Classical Shadow

Henry Lin

June 2025

1 Introduction

We have two ways to reconstruct an unknown quantum state now:

- Full quantum state tomography
- Classical Shadow

The first tool is full quantum state tomography reconstruct the full density matrix of the unknown quantum state, but it is very costly. The second tool is classical shadow which reconstruct the expectations and uses approximations.

2 Classical Shadow Pipline

Let us consider an unknown n-qubit quantum state ρ and apply a random unitary to the state:

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

The procedure of the classical shadow is the following:

- apply an random unitary
- measurement
- apply the inverse of the random unitary
- invert the measurement channel which obtain ρ'

If we apply the above procedure N times, then we will obtain the following(classical shadow):

$$S(\rho, N) = \{\hat{\rho_1} = \mathcal{M}^{-1}(U_1^\dagger |b_1\rangle\langle b_1| U_1), \ldots, \hat{\rho_N} = \mathcal{M}^{-1}(U_N^\dagger |b_N\rangle\langle b_N| U_N)\}$$

Note, the inverted channel is not physical. Now if we want to calculate the expectation of an observable which is the following:

$$< o_i >= \mathrm{Tr}(o_i\rho)$$

We can replace ρ by the estimator $\hat{\rho}$.

$$< o_i >= \mathrm{Tr}(o_i\hat{\rho}_i)$$

3 Notes

The measurement bases are choosen randomly and are choosen from the following set: $\{I, X, Y, Z\}$. Since each Pauli basis is unitary, the eigenvalues of it are lie on the unit circle of a complex plane $e^{i\theta}$. The measurement outcomes are either $+1$ or -1 which are associated with the eigenvalues.

4 Estimating observables with classical shadows in the Pauli basis

We only care about the case that all the Pauli operators matched with the observables. This implies that computing expectation values with classical shadows comes down to picking the specific subset of snapshots where those specific observables were already measured and discarding the remaining. If the desired observables are known prior to the measurement, one is thus advised to directly perform those measurements. This was referred to as derandomization

5 Measurement problems

The number of measurements that are needed for the quantum hardware is still grows polynomially for algorithms like VQE and QAOA. To mitigate this measurement problem, people can group the communiting terms of the Hamiltonian to significantely reduce the number of measurements. In some case, the reduction is up to 90%. The number of terms in the Hamiltonian scales like $O(N^4)$. Our goal is to reduce the number of measurements needed while maintaining a good accuracy.

Let's start with the Heisenberg uncertainty relationship for two observables $\hat{A}$ and $\hat{B}$, we know that

$$\sigma_A^2 \sigma_B^2 \geq |\langle [\hat{A}, \hat{B}] \rangle|$$

If A and B don't commute, then we cannot simultaneously measure the expectation values of the two observables.

If A and B do commute, then there exists a measurement basis where we can simultaneously measure the expectation value of both observables on the same state.

If A and B do commute, then there exists a complete orthnormal basis set $|\phi_n\rangle$ that simultaneously diagonalizes both A and B. This means:

$$A |\phi_n\rangle = \lambda_{A,n} |\phi_n\rangle$$
$$B |\phi_n\rangle = \lambda_{B,n} |\phi_n\rangle$$

or we can write it like this:

$$A = PD_AP^{-1}$$
$$B = PD_BP^{-1}$$

This means that they shared a same set of eigenbases, but with different eigenvalues.

Proof:

$$AB\left|\psi_n\right\rangle = \lambda_{A,n}\lambda_{B,n}\left|\psi_n\right\rangle$$
$$BA\left|\psi_n\right\rangle = \lambda_{A,n}\lambda_{B,n}\left|\psi_n\right\rangle$$
$$(AB - BA)\left|\psi_n\right\rangle = [A, B]\left|\psi_n\right\rangle = 0$$

This means $AB\left|\psi_n\right\rangle = BA\left|\psi_n\right\rangle$. Let say we have operator A and B, and use spectral decomposition:

$$A = \sum_i \lambda_{A,n}\left|\psi_n\right\rangle\left\langle\psi_n\right|$$
$$B = \sum_i \lambda_{B,n}\left|\psi_n\right\rangle\left\langle\psi_n\right|$$

Let's calculate $\left\langle\psi\right| A\left|\psi\right\rangle$:

$$\begin{aligned}\left\langle\psi\right| A\left|\psi\right\rangle &= \left\langle\psi\right|\sum_i \lambda_{A,n}\left|\psi_n\right\rangle\left\langle\psi_n|\psi\right\rangle \\ &= \sum_i \lambda_{A,n}\left|\left\langle\psi_n|\psi\right\rangle\right|^2 \\ \left\langle\psi\right| B\left|\psi\right\rangle &= \left\langle\psi\right|\sum_i \lambda_{B,n}\left|\psi_n\right\rangle\left\langle\psi_n|\psi\right\rangle \\ &= \sum_i \lambda_{B,n}\left|\left\langle\psi_n|\psi\right\rangle\right|^2\end{aligned}$$

This we can recover the expected value of A and B by a single measurement. If we combine every single $\left|\left\langle\psi_n|\psi\right\rangle\right|^2$ together, it will form a distribution. First rotated every single qubit into the Z-basis representation, then this will create a diagonal matrix for each of the Pauli string, then our eigenbasis is basically the computational basis. Each λ is pre-determined and easy to calculate, then weight each probability by it and sum them to recover the expected value.

We can use technique such as qubit-wise commutativity(QWC) which means we only looking at individual Pauli word in a Pauli string to determine if they commute or not. If individual Pauli word commutes then the two strings are

commute. This is sufficient but not necessary condition to determine if two operators are commute or anti-commute.

Let's say you have a Hamiltonian, then we can use Pauli decomposition to express it in terms of the Pauli strings. Next, we can use linearity of expectation to recover the expected value of the Hamiltonian. However, to measure every single Pauli string is a naive way to do this. We can determine the full-commutativity or QWS to reduce the number of Pauli strings that are needed for measurements.

6 Denormalization

The expected value for a observable for the tth classical shadow is:

$$\langle \bigotimes_i O_i \rangle^{(t)} = 3^q \prod (1 - 2b_i^{(t)})$$

q is the number of local gates.
The expectated value for the Pauli string is

$$\langle \bigotimes_i O_i \rangle = \frac{1}{\tilde{T}} \prod (1 - 2b_i^{(t)})$$

$\tilde{T} = \frac{T}{3^q}$.

This implies that computing expectation values with classical shadows comes down to picking the specific subset of snapshots where those specific observables were already measured and discarding the remaining. **If the desired observables are known prior to the measurement, one is thus advised to directly perform those measurements.**

7 Applications of classical shadow

- quantum state fidelity
- expectation values of Hamiltonians
- entanglement witnesses
- two-point correlators