Q)}{\eta_P \eta_Q} \lambda_P \lambda_Q \mathrm{tr}\,[PQ\rho] - (\mathrm{tr}\,[H\rho])^2\,,$$

where $f(P,Q) = \prod_{i=1}^{n} f_i(P,Q)$, and

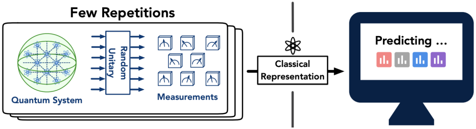
$$f_i(P,Q) = \begin{cases} 1 & \text{if } \; P_i = \mathbb{1} \; \text{or} \; Q_i = \mathbb{1} \; \text{or} \; P_i = Q_i\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Optimisation

- Noise parameters η_{PQ}, η_P and η_Q of parent POVM can be optimised to minimize $Var[\hat{H}]$

Connections to classical shadows

Classical shadows



- Randomised measurement protocol[7], sampling $U \in \mathcal{U}$, such that

$$\rho \longrightarrow U\rho U^\dagger \longrightarrow |b\rangle\langle b| \longrightarrow U^\dagger |b\rangle\langle b| U$$

- In expectation, the protocol defines a **shadow channel**

$$\mathcal{M} : \rho \longmapsto \mathbb{E}_{U\in\mathcal{U}} \sum_{b\in\{0,1\}^n} p(b|\rho) U^\dagger |b\rangle\langle b| U$$

- Applying the inverse channel $\mathcal{M}^{-1}$ gives a **classical snapshot**

$$\hat{\rho} = \mathcal{M}^{-1}(U^\dagger |b\rangle\langle b| U)$$

- $\hat{\rho}$ is not necessarily positive semidefinite but

$$\mathbb{E}[\hat{\rho}] = \mathcal{M}^{-1}(\mathbb{E}[U^\dagger |b\rangle\langle b| U]) = \mathcal{M}^{-1}(\mathcal{M}(\rho)) = \rho$$

[7] H.-Y. Huang, R. Kueng, and J. Preskill, *Nat. Phys.* **16**, 1050 (2020).

Estimating Hamiltonians with classical shadows

- The classical shadow for random Pauli measurements (i.e. X, Y and Z on each qubit), is

$$\hat{\rho} = \otimes_{i=1}^{n} (3 U_i^\dagger |b_i\rangle\langle b_i| U_i - \mathbb{1})$$

- An arbitrary set of m observables $O_1, \ldots, O_m$, can be estimated simultaneously via

$$\hat{O}_j^{sh} = \mathrm{tr}\left[O_j \hat{\rho}\right] \qquad j = 1, \ldots, m$$

  with variance bounded by the shadow norm, $Var[\hat{O}^{sh}] \leq \left\|O_j\right\|_{\text{shadow}}^2$
- For Hamiltonian $H = \sum_P \lambda_P P$, estimator of $\mathrm{tr}\left[H\rho\right]$ is

$$\hat{H}^{sh} = \sum_P \lambda_P \mathrm{tr}\left[P\hat{\rho}\right]$$

- Locally biased classical shadow[8]: X, Y and Z are sampled from a (biased) probability distribution $p(P_i|i)$ where $P_i \in \{X, Y, Z\}$ for each qubit i

Observation 1

The joint measurability (JM) and locally biased classical shadow (LBCS) estimation protocols (for Hamiltonians $H = \sum_P \lambda_P P$) have the same sample complexity bounds

---

[8] C. Hadfield, S. Bravyi, R. Raymond, and A. Mezzacapo, *Commun. Math. Phys.* **391**, 951 (2022).

From joint measurements to classical shadows

Observation 2

From the joint measurement F we can construct a locally biased classical shadow. Restricting to the unbiased setting we recover a shadow with similar form to Huang *et al.* [a]

[a] H.-Y. Huang, R. Kueng, and J. Preskill, *Nat. Phys.* **16**, 1050 (2020)

- Single shot classical approximation of the quantum state ρ is given by

$$\hat{\rho}^{\,\mathrm{JM}} = \bigotimes_{i=1}^{n} \frac{1}{2}(\mathbb{1} + \mathbf{e}_i \cdot \boldsymbol{\sigma}) = \bigotimes_{i=1}^{n} \left(\|\mathbf{e}_i\| \rho_{\tilde{\mathbf{e}}_i} + \frac{\mathbb{1}}{2}(1 - \|\mathbf{e}_i\|) \right) ,$$

  where $\mathbf{e}_i = (x_i/\eta_i^x, y_i/\eta_i^y, z_i/\eta_i^z)$ and $\rho_{\tilde{\mathbf{e}}_i} = \frac{1}{2}(\mathbb{1} + \tilde{\mathbf{e}}_i \cdot \boldsymbol{\sigma})$, with $\tilde{\mathbf{e}}_i = \mathbf{e}_i / \|\mathbf{e}_i\|$
- If $\eta_i^x = \eta_i^y = \eta_i^z = \frac{1}{\sqrt{3}}$, then $\|\mathbf{e}_i\| = 3$, and $\hat{\rho}^{\,\mathrm{JM}} = \bigotimes_{i=1}^{n}(3\rho_{\tilde{\mathbf{e}}_i} - \mathbb{1})$
- This has a similar form to $\hat{\rho}^{\,\mathrm{CS}} = \bigotimes_{i=1}^{n}(3U_i^\dagger |b_i\rangle\langle b_i| U_i - \mathbb{1})$

From classical shadows to joint measurements

Observation 3

Any classical shadow defines a joint measurement and provides a sufficient condition for the compatibility of an arbitrary set of measurements

- Protocol describes a single POVM, $\mathrm{G}(x, U) = \frac{1}{|\mathcal{U}|} U^\dagger |x\rangle\langle x| U$, where $U \in \mathcal{U}$, $x \in \{0, 1\}^n$.
- Produces a snapshot $\hat{\rho}_{x,U} = (d+1) U^\dagger |x\rangle\langle x| U - \mathbb{1}$
- $\hat{\rho}_{x,U}$ is not necessarily positive semidefinite, but $\mathrm{tr}\,[\hat{\rho}_{x,U}] = 1$ and $\mathbb{E}[\hat{\rho}_{x,U}] = \rho$
- For a set of POVMs $\mathrm{M}_j(s)$, we can compute

  $$q(s|j, x, U) = \mathrm{tr}\left[\mathrm{M}_j(s)\hat{\rho}_{x,U}\right]$$

  which, in expectation, yields the outcome statistics of M_j
- Add noise such that $\mathrm{tr}\left[\mathrm{M}_j^\eta(s)\hat{\rho}_{x,U}\right] \geq 0$, where $\mathrm{M}_j^\eta = \eta \mathrm{M}_j + (1-\eta)\mathrm{tr}\left[\mathrm{M}_j\right] \mathbb{1}/d$
- Joint measurability of M_j^η holds for $\eta \leq \frac{1}{d+1}$

Estimating in the presence of physical noise

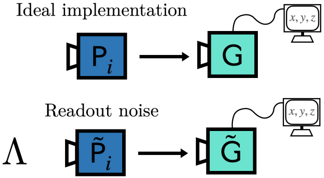
Incorporating readout noise

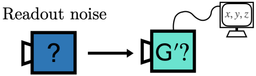
Readout noise

Ideal projective measurement P affected by stochastic readout noise, i.e.,

$$\widetilde{\mathrm{P}}(j) = \sum_k \Lambda_{jk} \mathrm{P}(k)\,.$$

Modified parent POVM $\widetilde{\mathrm{G}}$, implemented via randomisation of $\widetilde{\mathrm{P}}$, may no longer be optimal





- Which POVMs are noisy projective simulable?
- What is the optimal G′ that is noisy projective simulable?

Numerics

- Evaluate $Var[\hat{H}]$ of Hamiltonian H using noisy projective simulable POVM G'
- Optimise G' on each qubit to minimise $Var[\hat{H}]$
- Compare with readout noise in classical shadows [9,10]

Joint measurability vs noisy classical shadows

Encoding / Molecule	H_2	LiH	BeH_2	H_2O
Jordan-Wigner	1.00	0.06	0.04	0.1
Bravyi-Kitaev	0.13	0.78	0.55	0.61
Parity	0.38	0.02	0.009	0.02

Table: Upper bounds on the variance of the estimators of Hamiltonians in the presence of readout noise, normalised by classical shadows

- Optimised strategies allow us to obtain a reduction of the variance upper bound by as much as a factor of ≈ 100

[9] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, *PRX Quantum* **2**, 030348 (2021).

[10] D. E. Koh and S. Grewal, *Quantum* **6**, 776 (2022).

Concluding remarks

- Are there deeper fundamental connections between shadows and joint measurements?
- Can we gain further insight into the efficiency of computational tasks from the limits of joint measurability?
- Can we construct optimal joint measurements from the performance limits of classical shadows?
- Is joint measurability a useful strategy in other quantum computing applications?
- Motivates further studies of incompatibility, e.g. characterisation of optimal joint measurements which are projective (or noisy projective) simulable
- **Future work**: Joint measurability strategies in fermionic systems